Boletín de la

SOCIEDAD MATEMÁTICA MEXICANA

Tercera Serie

Volumen 10

Número Especial

2004

Homenaje a
Francisco González Acuña

SOCIEDAD
MATEMATICA
MEXICANA

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

A CUT-AND-PASTE APPROACH TO CONTACT TOPOLOGY

WILLIAM H. KAZEZ

Abstract. This expository paper gives an introduction to some of the techniques used to study tight contact structures on 3-manifolds. The goal is to develop cut-and-paste techniques that are analogous to Haken and sutured manifold decompositions. Many examples and sketches of ideas behind some of the main theorems are given.

Contents

1. Convex Surfaces

Unless otherwise stated, M denotes a compact oriented 3-dimensional manifold that may have nonempty boundary. Throughout this paper we assume that all manifolds and submanifolds are oriented.

Definition (1.1). A (positive) *contact structure*, ξ on M, is a smooth 2-plane bundle $\xi_p \subset TM$ such that there exists a $1-$form α, satisfying
1. $\ker_p(\alpha) = \xi_p$ for all $p \in M$ and
2. $\alpha \wedge d\alpha > 0$.

Example (1.2). Figure 1 shows a family of planes in $\mathbb{R}^3$ that is invariant under rotation about the z-axis or translation in the z-direction. The indicated line L is *Legendrian*, that is, $T_xL \subset \xi_x$ for all $x \in L$. Note the planes twist slowly to the left as you move along L in either direction. This example can be made more explicit by taking ξ to be the kernel of $\alpha = r^2 d\theta + dz$ and checking that it is a contact structure.

The next definition and many of the results in the section are due to Giroux, [18].

2000 *Mathematics Subject Classification*: Primary 57M50; Secondary 53C15.
Keywords and phrases: tight, contact structure.
Supported in part by NSF grants DMS-0073029 and DMS-0406158.

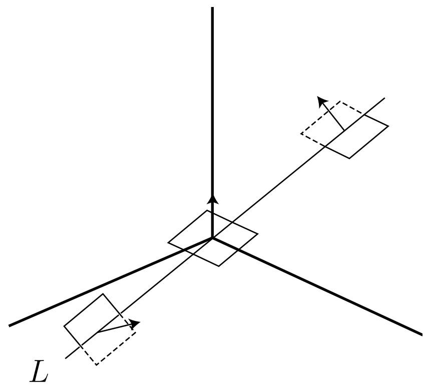


Figure 1. A rotationally symmetric contact structure on $\mathbb{R}^3$.

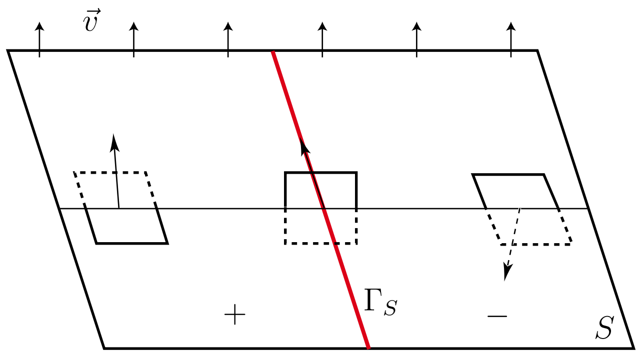


Figure 2. Dividing curve on a convex surface.

Definition (1.3). A vector field $\vec{v}$ is called a *contact vector field* for ξ if flowing in the $\vec{v}$ direction preserves the contact planes. A surface $S \subset (M, \xi)$ is *convex* if there exists a contact vector field $\vec{v}$ transverse to S. If $\partial S \neq \emptyset$, we also require, for S to be convex, that ∂S be Legendrian.

Example (1.4). If ξ is the contact structure of Example (1.2), it follows that any horizontal plane is convex by considering the constant vector field $\vec{v} = \frac{\partial}{\partial z}$. Indeed it follows that any surface in $\mathbb{R}^3$ transverse to the vector field $\frac{\partial}{\partial z}$ is convex.

Roughly, S is convex if and only if S has a *product neighborhood*. Convexity is a global condition; all smooth surfaces are locally convex.

Definition (1.5). If $S \subset (M, \xi)$ is convex, *the dividing set* is denoted Γ_S and is defined to be $\{x \in S \mid \vec{v}(x) \in \xi_x\}$.

Intuition: If we think of the vector field $\vec{v}$ as vertical or perpendicular to S, then Γ_S are those points whose contact planes are perpendicular to S.

Definition (1.6). The *induced (singular) foliation* $\mathcal{F}_S$ on S is defined by integrating the line field $\xi_p \cap T_pS$ on S.

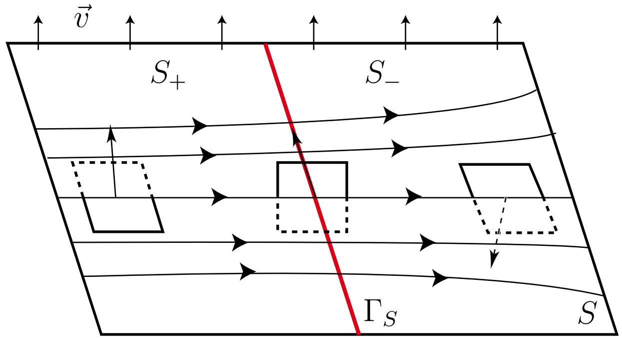


Figure 3. The induced foliation near a dividing curve.

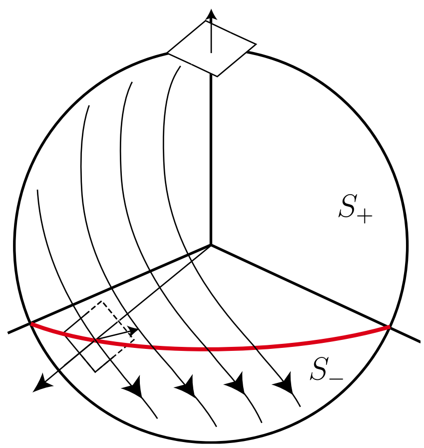


Figure 4. The induced foliation on a small sphere.

Since we are assuming that manifolds and submanifolds are oriented, and since our contact structures are positive, the contact planes inherit a transverse orientation. Orientations on M, S, and ξ allow us to orient the leaves of $\mathcal{F}_S$. Comparing these orientations leads to two ways of thinking about dividing sets:

1. The dividing set Γ_S divides S into regions where the contact planes are right side up or upside down relative to S and as shown in Figure 2.

2. With respect to the induced foliation on S, Γ_S divides S into *source* and *sink regions*, labeled S_+ and S_-, respectively, in Figure 3.

Example (1.7). Figure 4 shows the induced foliation and dividing set on a small round sphere about the origin of Example (1.2). Contact structures are locally homogeneous by Pfaff's Theorem. It follows that there exist small spheres like this about every point of any contact structure.

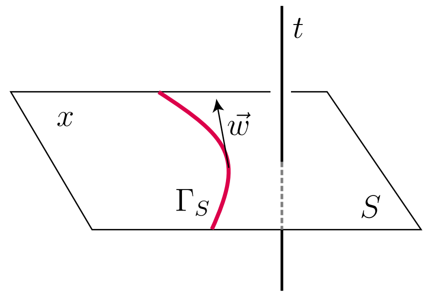


Figure 5. Coordinates near a convex surface.

PROPOSITION (1.8). *The dividing set Γ_S is a 1-dimensional submanifold of S transverse to $\mathcal{F}_S$.*

Proof. Choose coordinates $x \in S$ and t in the $\vec{v}$ direction. Then the 1-form defining ξ may be written $\alpha = \beta(x) + f(x)dt$ where $\beta(x)$ is a 1-form on S, and f is a function on S. Since $\ker\alpha = \xi$ we have:

1. $\alpha_x\left(\frac{\partial}{\partial t}\right) = 0$ if and only if $f(x) = 0$, and therefore $\Gamma_S = f^{-1}(0)$.

2. $0 \neq \alpha\wedge d\alpha = (\beta+f dt)\wedge(d\beta+df dt) = \beta\wedge d\beta+\beta df dt+f dt d\beta = \beta df dt+f dt d\beta$. Therefore, if $f(x) = 0$, then $\beta df dt \neq 0$, and in particular $df \neq 0$. It now follows that $\Gamma_S = f^{-1}(0)$ is a submanifold of S.

3. Let $\vec{w}$ be tangent to Γ_S. Then $dt(\vec{w}) = df(\vec{w}) = 0$, so $\beta df dt \neq 0$ implies that $\beta(\vec{w}) \neq 0$, that is, $\vec{w} \notin ker\beta = T\mathcal{F}_S$. It now follows that Γ_S is transverse to the induced foliation on S. ☐

The next several results are used throughout the paper.

PROPOSITION (1.9). *The isotopy class of Γ_S does not depend on the choice of the contact vector field $\vec{v}$.* ☐

Definition (1.10). Let L be a Legendrian curve that is a boundary component of a surface S. The *twisting of ξ with respect to the framing induced by S* is denoted by $t(L, Fr_S)$. Since we are assuming S and ξ are oriented, $t(L, Fr_S)$ will be an integer. We use the convention that if the planes of ξ twist to the left with respect to S as we move around L, then $t(L, Fr_S)$ is negative.

THEOREM (1.11) (Existence of convex surfaces). *Every closed surface can be approximated by a convex surface. If S is a surface with Legendrian boundary, and if the twisting of ξ with respect to S is negative on each boundary component of S, then S can be approximated, relative to ∂S, by a convex surface.*

Theorem (1.11) was proved by Giroux [18] for closed surfaces and by Honda for surfaces with boundary [20]. The approximating convex surfaces can be chosen to be C^∞ close to S on the interior of S and C^0 close to S along ∂S. Convex surfaces with Legendrian boundary were first used by Kanda [25]. Theorem (1.11) follows from

PROPOSITION (1.12). *If ∂S is Legendrian and the induced foliation $\mathcal{F}_S$ is Morse-Smale, that is,*

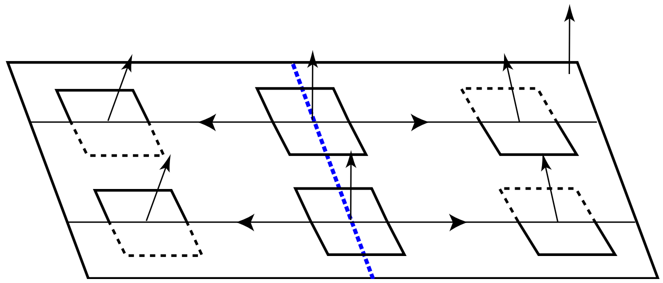


Figure 6. Legendrian divide.

1. *$\mathcal{F}_S$ has a finite number of closed leaves and Morse type singularities,*
2. *there are no saddle-saddle connections,*
3. *the holonomy about closed leaves is linear and either attracting or repelling,*

then S is convex. □

Example (1.13)*.* Proposition (1.12) gives sufficient, but not necessary conditions for a surface to be convex. Figure 6 shows a portion of convex surface whose induced foliation has a circle's worth of singularties. The circle of singularities is called a *Legendrian divide*. The contact structure hinted at in the figure is invariant under translation in the vertical direction or parallel to the Legendrian divide. Note that a Legendrian divide is not a dividing curve.

Figure 7 shows a product neighborhood of the Legendrian divide and the effect of a slight perturbation of the original surface on the induced foliation.

This example hints at a remarkable theorem about the possible induced foliations that can occur on perturbations of a convex surface. Roughly, the Giroux Flexibility Theorem states that we can force the induced foliations to be whatever we like, within reason. The statement of the theorem will make more sense after reading the definitions which follow it.

Theorem (1.14) (Giroux Flexibility Theorem [18]). *Let $S \subset (M, \xi)$ be a convex surface with dividing set Γ_S, and let $\mathcal{F}$ be an arbitrary singular foliation on S* divided by *Γ_S, then there exists an isotopy of S fixing Γ_S (and keeping S transverse to $\vec{v}$) such that at the end of the isotopy, $\mathcal{F}$ is the induced foliation on S.*

Definition (1.15)*.* We say Γ_S *divides* $\mathcal{F}$ if Γ_S cuts S into a maximal number of sink and source regions, that is, regions in which the induced foliation either points in at every boundary component or out at every component of each region, respectively.

Example (1.16)*.* Figure 8 shows three foliations. The first and last are divided by the indicated curves, but the middle example is not; it is not cut into a maximal number of sink and source regions as the third example shows. Another example to which the Giroux Flexibility Theorem can be applied is to replace the indicated foliation on the annular region between the two dividing curves of the third example with a Legendrian divide.

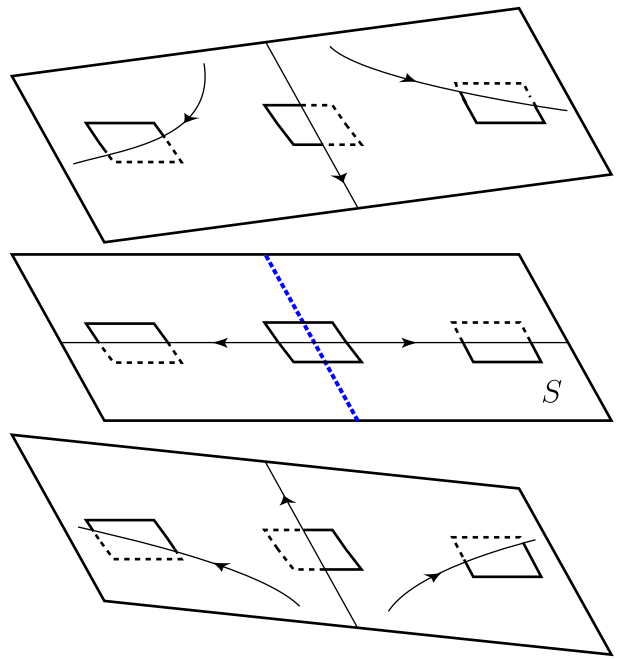


Figure 7. Flexibility near a Legendrian divide.

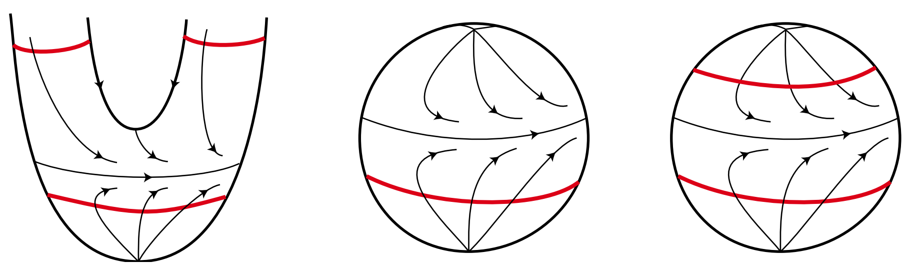


Figure 8. Hypothetical induced foliations.

Definition (1.17). (M, ξ) is *tight* if there does not exist an embedded disk $D \subset M$ such that D is tangent to ξ along its boundary (i. e., $T_xD = \xi_x$ for all $x \in \partial D$). (M, ξ) is called *overtwisted* if it is not tight.

Overtwisted contact structures are classified by their underlying 2-plane bundles [11]. The notion of tightness is analogous to tautness or non-existence of Reeb components in foliation theory or incompressibility of surfaces. We shall see that tight contact structures reflect the underlying topology of the 3-manifolds which carry them.

Figure 9 shows an overtwisted disk that would live in the contact structure described in Example (1.2) if the contact planes were allowed to rotate too quickly along rays leaving the origin.

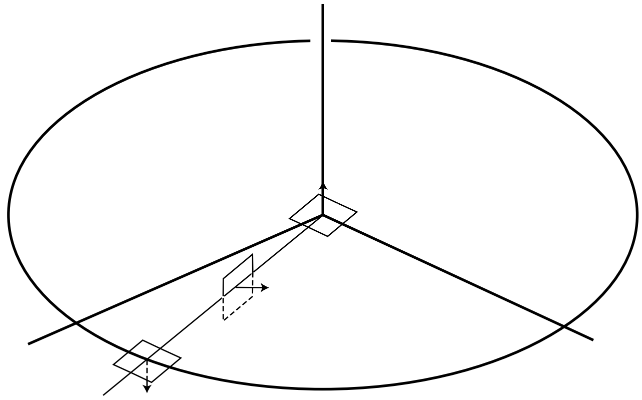


Figure 9. An overtwisted disk.

Proposition (1.18) (Giroux [19]). *If $S \subset (M, \xi)$ is convex, a product neighborhood of S is tight if and only if one of the following is satisfied:*

1. *$S = S^2$, and Γ_S is connected*
2. *$S \neq S^2$, and no component of Γ_S is null-homotopic in S.*

Sketch. ($\Rightarrow$) If either (1) or (2) is false, use the Giroux Flexibility Theorem to realize a null-homotopic Legendrian divide, as discussed in Example (1.16). The disk in S bounded by the Legendrian divide is an overtwisted disk.

($\Leftarrow$) We need a starting point and gluing theorems. That is, until this point, we have not even stated that there are any tight contact structures on any manifold. The next theorem addresses this. Given simple examples of tight contact structures we require gluing theorems to produce more complicated examples. This paper will eventually describe several gluing theorems. Another strategy, used by Giroux, is to produce models in which the desired S and Γ_S exist and must be tight. □

Theorem (1.19). *There exists a tight contact structure on B^3. Moreover, two tight contact structures which induce the same foliations on ∂B^3 are diffeomorphic.* □

The existence portion of the theorem is due to Bennequin [1], and the uniqueness portion is due to Eliashberg [12]. In light of the Giroux Flexibility Theorem, we paraphrase Theorem (1.19) by saying that there is a unique tight contact structure on B^3.

Convex surfaces are required to have Legendrian boundary. Therefore to decompose manifolds with convex boundaries along convex surfaces, we need to know which curves on a convex surface S can be "made Legendrian". That is, we need to know which curves are contained in the leaves of some foliation $\mathcal{F}$ divided by Γ_S. Knowing this will allow us to decide if S can be perturbed so that ∂S becomes Legendrian. The next definition and theorem of Honda's [20] exactly answers this question.

Definition (1.20). A properly embedded 1-submanifold C of a convex surface S is *non-isolating* if

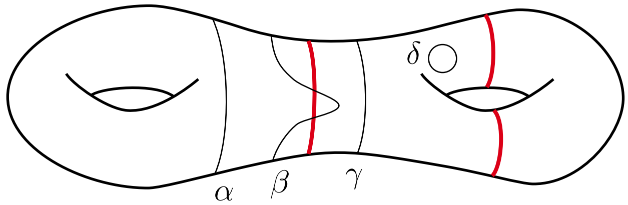


Figure 10. Hypothetical Legendrian curves.

1. C is transverse to Γ_S and
2. the closure of every component of $S \backslash C$ intersects Γ_S.

Theorem (1.21) (Legendrian Realization Principle). *If C is non-isolating then C can be made Legendrian.*

Sketch. The non-isolating condition guarantees that C can be extended to a foliation divided by Γ_S. Then use Giroux Flexibility to realize this foliation on S. ☐

Example (1.22). Of the curves shown in Figure 10, only β and γ are non-isolating. Notice that any curve, such as β, which intersects Γ_S is non-isolating. It is not too hard to extend, say β, to a singular foliation on S divided by Γ_S; however, β will end up passing through singularities, that is, it may not be a smooth curve on S. Note also that in the definition of non-isolating, C is not necessarily connected or closed.

2. Preview

At this point we have enough of the foundational tools in place to sketch, in general terms, some of the issues and techniques involved in studying contact structures from a cut-and-paste point of view.

Classification: Given a 3-manifold M and a collection of curves Γ contained in ∂M, how many tight contact structures, up to equivalence, are there on M with $\Gamma_{\partial M} = \Gamma$? Equivalence might be either diffeomorphism or isotopy taking one contact structure to another.

To be specific, consider the case of a solid torus with four dividing curves on its boundary shown in Figure 11.

Decomposition: How many "sensible" ways are there to decompose such an (M, Γ)?

Continuing with the solid torus example, Figure 11 suggests that there are just two possible decompositions, thus, there are at most two tight contact structures carried by (M, Γ).

Gluing: Which of the decompositions into tight pieces can be glued to form a tight union?

Unlike many situations in 3-dimensional topology, it is very difficult to give general conditions under which the union of tight pieces is tight. The problem

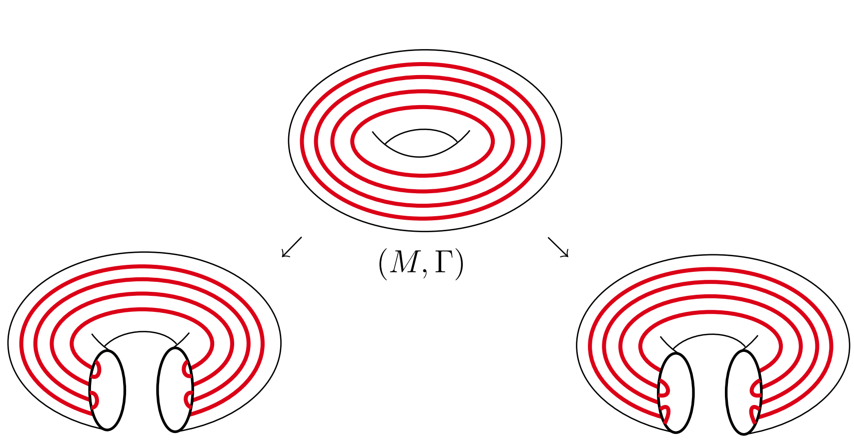


Figure 11. Different convex decompositions.

is that a manifold can contain a large overtwisted disk, but when it is chopped into small pieces, none of the pieces may contain overtwisted disks themselves.

It turns out that regluing either of the decompositions shown in Figure 11 gives a tight contact structure. *A priori*, we do not know that these two contact structures are different. By gluing we can conclude only that (M,Γ) carries at least one tight contact structure.

Invariants: Of the various ways of gluing into a tight union, which result in non-isotopic contact structures?

In our example, an Euler charactistic type invariant shows that the two gluings result in different contact structures. It follows that (M,Γ) carries exactly two tight contact structures.

3. Convex Decompositions

A convex decomposition can be viewed in two ways. First, you can start with a contact structure on a 3-manifold M and keep splitting M along convex surfaces until the pieces are balls. Alternatively, you can start with M and a collection of curves Γ on ∂M that you hope will end up being dividing curves for a contact structure that you are trying to build, and then split along surfaces which you hope will end up being convex. We need to see how actual convex surfaces intersect so that this structure can be correctly modelled in the definition of a convex decomposition.

Example (3.1). The kernel of $\alpha_k = \sin(2\pi kz)dx + \cos(2\pi kz)dy$ defines a contact structure on $\mathbb{R}^3$ shown in Figure 12. In this example, the contact planes all contain the z-axis, that is, any vertical line is Legendrian. The foliation induced on horizontal planes is a linear foliation with slope changing as the height of the plane increases. The vector field given by $\frac{\partial}{\partial r}$ in cylindrical coordinates is a contact vector field, thus a cylinder at constant distance from the z-axis is convex. The dividing curves on this cylinder start on the x-axis and spiral upwards at a rate depending on k. Figure 12 also shows the tangencies of the contact planes and the cylinder as long dashed lines starting on the y-axis.

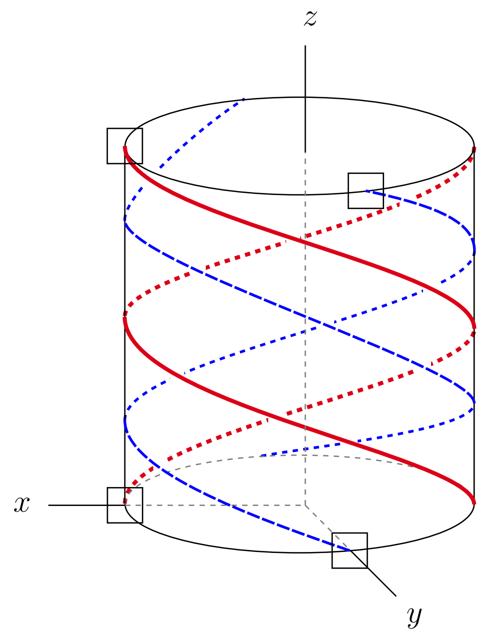


Figure 12. The neighborhood of a Legendrian curve.

Let D^2 be the unit disk. By restricting to the cylinder $D^2 \times [0,1]$ and identifying $D^2 \times \{0\}$ and $D^2 \times \{1\}$, α_k defines a contact structure on $D^2 \times S^1$. The key features of this contact structure are:

1. $T = \partial(D^2 \times S^1)$ is convex
2. $\#\Gamma_T = 2$
3. $\text{slope}(\Gamma_T) = -\frac{1}{k}$

The next theorem may be paraphrased by saying that Legendrian curves, such as the quotient of the z-axis in the previous example, have standard neighborhoods.

Theorem (3.2) (Kanda [25], Makar-Limanov [26]). *There is a unique tight contact structure on $D^2 \times S^1$ such that (1), (2), and (3) hold.* □

When we start with a manifold with convex boundary and cut it along a convex surface, the cutting surface, by definition of convexity, intersects the boundary in a Legendrian curve. The next example is a portion of the region shown in Figure 12. From it we see how the dividing curves on a pair of intersecting surfaces are related near their Legendrian curve of intersection.

Example (3.3). In Example (3.1) the xz-plane is convex with respect to the contact vector field $\frac{\partial}{\partial y}$ and similarly the yz-plane is convex with respect to $\frac{\partial}{\partial x}$. Figure 13 shows portions of these planes, labelled F and G and their dividing curves. Notice that Γ_F and Γ_G are horizontal lines starting on the z-axis and ending at a point of tangency of a contact plane and the vertical cylinder.

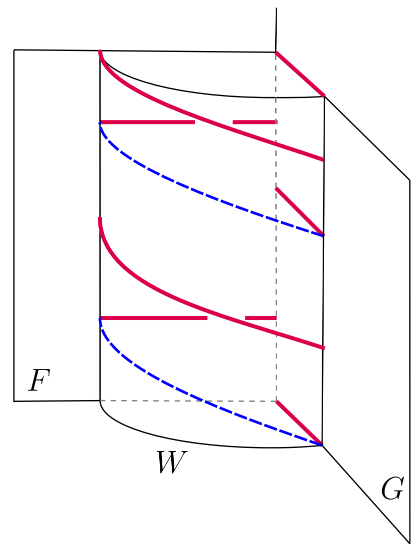


Figure 13. Intersecting convex surfaces.

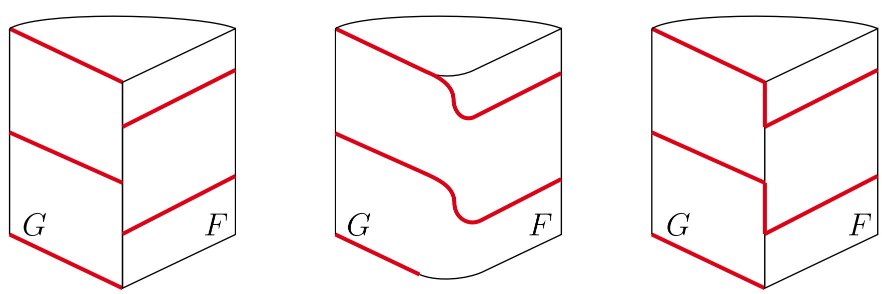


Figure 14. Dividing curves before and after smoothing.

From this we see that for general intersecting convex surfaces F and G, the endpoints of Γ_F and Γ_G alternate along curves of $F \cap G$. Further examination of Figure 13 shows that if the corner of the wedge W subtended by F and G is smoothed, the manifold produced has convex boundary and the dividing curves of F and G are joined by turning to the right (when viewed from the outside of W). The "turn to the right" rule that is forced on us in the presence of a *positive* contact structure serves as the model for defining the orientation conventions in convex decompositions.

Figure 14 shows three views of W. The first shows W before rounding corners, the second is after rounding corners. The last picture shows W without the corner rounded, but it shows the effect on the dividing curves of corner rounding. Most of the figures in this paper are drawn in this fashion.

Definition (3.4). (M, γ) is a *sutured manifold* if:

1. $\gamma \subset \partial M$ is a union annuli and tori,
2. $(\partial M)\backslash\gamma$ is a disjoint union of two subsurfaces $R_+(\gamma)$ and $R_-(\gamma)$, and

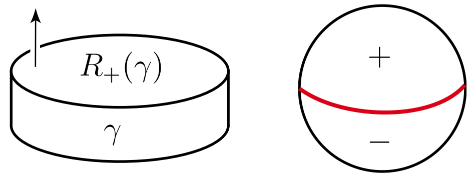


Figure 15. B^3 with a single suture.

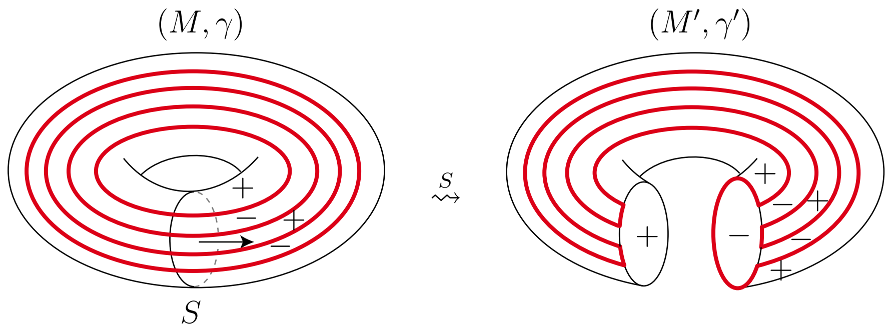


Figure 16. Sutured manifold splitting.

3. crossing an annular suture takes you from $R_\pm(\gamma)$ to $R_\mp(\gamma)$.

Gabai [14] defined sutured manifolds to study taut foliations. We are primarily concerned with the case that all sutures are annuli. Figure 15 shows two views of a solid ball with a single annular suture. The first view shows a manifold with corners, as it should be drawn. The second shows how sutures are drawn; the manifold appears smooth, and the sutures are very skinny.

Definition (3.5). If S is an oriented properly embedded surface in (M, γ), $(M, \gamma) \overset{S}{\rightsquigarrow} (M', \gamma')$ is defined by $M' = M \backslash S$ and introducing sutures as needed to separate the positively and negatively oriented portions of $\partial(M \backslash S)$ as shown in Figure 16.

Definition (3.6). A *convex structure* is a pair (M, Γ) such that:
1. Γ is a disjoint union of curves in ∂M,
2. ∂M split along Γ is the disjoint union of two subsurfaces, $R_+(\Gamma)$ and $R_-(\Gamma)$, and
3. crossing a dividing curve takes you from $R_\pm(\Gamma)$ to $R_\mp(\Gamma)$.
Also assume each component of ∂M has dividing curves on it.

A sutured manifold (M, γ) is a manifold with corners. Convex structures (M, γ) are smooth. Figure 17 shows portions of a sutured manifold near a suture and a convex structure near a dividing curve. Notice that the 2-planes along each arc α turn over as α is traversed, but they do so in different fashions.

Of course we hope that (M, Γ) will carry an actual tight contact structure, just as Gabai would like (M, γ) to carry a taut foliation; however there is no *a priori* reason that it will. When discussing a surface with curves on it, such as

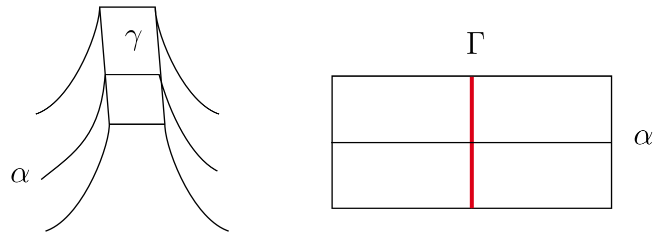


Figure 17. An arc crossing a suture compared to an arc crossing a dividing curve.

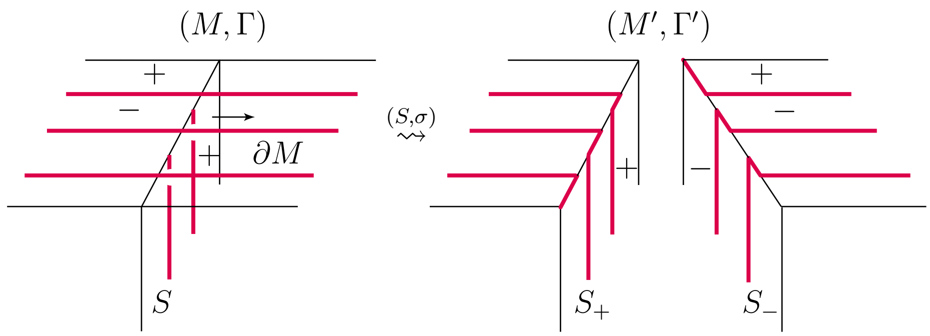


Figure 18. A convex splitting,

(S,σ) in the next definition, there is no need to distinguish between an "abstract" convex surface and an actual convex surface, for a contact structure is uniquely determined in a (product) neighborhood of S by the dividing curve configuration σ.

Definition (3.7). Let (S,σ) be a convex surface in (M,Γ) such that ∂S is non-isolating in ∂M, and the endpoints of σ alternate with points of $\Gamma\cap\partial S$ along ∂S. Define $(M,\Gamma) \overset{(S,\gamma)}{\rightsquigarrow} (M',\Gamma')$ by $M' = M\backslash S$ and by adding new portions of dividing curves to $(\sigma\cup\Gamma)\backslash S$ using the "turn to the right" rule shown in Figure 18.

A sutured manifold with annular sutures (M,γ) naturally determines a convex structure (M,Γ) by replacing each annulus of γ by its core. To be able to use Gabai's existence theorems for sutured manifold decompositions in our setting, we must be able to start with a sutured manifold splitting (the top row of the following diagram) and then produce a convex surface, (S,σ), such that the diagram commutes. We discuss how this can be done through a series of examples.

$$\begin{array}{ccc} (M,\gamma) & \overset{S}{\rightsquigarrow} & (M',\gamma') \\ \downarrow & & \downarrow \\ (M,\Gamma) & \overset{(S,\sigma)}{\rightsquigarrow} & (M',\Gamma') \end{array}$$

Example (3.8). Figure 19 shows how to introduce *boundary-parallel* dividing curves, σ, so that the diagram commutes. This technique works, provided that every component of ∂S has nonempty intersection with Γ.

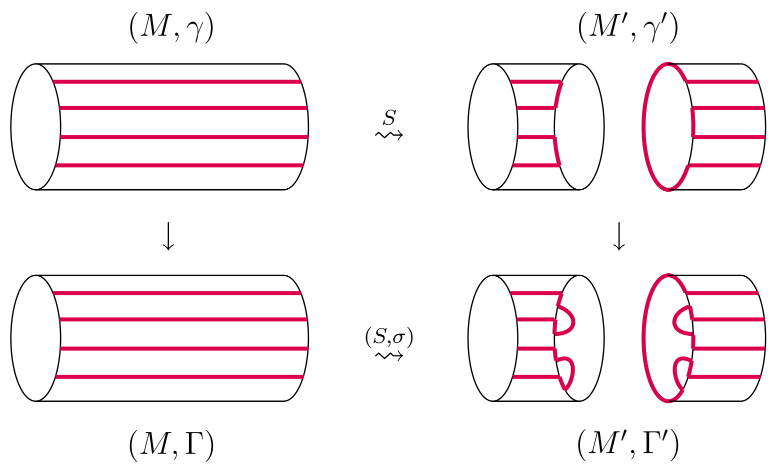


Figure 19. A commuting diagram of sutured manifold and convex splittings.

Definition (3.9). A convex surface (S,σ) has *boundary-parallel dividing curves* if ∂S is nonempty, every component of ∂S intersects σ, and σ is collection of arcs each of which bounds a half disk that contains a portion of ∂S but no other arcs of σ.

Example (3.10). Now consider the possibility that $\partial S \cap \Gamma = \emptyset$. Such an S might have isolating boundary, that is, it might not be possible to make it Legendrian and hence S convex. Figure 20 shows first a sutured manifold splitting along a surface S with $\partial S \cap \Gamma = \emptyset$. The second two portions of the figure show two possible ways of introducing intersections between ∂S and Γ and of adding boundary-parallel σ to S.

There are two key features in this example. First, the strategy of introducing intersections can only work if there are dividing curves on the same component of ∂M as ∂S – this will show up in the definition of "sutured manifold with annular sutures" below. And second, only one of the perturbations of S makes the splitting diagram commute.

Figure 21 is similar to Figure 20 in that $\partial S \cap \Gamma = \emptyset$, but in this case there are multiple portions of ∂S, each with its own orientation preference for creating a pair of intersections with Γ, and they cannot all be satisfied simultaneously. Rather than describe how to get around this, we just point out that Gabai confronted a similar situation in developing sutured manifold theory. He introduced a notion of "well-groomed" sutured decompositions, that is, he showed that splittings could be assumed to have coherently oriented boundary components, and for such splittings we can produce a commutative splitting diagram using the technique of Example (3.10).

Theorem (3.11). *Let* (M,γ) *be an irreducible* sutured manifold with annular sutures*, and let* (M,Γ) *be the corresponding convex structure. The following are equivalent.*

1. (M,γ) *is* taut.

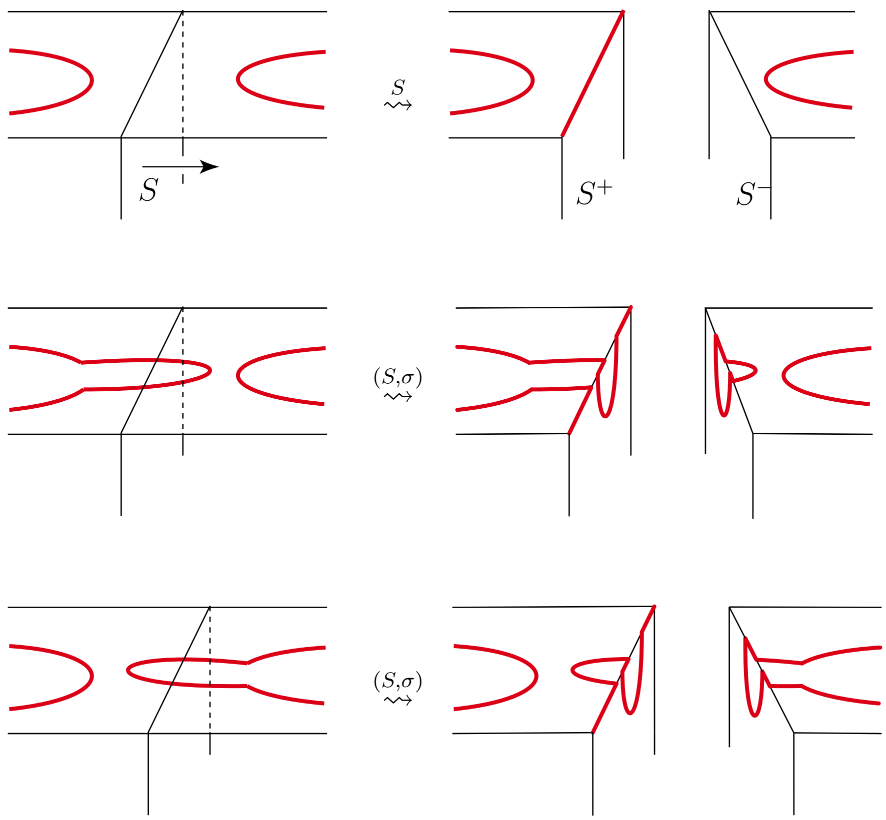


Figure 20. Two possibilities for (S, σ).

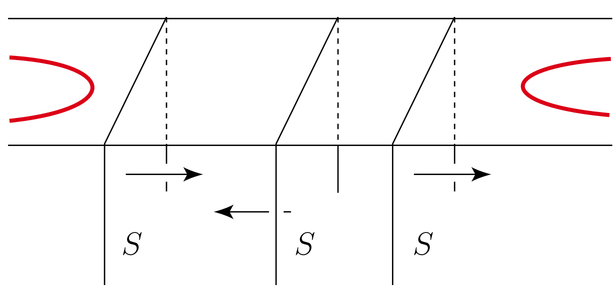


Figure 21. There is no consistent way to introduce intersections between ∂S and Γ.

2. (M, γ) *has a* sutured manifold decomposition*.*
3. (M, γ) *carries a* taut foliation*.*
4. (M, Γ) *carries a* universally tight *contact structure.*
5. (M, Γ) *carries a tight contact structure.*

Definition (3.12). A sutured manifold has *annular sutures* if each component of M has nonempty boundary, every boundary component contains at least one annular suture, and if there are no toroidal sutures.

(M, γ) is *taut* if $R_+(\gamma)$ and $R_-(\gamma)$ are Thurston norm minimizing in their homology class in $H_2(M, \gamma)$. A *sutured manifold decomposition* of M is a sequence of splittings

$$(M, \gamma) \overset{S_1}{\rightsquigarrow} \cdots \overset{S_m}{\rightsquigarrow} \cup(B^3, S^1 \times I)$$

where $(B^3, S^1 \times I)$ denotes the sutured manifold shown in Figure 15.

A foliation is *taut* if every leaf intersects a closed transversal.

A contact structure is universally tight if $(\widetilde{M}, \tilde{\xi})$ is tight.

Thurston [28] proved (3) implies (1). Gabai [14] proved (1) implies (2) and (2) implies (3). Eliashberg and Thurston [13] showed (3) implies (4). It is immediate that (4) implies (5). All of these results apply without the additional assumption of annular sutures. Since S^3 carries a tight contact structure but cannot support a taut foliation, some additional hypothesis is necessary for (5) to imply (1).

The techniques of (5) implies (1) are not used in the rest of the paper, so we instead sketch a direct proof of (2) implies (4) that has the advantages of making the importance and utility of universal tightness clear. The proof introduces a gluing strategy that is used repeatedly.

Proof of (2) implies (4). First replace the given sutured manifold decomposition with a corresponding convex decomposition

$$(M, \Gamma) \overset{(S_1,\sigma_1)}{\rightsquigarrow} \cdots \overset{(S_m,\sigma_m)}{\rightsquigarrow} \cup(B^3, S^1)$$

By Theorem (1.19), (B^3, S^1) carries a (universally) tight contact structure. By construction, the surfaces (S_i, σ_i) have boundary-parallel dividing curves (see Definition (3.9)). This portion of the theorem follows from the next gluing theorem. □

Theorem (3.13) (Colin [7]). *Suppose that* (M', Γ') *is obtained from* (M, Γ) *by splitting along a convex surface* (S, σ)*, that is,* $(M, \Gamma) \overset{(S,\sigma)}{\rightsquigarrow} (M', \Gamma')$*. If* M *is irreducible,* S *has boundary parallel dividing curves, and* (M', Γ') *carries a universally tight contact structure, then so does* (M, Γ)*.*

Sketch. We will illustrate key ideas of our interpretation [22] of Colin's gluing theorem with examples. The proof strategy is to:

1. Suppose the contact structure on (M, Γ) obtained by gluing (M', Γ') along (S, σ) is overtwisted, and let D be an overtwisted disk.
2. (In small steps) isotop S to S' and eventually off D.
3. While isotoping S, make sure that $M \backslash S'$ stays universally tight.

This strategy gives a contradiction once $S' \cap D = \emptyset$, for $M \backslash S'$ is both tight and contains the overtwisted disk D.

"In small steps" refers to a fundamental idea due to Honda [21]. That is, any isotopy of a convex surface S can be expressed as a sequence of *bypasses*.

Definition (3.14). A *bypass* consists of:

1. a Legendrian arc α connecting 3 dividing curves in S,
2. a Legendrian arc β joining $\partial\alpha$,

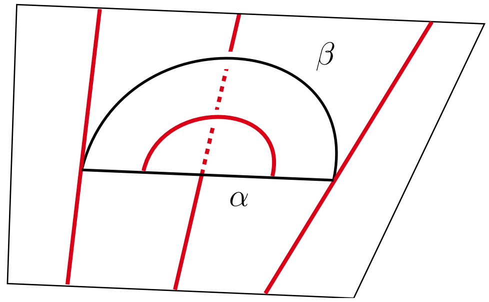


Figure 22. A bypass attached to S along α.

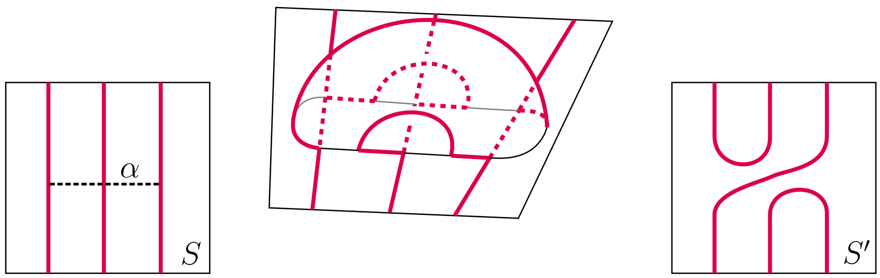


Figure 23. The effect on Γ_S of isotoping S through a bypass attached along α.

3. a convex half disk in $M\backslash S$ with boundary equal to $\alpha \cup \beta$ which contains a single dividing curve.

Figure 23 shows the effect on σ of isotoping S across a bypass. Notice that the "turn to the right" rule for dividing curves going around corners looks more like a "turn to the left" rule when it is viewed from inside the manifold.

Example (3.15). The global effect on the dividing curves of a bypass move depends very much on how the local picture sits with respect to the entire surface and dividing curve set. Figure 24 shows an example in which the arc of attachment connects two parallel dividing curves to a third. Isotoping S across this bypass has the effect of removing two dividing curves from S.

Continuing with the gluing theorem, we now consider some examples of bypasses that S might have to be isotoped through while moving S off of D. Hopefully universal tightness of $M\backslash S'$ follows from universal tightness of $M\backslash S$ in each case.

Example (3.16). Notice that the dividing curves in Figure 25 are boundary-parallel. If such a bypass were to exist in M, then, as shown, S' would contain a null-homotopic dividing curve. Proposition (1.18) implies the existence of an overtwisted disk near S'. Since S' may be thought of as living in the complement of S, and we are assuming $M\backslash S$ is universally tight, the bypass drawn in Figure 25 cannot exist.

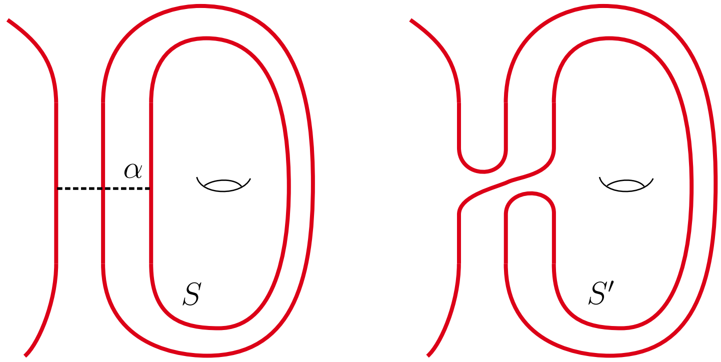


Figure 24. This bypass removes two parallel dividing curves.

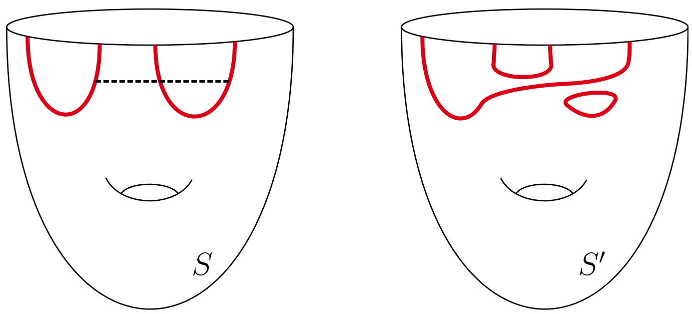


Figure 25. This bypass cannot exist in a tight contact structure.

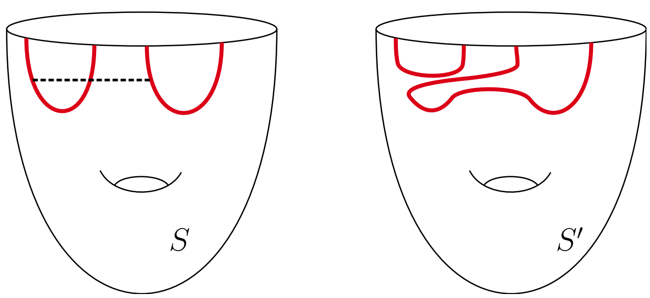


Figure 26. A trivial bypass.

Example (3.17). Figure 26 shows a similar-looking bypass that has a very different effect on the dividing curves of σ. Up to isotopy, the dividing curves are unchanged. Though we omit the proof here (see [22]), it is a consequence of the uniqueness of tight contact structures on a ball that S and S' cobound a contact product. It follows that $M \backslash S$ and $M \backslash S'$ are contactomorphic, hence both are universally tight.

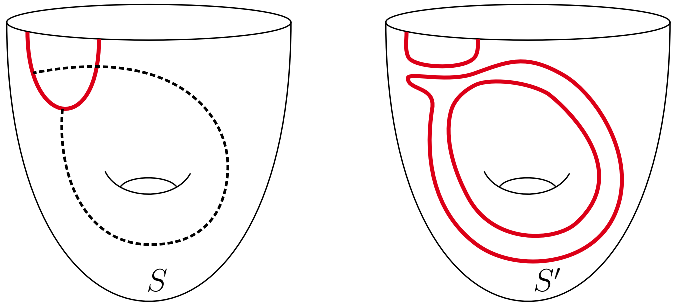


Figure 27. It is not clear what the implications of such a bypass are in general.

Example (3.18). Figure 27 shows the most mysterious of these examples. There is no reason for $M\backslash S'$ to be tight if we only assume tightness of $M\backslash S$. However, α represents a nontrivial element of $\pi_1(S)$ or $\pi_1(M)$, and any cover $\widetilde{M\backslash S}$ is still tight (by universal tightness). If we lift to the right cover, α is unwound, and this example becomes the same as the previous example. Thus the proof strategy may be continued in this and subsequent covers.

There are several other types of bypass configurations to check, but this pattern repeats itself. Bypasses are of three types: those which cannot exist, those which cause no trouble if they do exist, and all of the rest. The typical situation is that troublesome bypasses can be dealt with in the right cover. This is the point and power of the assumption of universal tightness. □

4. Tori

Example (4.1). Let $\alpha_k = \sin(2k\pi z)dx + \cos(2k\pi z)dy$ as in Example (3.1). Restricting α_k to the cube $[0,1]\times[0,1]\times[0,1]$ and identifying the front with the back face and the left with the right face defines a contact structure on $T\times I$. Neither $T\times\{0\}$ nor $T\times\{1\}$ is convex. Perturbing $T\times\{0\}$ and $T\times\{1\}$ so that they are convex gives the contact structure ξ_k on $T\times I$ shown in Figure 28. A vertical annulus, such as the one shown on the front face, is convex and has $2k$ closed dividing curves.

The next theorem gives a very general, but rough, classification theorem for tight contact structures.

Theorem (4.2). *If M is irreducible, then M carries finitely many tight (or universally tight) contact structures if and only if M is atoroidal.*

The if direction is due to Colin, Honda, and Giroux [10], and the only if direction is due to Colin [8, 9] and Honda, Kazez, and Matić [22].

Proof. We will explain the following portions of the proof of finiteness direction:

1. There is a finite collection of branched surfaces in M which carry every tight contact structure.

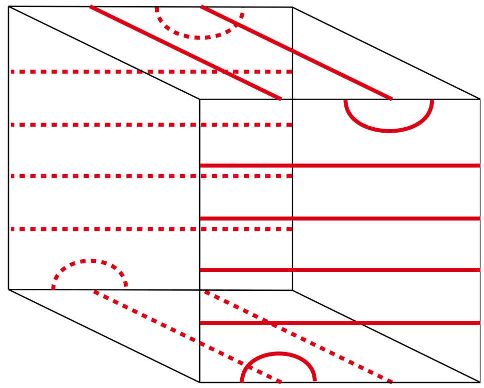


Figure 28. ξ_2 on $T \times I$.

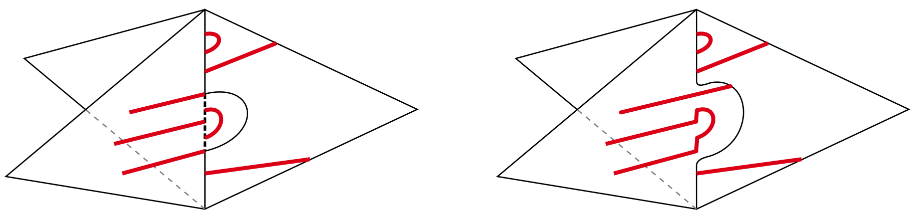


Figure 29. Isotoping an edge of τ across a bypass.

2. If a branched surface carries infinitely many tight contact structures then it carries tori.

Proof of (1).

• Pick a triangulation τ of M, and isotop it until τ^1 is a collection of Legendrian arcs.

• Isotop τ^2 relative to τ^1 so that each face is convex.

• Isotop τ to remove interior ∂-parallel dividing curves. Figure 29 shows how to pry a two cell open along an edge to effect a bypass move and accomplish this.

• For each $\Delta \in \tau^3$, group the dividing curves in $\partial\Delta^3$ so that, except for a bounded number of dividing curves near the vertices, they are contained in at most 5 prisms P_i. See Figure 30.

• Use Giroux Flexibility to force the foliation induced by ξ on each $\partial_v P_i$ to be a union of vertical arcs and on $\partial_h P_i$ to be a fixed non-singular foliation.

• By the uniqueness of tight contact structures on B^3 we may assume all vertical arcs in P_i are Legendrian.

• The union over $\Delta \in \tau^3$ is naturally a neighborhood $N(B)$ of a branched surface B, and by construction there are only finitely many such B.

• Each component of $\Delta^3 - N(B)$ is a polygonal ball. The number of such polygonal balls is bounded, and the possible dividing curve configurations on the boundary faces of each polygonal ball is also bounded. Thus ξ is defined by $\xi\,|_{N(B)}$ up to finitely many choices.

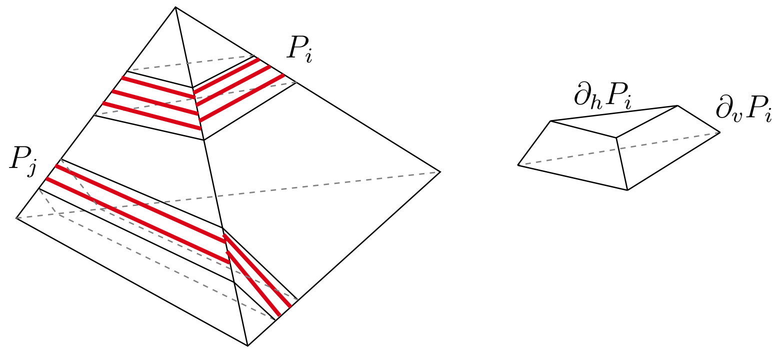


Figure 30. Dividing curves on $\partial\Delta$ are carried by a family of prisms.

Proof of (2). Suppose two contact structures ξ_0, ξ_1 are carried by B. By construction the foliations induced on $\partial_h N(B)$ agree, thus ξ_1 is defined by ξ_0 and a finite set of integer weights on the sectors of B which describe the twisting of the planes of ξ_1 relative to the planes of ξ_0 along vertical Legendrian arcs of $N(B)$.

An infinite collection of contact structures all carried by one branched surface give an infinite collection of weights. Since the contact structures are all positive, there is a lower bound, perhaps negative, on these weights. It follows that there must be a non-negative collection of integer weights on B. In the standard way, these non-negative weights can be used to piece together a surface in $N(B)$ that is transverse to the vertical Legendrian arcs. The induced foliation on such a surface has no singularities, thus the surface is either a torus or a Klein bottle.

We draw two conclusions from this portion of the argument.

• Changing weights along a torus does not change the homotopy class of the 2-plane bundle, thus it follows that only finitely many 2-plane bundles support tight contact structures.

• The only way to produce infinitely many contact structures on a given space is to insert twisting in a neighborhood a torus of the sort exemplified by ξ_k, as defined in Example (4.1).

With this in mind we sketch some of the remaining steps in the infinitely many portion of the theorem.

3. A toroidal manifold has a universally tight contact structure.
4. Inserting ξ_k near the torus preserves universal tightness
5. and changes the contact structure.

Proof of (3).We will assume $\partial M = T$. This is just one gluing theorem away from full generality. The sutured manifold (M, T), where T is a toroidal suture, is automatically taut, and by Gabai's theorem it has a sutured manifold decomposition. For simplicity, assume the first splitting surface S intersects T in a single curve, and say $(M, T) \overset{S}{\rightsquigarrow} (M', \gamma')$. The new suture γ' is the annulus obtained by splitting T along ∂S.

We would like to consider the corresponding convex decomposition, but first we must fix the toroidal suture. Pick two parallel curves on T dual to $S \cap T$

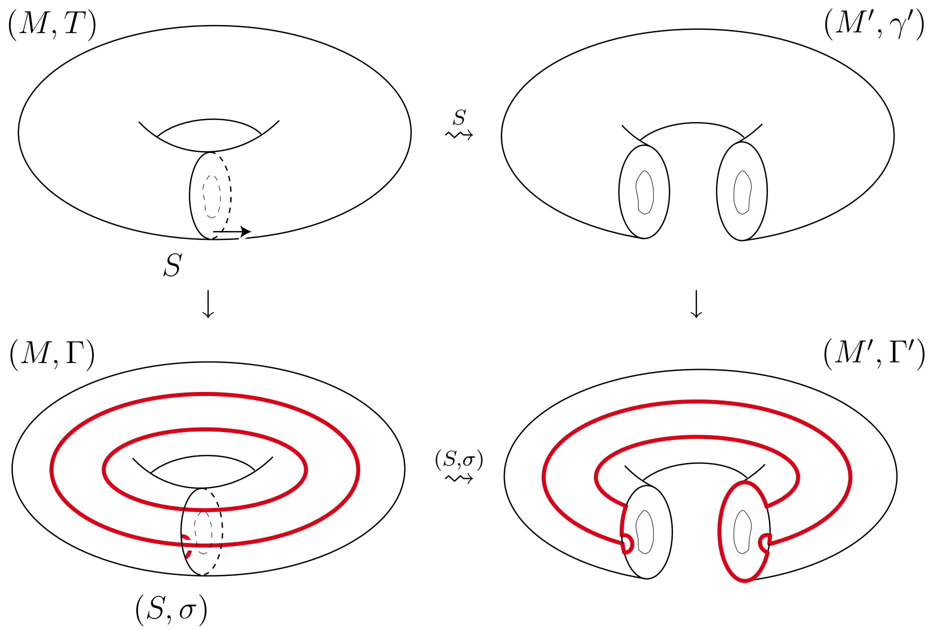


Figure 31. Introducing dividing curves on a torus suture.

and define them to be Γ. Figure 31 shows how we can force the usual correspondence between the sutured manifold decomposition on the first row and the convex structure on on the second row. It is particularly important to note that (S,σ) has boundary-parallel dividing curves. Since (M',γ') has a sutured manifold decomposition, (M',Γ') carries a universally tight contact structure. By Theorem (3.13) (M,Γ) does also.

The technique of adding a pair of parallel dividing curves to a boundary component with no sutures can be used in other settings as well.

Proof of (4).

Continuing with the same M, S, and T, we need to show that the contact structure on $M \cup (T \times I)$ obtained by gluing the structure built in (3) and ξ_k is universally tight. Gluing along T is beyond the scope of Theorem (3.13). Instead we will compare

$$(M,\Gamma) \overset{(S,\sigma)}{\rightsquigarrow} (M\backslash S,\Gamma')$$

and

$$M \cup (T \times I) \overset{S\cup A}{\rightsquigarrow} (M\backslash S) \cup (T \times I\backslash A)$$

where the contact structure on $T \times I$ is ξ_k, S is the first decomposing surface, and A is an annulus extending ∂S that is used to keep track of the k twists in ξ_k.

The first row of Figure 32 shows two views of $M\backslash S$ near $T\backslash S$. In the first 3-dimensional picture, a pair of dividing curves becomes a single dividing curve after corner rounding. In the second, the same neighborhood is expressed as a product with S^1, and the single dividing curve is shown as a point. The second row of the figure gives a similar view of $(M\backslash S) \cup (T \times I\backslash A)$ in the case $k = 1$.

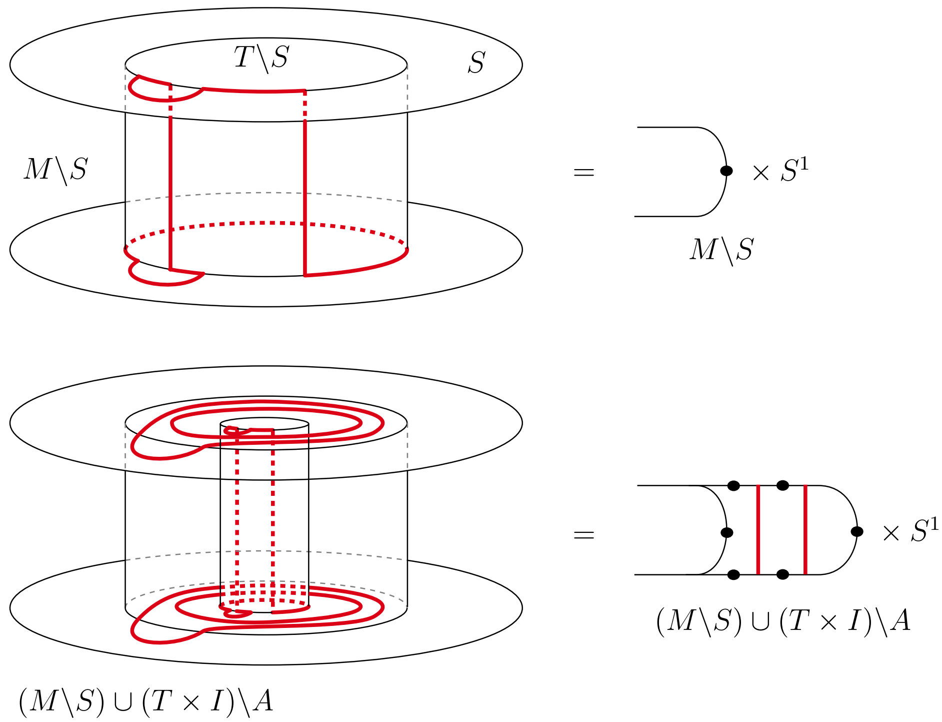


Figure 32. Views of $M\backslash S$ near the boundary before and after adding $T \times I\backslash A$.

The second figure in the second row shows a convex surface transverse to the S^1 direction that detects the twisting along ξ_k.

We now show that our contact structure $(M\backslash S) \cup (T \times I\backslash A)$ is tight. This is done by finding an embedded copy of this space in the tight space $M\backslash S$, and then using the obvious (but useful!) fact that a subset of a tight space is tight. Here is how this is done.

The curve C parallel to ∂S shown in Figure 33 is isolating. Pass to a cover, without changing notation, in which the number of boundary components of S is increased, and then C becomes non-isolating. Then use flexibility to make C a Legendrian divide. Figure 34 shows a product with S^1 view of $M\backslash S$ near $T\backslash S$. The region shown is the product of a convex disk and S^1. This disk is shown with part of two dividing curves that end on the two points of intersection with C, and these two points of C are shown as hollow dots.

Figure 35 shows $(M\backslash S) \cup (T \times I\backslash A)$. Finally, Figure 36 shows a larger version of $(M\backslash S)$ than Figure 34. The shaded subset of Figure 36 is contact isomorphic to $(M\backslash S) \cup (T \times I\backslash A)$ and is necessarily (universally) tight.

Next we show the contact structure on $M \cup (T \times I)$ is tight. This requires a gluing theorem along $S \cup A$, which unlike S, does not have boundary-parallel dividing curves. Figure 37 shows $\Gamma_{S \cup A}$.

We use the same gluing strategy:

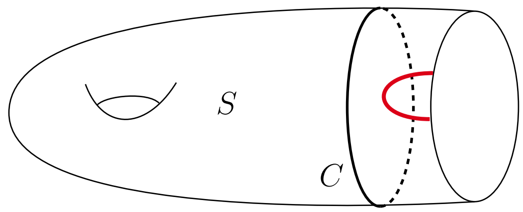


Figure 33. The curve C is isolating.

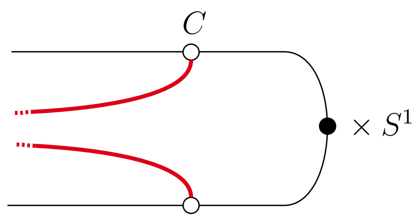


Figure 34. C is now a Legendrian divide.

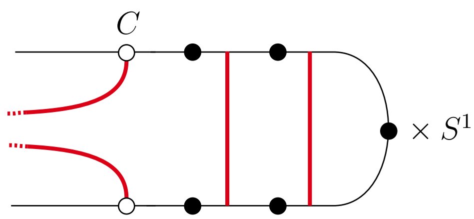


Figure 35. $M \cup (T \times I)$ split along $S \cup A$.

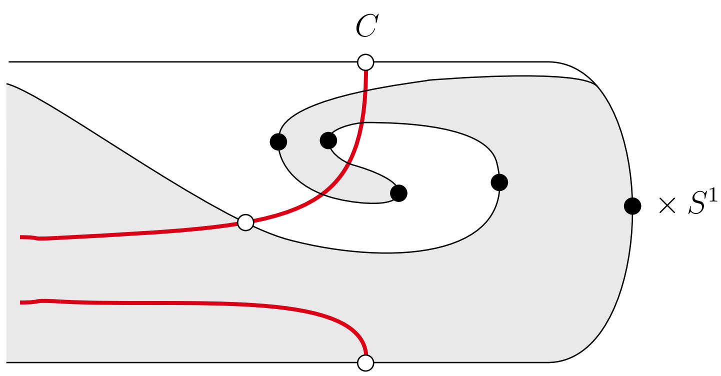


Figure 36. Enlarged view of $M \cup T$ split along S shown with a distinguished subset.

(a) Assume the union along $S \cup A$ is overtwisted.
(b) Isotop, via bypasses, $S \cup A$ off of the overtwisted disk, and
(c) argue that the split manifold stays tight during (b).

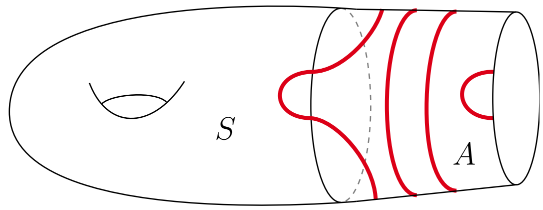


Figure 37. Dividing curves on $S \cup A$.

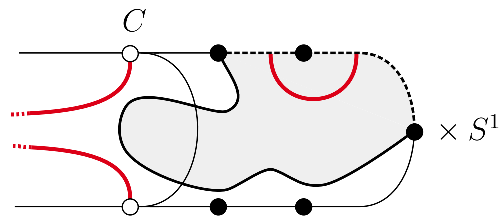


Figure 38. A possible location of a bypass.

There is an important change in perspective though. From the point of view of $S \cup A$, performing an isotopy in (b) is equivalent to digging a bypass out of one side of $S \cup A$ and adding it to the other side. We prove (c) by showing that there are no troublesome bypasses that can be dug from either side of $S \cup A$ in $M \cup (T \times I)$.

We have discussed bypasses which involve only boundary-parallel dividing curves in the proof of Theorem (3.13), so now we consider the existence of a bypass involving the closed dividing curves on $S \cup A$.

Example (4.3). Figure 38 shows a bypass attached along a dotted curve α connecting three different dividing curves on the boundary of $(M \cup T \times I) \backslash (S \cup A)$. Figure 35 showed the same space, but it showed a convex disk with a different set of dividing curves. Certainly this bypass, if it exists, is not a subset of that convex disk. Indeed the bypass itself may be very large and reach out of the portion of the manifold shown in either of these figures.

Proposition (4.4). *The bypass shown in Figure 38 cannot exist.*

Proof. Consider the space obtained by adding the product of a bypass attached along α and S^1 that is shown in Figure 39. The null-homotopic dividing curve that is created implies the contact structure is overtwisted.

The opposite conclusion is reached if we add the same set and consider its relationship to the convex disk that we know is contained in $(M \cup T \times I) \backslash (S \cup A)$. The union is shown in Figure 40. We see that this space must be tight, for it also is a "fold along C", that is, it can be found as a subset of $M \backslash S$. This completes the proof of the Proposition and hence Step (4). $\square$

Proof of (5).

We will only give the idea behind the technique used in producing invariants that distinguish the various contact structures on $M \cup (T \times I)$. Let F be a

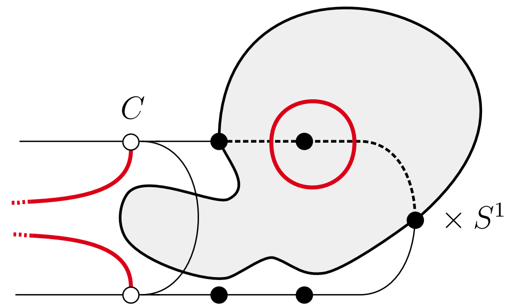


Figure 39. Detecting a bypass by adding a "template".

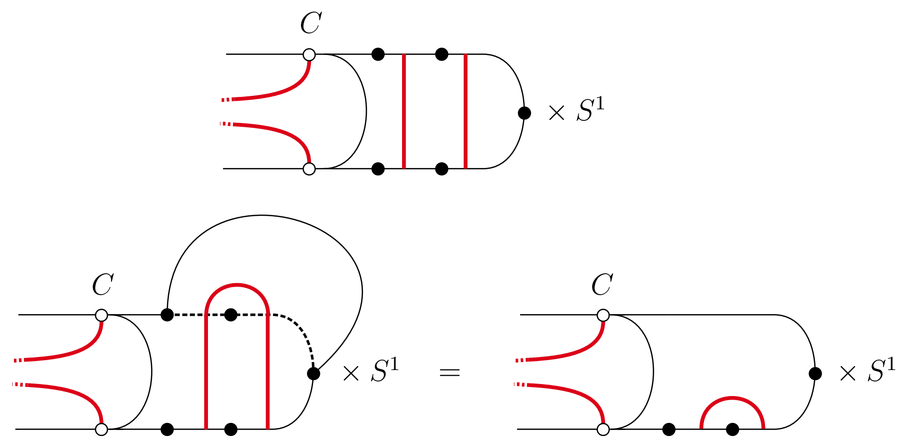


Figure 40. The actual result of adding the template.

properly embedded convex surface which intersects the boundary component $T \times \{1\}$ of $M \cup (T \times I)$, and let δ be a homotopically essential arc in F which starts and ends on $T \times \{1\}$. The minimum of $\#|\delta \cap \Gamma_F|$ over all such δ and F is an invariant which tends to infinity as the twisting, that is, the k in ξ_k, is increased.

□

5. Surface bundles

We will give a classification of tight contact structures on $\Sigma \times I$ such that

(*) Σ is a closed surface with genus at least two, and the dividing curves on each component of $\partial(\Sigma \times I)$ are a pair of parallel non-separating curves.

Example (5.1). In Figure 41, notice that on Σ_1, $\chi(R_+) = \chi(\Sigma_1)$ and $\chi(R_-) = 0$. This is analogous to the product foliation on $\Sigma \times I$ in the sense that the contact 2-planes have outward pointing normal vectors everywhere on Σ_1, at least as measured by Euler charactertistic. This structure is a special case of an *extremal* contact structure.

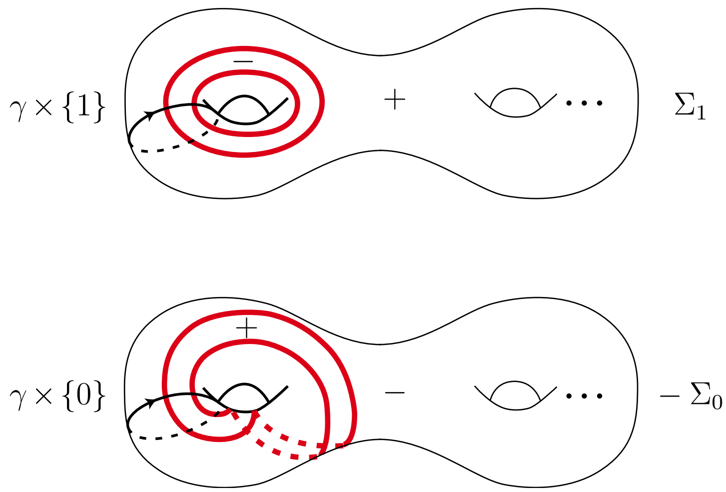


Figure 41. Dividing curves and splitting annulus on $\Sigma \times I$.

Definition (5.2). A contact structure on a surface bundle with fibre Σ is *extremal* if the Euler class, $e(\xi)$, of the 2-plane bundle satisfies $e(\xi)(\Sigma) = \pm\chi(\Sigma)$. Equivalently, if Σ is convex either $\chi(R_+) = \chi(\Sigma)$ or $\chi(R_-) = \chi(\Sigma)$.

Theorem (5.3). [23] *There are exactly 4 (universally) tight non-product contact structures satisfying* (*). *They correspond to a choice of dividing curve on each of Σ_0 and Σ_1.*

This theorem is used to prove the following theorems.

Theorem (5.4). *Let φ be a pseudo-Anosov map of a closed surface Σ. There is a unique extremal, tight (or universally tight) contact structure on $(\Sigma \times I)/(\varphi(x), 1) \sim (x, 0)$.*

Theorem (5.5) (Gabai-Eliashberg-Thurston Theorem). *If M is Haken and $H_2(M) \neq 0$, then M carries a universally tight contact structure.*

Theorem (5.5) follows from Gabai's work [14] on the existence of taut foliations and Eliashberg and Thurston's perturbation technique [13] for producing universally tight contact structures from taut foliations. The proof we give [24] is a direct construction which has the advantage of helping us discover new gluing theorems.

Sketch of proof of Theorem (5.3). We concentrate only on the dividing curve configuration shown in Figure 41. What the proof strategy lacks in subtlety it makes up in directness. We start by decomposing $\Sigma \times I$ along a vertical annulus $\gamma \times I$ whose boundary components are shown in Figure 41 and analyze all dividing curve configurations that can occur on $\gamma \times I$.

Figure 42 shows $\gamma \times I$ cut by a vertical arc into a rectangle, and it lists all possible dividing curve configurations such that the boundary components of $\gamma \times I$ intesect the dividing curves twice each.

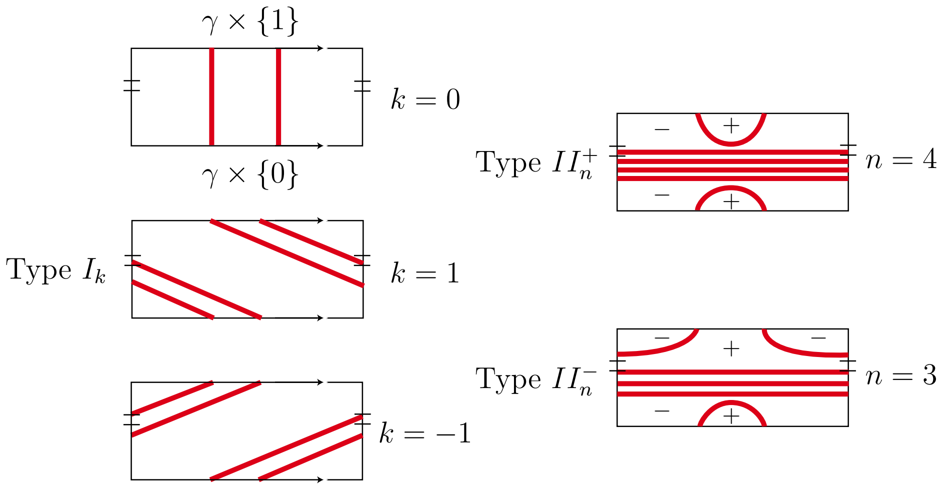


Figure 42. Possible dividing curves on $\gamma \times I$.

We will show how Type II_2^+ can be reduced to Type II_0^+, that is, if we start with a convex annulus of Type II_2^+, we can find another convex annulus of Type II_0^+. This case is fairly typical of the type of arguments we use to prove this classification theorem.

Figure 43(A) shows the result of splitting $\Sigma \times I$ along an an annulus of Type II_2^+, and Figure 43(B) shows the dividing curves on $\Sigma \backslash \sigma \times \{0\}$. After rounding the corners and gathering the dividing curves on the two vertical annuli of $(M \backslash \sigma) \times I$, the result is shown in Figure 43(C). Cut this along a convex rectangle $\delta \times I$, and again we must consider all possible dividing curve configurations on a splitting convex surface.

Suppose first that the dividing curve configuration on $\delta \times I$ has a boundary-parallel component whose half disk contains the point labelled 2 in $\delta \times I$. (Figure 43(D) shows a different configuration.) This would imply the existence of a bypass with arc of attachment running from 1 to 3. Consider how this bypass is situated relative to $\sigma \times I$ – it crosses two parallel dividing curves and ends on a third curve in Figure 43(A). But a quick computation (Example (3.15)) shows that the effect of pushing $\sigma \times I$ across such a bypass removes both closed dividing curves from $\sigma \times I$, that is, it produces a Type II_0^+ annulus that we had promised to find.

The same logic shows that if any dividing curve is boundary-parallel and centered on the point 3, 4, or 5, a Type II_0^+ can be shown to exist. We are left with the dividing curve configuration of Figure 43(D). The result of cutting along $\delta \times I$ and rounding corners is shown in Figures 43(E,F).

Lemma (5.6). *There must exist a bypass along the arc of attachment shown in Figure 43(E).*

Sketch of proof of lemma. If such a bypass exists, then pushing the vertical annulus of Figure 43(E) through it produces three dividing curves rather than just the one shown in Figure 43(F). One step in a complete proof is arguing

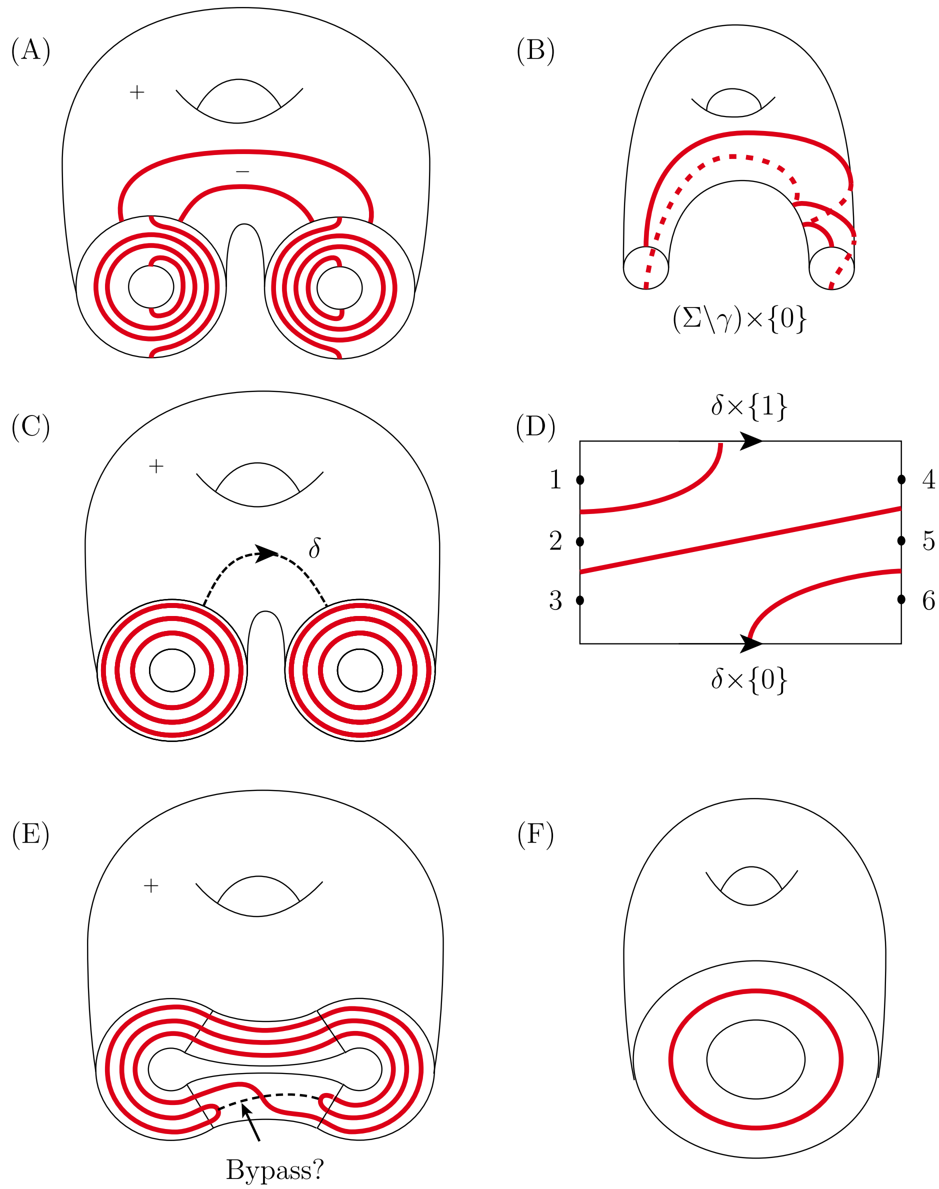


Figure 43. Decomposing $\Sigma \times I$.

that there is a vertical annulus with three dividing curves near the given vertical annulus. This is very similar to the part of the proof of tightness in the toroidal case, shown in Figure 36, in which a more complicated space was shown to exist inside a manifold with a single dividing curve on a vertical annulus in its boundary. In short, the annulus is shown to exist by a folding argument along an isolating curve that is shown to exist near the original annulus. This more complicated annulus implies the existence of the desired bypass. $\square$

Pushing $\delta \times I$ through this bypass produces a new dividing curve configuration on $\delta \times I$ which has boundary-parallel dividing curves centered on the points labelled 3 and 4. As above, this allows us to modify $\gamma \times I$ and produce an annulus of Type II_0^+. The rest of the proof of Theorem (5.3) requires:

1. Many other similar reductions.
2. Existence of the four types of contact structures must be shown.
3. Uniqueness of these contact structures must be established.

We will show the existence and uniqueness of Type II_0^+ tight contact contact structures. Figure 44(A,B) shows $\Sigma \times I$ cut along $\gamma \times I$. Notice that the dividing curves on $\gamma \times I$ are boundary-parallel. The result of rounding corners and the next splitting surface is shown in Figure 44(C). Since $\delta \times I$ intersects only two dividing curves, there is no choice for the diving curve set on $\delta \times I$, it must be the single boundary-parallel arc shown in Figure 44(D). The result of cutting along $\delta \times I$ and rounding corners is shown in Figure 44(E,F). Next, choose ϵ such that $\epsilon \times I$ intersects dividing curves just twice. Continuing in this fashion we produce a convex decomposition by splitting surfaces which have boundary-parallel dividing curves. Morever, as long as null-homotopic dividing curves on the splitting surfaces are not allowed, the choice of dividing curves is unique.

Since the splitting surfaces all have boundary-parallel dividing curves in this case, (2) follows from Theorem (3.13). It is worth emphasizing that a convex decomposition determines the contact structure near ∂M and the splitting surfaces. Furthermore, by the uniqueness of tight contact structures on B^3, the contact structure is determined on the rest of M as well. Since there is just one possible choice of dividing curves on the splitting surfaces, (3) follows.

This completes the sketch of the proof of Theorem (5.3). □

The statement of Theorem (5.3) refers to a choice of dividing curves. This choice can be described explicitly using the notion of a *straddled* dividing curve.

Definition (5.7). A dividing curve in $\partial(\Sigma \times I)$ is *straddled* if there exists a dual convex annulus with a boundary-parallel dividing curve centered on it.

We record a couple of consequences of Theorem (5.3) that are used in applications.

Corollary (5.8) (Addition). *Let ξ_1 and ξ_2 be tight, non-product contact structures on $\Sigma \times [0,1]$ and $\Sigma \times [1,2]$, respectively, that agree on $\Sigma \times \{1\}$. Then $\xi_1 \cup \xi_2$ is tight if and only if no dividing curve on $\Sigma \times \{1\}$ is straddled in both $\Sigma \times [0,1]$ and $\Sigma \times [1,2]$.* □

Corollary (5.9) (Freedom of Choice). *Let ξ be a non-product tight contact structure on $\Sigma \times I$, and let α_1 and α_2 be a pair of parallel, non-separating curves on Σ. Then there exists a convex embedding of Σ in $\Sigma \times I$ that is isotopic to the inclusion of a boundary component, such that $\Gamma_\Sigma = \alpha_1 \cup \alpha_2$.* □

The next proposition is self-evident and very useful [20].

Proposition (5.10) (Imbalance Principle). *Let $S^1 \times [0,1]$ be a properly embedded convex annulus in M such that $S^1 \times \{0\}$ intersects fewer dividing curves than $S^1 \times \{1\}$. Then $S^1 \times [0,1]$ contains a bypass centered on a dividing curve intersecting $S^1 \times \{1\}$.* □

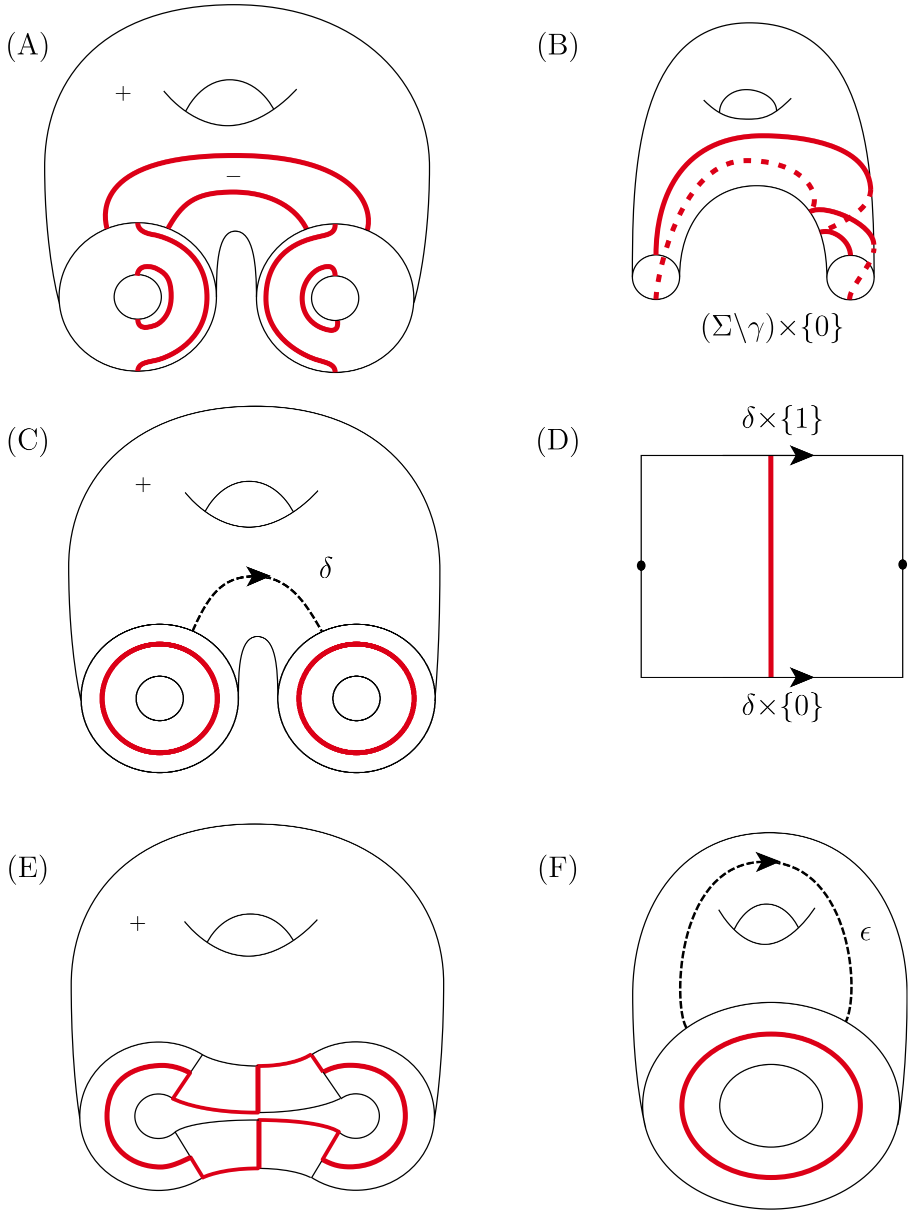


Figure 44. Existence of uniqueness of Type II_0^+.

Sketch of the proof of Theorem (5.4).

• Given a surface bundle with pseudo-Anosov monodromy φ, pick a fibre, isotop it until it is convex, and cut the bundle along the fibre. The dividing curves on each boundary component consist of a family of parallel pairs of curves.

• If there are more than one pair of parallel curves on either boundary component, then since φ is pseudo-Anosov, there exists an imbalance annulus.

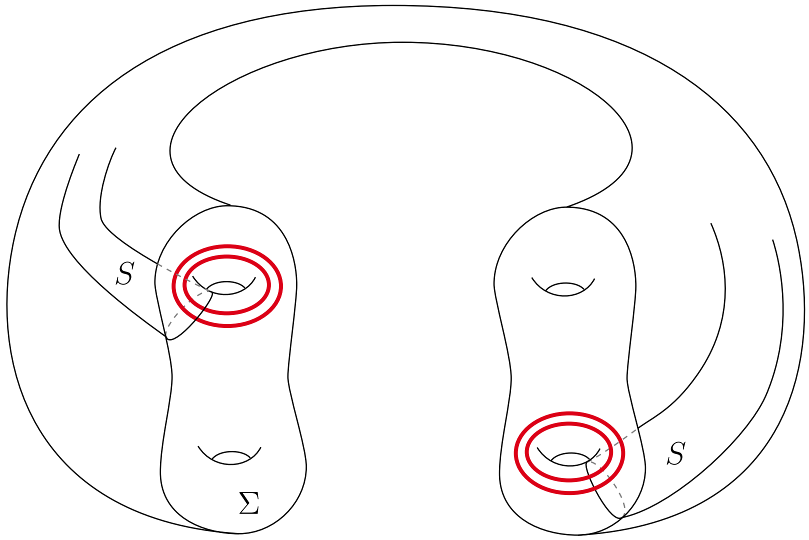


Figure 45. The first splitting surface S in $M\backslash\Sigma$.

• Isotoping the fibre through the bypass guaranteed by the Imbalance Principle reduces the number of dividing curves. Continue until there is just a single pair on each boundary component.

• By Freedom of Choice a new fibre can be chosen with a fixed pair of non-separating dividing curves.

• Splitting the bundle along this fibre reduces an arbitrary bundle to one of the four standard forms given in Theorem (5.3). Of the four possible straddlings, two are ruled out because of the tightness of the gluing that recreates the original surface bundle. The other two are related by another application of Freedom of Choice. □

Sketch of the proof Theorem (5.5). By Theorem (3.11) we may assume M is a closed manifold. Let $\Sigma \subset M$ be a Thurston norm minimizing surface corresponding to a non-zero element of $H_2(M)$ and split M along Σ.

The sutured manifold $M\backslash\Sigma$ has no sutures, but it does have a sutured manifold decomposition. Let the first splitting surface be S, and we shall assume that S intersects each copy of Σ in a single closed curve as shown in Figure 45. Make $M\backslash\Sigma$ a convex structure by adding a pair of parallel dividing curves dual to ∂S on each boundary component. Make S a convex surface by adding boundary-parallel dividing curves σ straddling a component of Γ, the dividing set on the boundary of $M\backslash\Sigma$, on each copy of Σ. If the right curves are straddled, splitting the sutured manifold $M\backslash\Sigma$ along S corresponds to the convex splitting defined by (S, σ). The remaining steps of the convex decomposition are directly inherited from the sutured manifold decomposition.

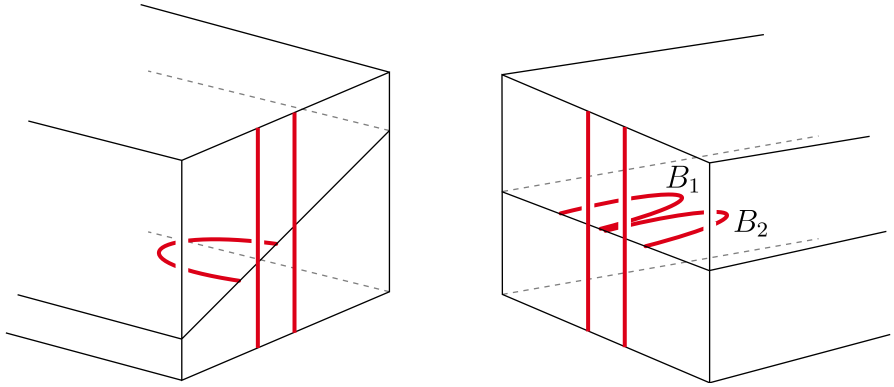


Figure 46. Possible location of bypasses before gluing.

This convex decomposition of $(M\backslash\Sigma, \Gamma)$ is by surfaces, all of whose dividing curves are boundary-parallel. Thus by Theorem (3.13), there is a tight contact structure on $M\backslash\Sigma$. Moreover, by construction, one dividing curve on each boundary component of $M\backslash\Sigma$ is straddled by a dividing curve on S.

By Theorem (5.3) there are four choices of tight contact structure on $\Sigma \times I$ that could be used to attach to $M\backslash\Sigma$ and produce a tight contact structure on M. It should seem very plausible, and it is true, that a curve straddled on both sides gives rise to an overtwisted disk. Thus we insert the unique, non-product, contact structure on $\Sigma \times I$ that gives $(M\backslash\Sigma) \cup \Sigma \times I$ a chance of being tight.

We are two gluing theorems away from a complete proof of tightness on M; we must glue along each of the boundary components of $\Sigma \times I$. As we have seen, the general form of these gluing proofs is:

1. Given an overtwisted disk in M, push Σ off it using bypasses while keeping $M\backslash\Sigma$ tight.

2. Analyze which bypasses exist on one component of $M\backslash\Sigma$ and which can be added to the other component while preserving tightness.

Rather than do this in generality, consider the local version of this that is shown in Figure 46. On the left are two dividing curves, one of which is straddled. On the right are the two dividing curves about to be identified with the curves on the left. Also shown are two boundary-parallel dividing curves. The first, B_1, is known to exist by construction, thus if it is removed and added to the other side, tightness must be shown to be preserved. The second, B_2, if added to the left would produce an overtwisted disk, thus, as part of a sufficient gluing theorem, these must be shown not to exist. These local gluing results follow from the next lemma.

□

Lemma (5.11). *Let γ_u and γ_s be a pair of parallel dividing curves on ∂M, and assume γ_s is straddled and the contact structure on M is tight. Then, adding a bypass to M across γ_u produces a tight contact structure.*

Since adding a bypass to M across γ_s produces an overtwisted structure, it follows that γ_u is not straddled.

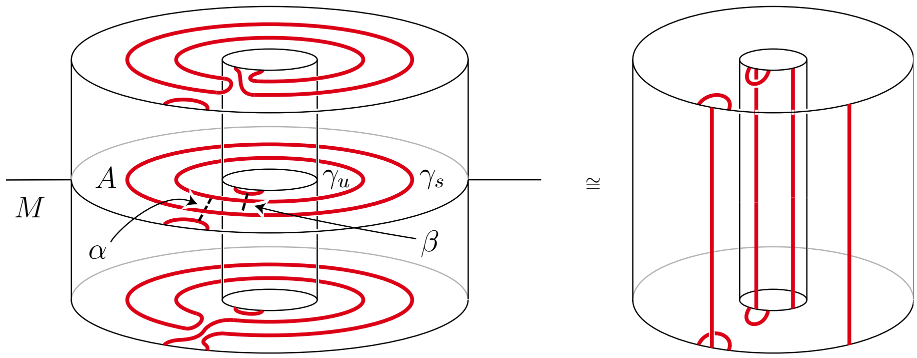


Figure 47. Adding a bypass across β is the same as removing a bypass across α.

Proof. Figure 47 shows a neighborhood $A \times I$ of an annular neighborhood A of γ_s and γ_u in ∂M. It also shows the arc of attachment α which straddles γ_s and the arc of attachment β to which a bypass is being added. The annulus parallel to and below A shows the result of removing the bypass attached along α. The annulus above A shows the result of adding a bypass along β. The figure on the right shows the dividing curves on the boundary $A \times I$.

At least on the boundary, the figure on the right looks like a product contact structure on $A \times I$, and indeed it is. Since the attaching curves, α and β, are disjoint, the contact structure on A can be built by first attaching a bypass to the bottom annulus along β and then attaching a bypass along α. From this point of view, the isotopy class of the dividing curves remains unchanged after adding each bypass (see also Example (3.17)), and thus the contact structure is a product. It now follows that adding a bypass across β is the same as removing a bypass in M attached along α, and this operation preserves tightness. □

6. Open Questions

There are two fundamental classes of open questions:

1. Which M^3 carry tight contact structures?
2. What are the topological implications of carrying a tight contact structure?

The central existence question, particularly from the point of view developed in this paper, is the question of whether or not Haken homology spheres M always carry tight contact structures.

In such a manifold, every surface $\Sigma \subset M$ must separate. In particular if there is a tight contact structure ξ on M, then $e(\xi)(\Sigma) = 0$. This means that if Σ is convex, then $\chi(R_+) = \chi(R_-)$. This is exactly opposite to the extremal case when $\chi(R_+) = \chi(\Sigma)$ and $\chi(R_-) = 0$. Presumably constructing contact structures will involve:

- Classification of such structures on $\Sigma \times I$ and
- new gluing theorems.

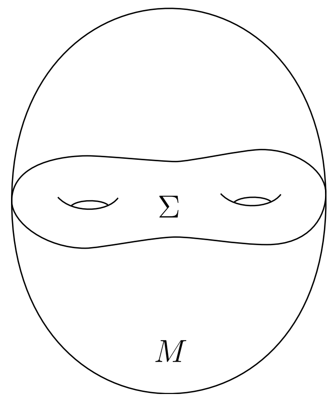


Figure 48. Haken homology sphere.

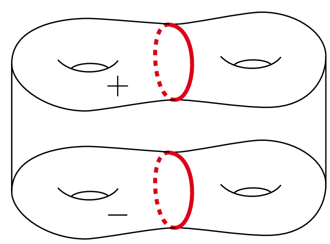


Figure 49. A non-extremal boundary configuration.

Example (6.1). Perhaps the simplest example of this sort of classification question on $\Sigma \times I$ is shown in Figure 49. Preliminary work of Cofer [6] shows there is exactly one tight, non-product, contact structure with these dividing curves. This example has the bizarre property that if you add any non-trivial bypass, it becomes overtwisted. It follows that it does not occur as a subset of any tight contact structure on $\Sigma \times I$ other than itself, and it may not show up in any tight closed 3-manifold.

Very little is known about (2), implications of carrying a tight contact structure, so we will describe results that have been obtained in lamination theory that perhaps have analogues in contact topology.

Definition (6.2). A *lamination* of M^3 is a disjoint union of surfaces which are locally homeomorphic to the product of D^2 and a closed subset of I.

A lamination is *essential* if the leaves are incompressible, the complementary regions are irreducible, and there are no folded leaves. A lamination is *genuine* if it is essential and some complementary region is not a product of a boundary leaf and I.

Figure 50 shows, in order, a folded leaf, a complementary region that is a product of a boundary leaf and I, and a complementary region that is not such a product.

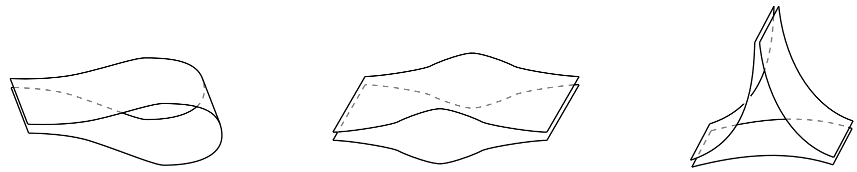


Figure 50. Complementary regions of a general lamination.

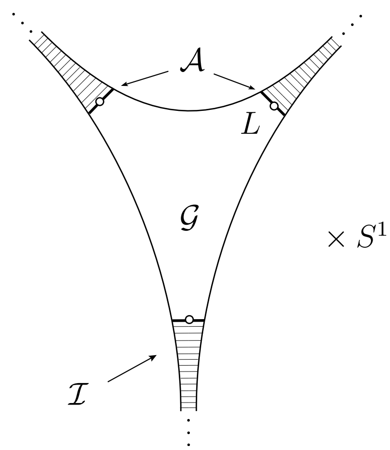


Figure 51. Structure on the complementary region of a genuine lamination.

Definition (6.3). The *Euler characteristic* of a surface with cusped boundary is defined to be the usual Euler characteristic of the underlying space minus half of the number of cusps.

The cross-sections of the complementary regions shown in Figure 50 are a disk with one cusp ($\chi = 1/2$), a disk with two cusps ($\chi = 0$), and a disk with three cusps ($\chi = -1/2$). The definition of essential consists of bans on various types of positive Euler characteristic, while the notion of a genuine lamination postulates the existence of some negative Euler characteristic in M. We shall see that atoroidal manifolds are group negatively curved (Theorem (6.6)). It is not clear what additional structure should be made for contact structures that might make the this theorem apply in that setting as well.

By the JSJ decomposition theorem, there is a unique $I-$bundle structure $\mathcal{I}$ on the ends of each complementary region. Thus each complementary region decomposes as the union of a $\mathcal{I}$ and the *guts* $\mathcal{G}$ as shown in Figure 51.

The key features of this decomposition are:

• $\mathcal{G}$ is compact.

• By maximality of $\mathcal{I}$, $\mathcal{G}$ has no *product disks*, that is, there are no non-trivial rectangles in $\mathcal{G}$ with sides that alternately consist of I-bundle fibres of $\mathcal{I}$ and arcs in leaves of the lamination.

• An essential lamination is genuine if and only if $\mathcal{G} \neq \emptyset$.

• $\mathcal{G} \cap \mathcal{I}$ is a finite union of annuli $\mathcal{A}$.

Definition (6.4). M is *group negatively curved* if there exists a constant C such that for every null-homotopic curve, $f : S^1 \to M$, there exists an extension of f to a disk D such that

$$\mathrm{area}(f(D)) < C \cdot \mathrm{length}(f(\partial D)).$$

M is *group negatively curved with respect to a link L in M* if there exists a constant C such that for every null-homotopic curve $f : S^1 \to M$, there exists an extension of f to a disk D such that

$$\mathrm{area}(f(D)) < C \cdot (\mathrm{length}(f(\partial D)) + \mathrm{wr}(f(\partial D), L)).$$

The *wrapping number* $\mathrm{wr}(f(\partial D), L)$ is a geometric linking number and is defined to be the minimum, taken over all disks E with $\partial E = f(\partial D)$ of the number of points of intersection of E with L.

The inequality in the definition of *group negatively curved with respect to a link L in M* is equivalent to the existence of a constant such that at least one of the two inequalities is satisfied:

$$\mathrm{area}(f(D)) < 2C \cdot \mathrm{length}(f(\partial D))$$

or

$$\mathrm{area}(f(D)) < 2C \cdot \mathrm{wr}(f(\partial D), L).$$

We need the following remarkable theorem.

Theorem (6.5) (Gabai's Ubiquity Theorem [15]). *If M is closed, irreducible, and atoroidal, and if $L \not\subset B^3$, then M is group negatively curved with respect to L.* □

Theorem (6.6). [17] *If M is atoroidal and contains a genuine lamination λ, then M is group negatively curved.*

Before applying Gabai's Ubiquity Theorem to the proof of Theorem (6.6), we need the following lemma which says that to prove an isoperimetric inequality for all null-homotopic curves, it is enough to prove the inequality on a "dense" subset.

Lemma (6.7). *Let A be the set of all null-homotopic curves $g : S^1 \to M$, and let S be a subset of A. If*

- *all $f \in S$ satisfy an isoperimetric inequality,*
- *each $g \in A$ is approximated by an $f \in S$ by a small area homotopy, and*
- *length(f) is not drastically bigger than length(g),*

then all $g \in A$ satisfy an isoperimetric inequality.

Proof. This follows by piecing together the homotopies shown in Figure 52. □

Sketch of the proof of Theorem (6.6). To apply this lemma think of $\mathcal{G}$ as a big, fat subset of M. Then to show M is group negatively curved, it is enough to prove an isoperimetric inequality for the set of null-homotopic curves $f : S^1 \to M$ such that

1. f is transverse to λ.
2. Each component of $f^{-1}(\mathcal{G})$ has length greater than some constant ε.

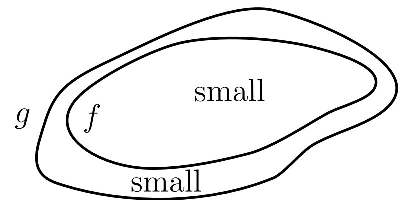


Figure 52. A small area homotopy that does not increase length much.

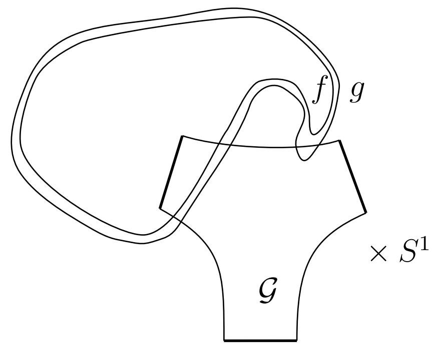


Figure 53. Short portions of $g^{-1}(\mathcal{G})$ can be removed efficently.

In other words, short bits of $f^{-1}(\mathcal{G})$ can be efficiently removed as in Figure 53.

Since $\mathcal{G} \cap \mathcal{I}$ is a finite union of annuli $\mathcal{A}$, we define L to be the union of the cores of $\mathcal{A}$. We now apply Theorem (6.5) to this choice of L. Given a null-homotopic $f : S^1 \to M$ satisfying (1) and (2), there exists a disk of null-homotopy, D, such that at least one of these inequalities is satisfied:

$$\text{area}(f(D)) < 2C \cdot \text{length}(f(\partial D))$$

or

$$\text{area}(f(D)) < 2C \cdot |f(D) \cap L|.$$

In the first case, we have exactly the isoperimetric inequality we are looking for. Thus it is enough to assume the second inequality is satisfied, and then show there exists a constant C' such that

$$2C \cdot |f(D) \cap L| < C' \cdot \text{length}(\partial f(D)).$$

Figure 54 shows $f^{-1}(D)$. The figure shows $f^{-1}(\mathcal{G})$ as shaded, and $f^{-1}(\mathcal{I})$ as white. Since we are only trying to give a sketch of the main ideas, we will think of f as an embedding.

Figure 55 shows regions that might occur as subsets of $f^{-1}(D)$. The first region, a null-homotopic circle, can be removed by choosing a new map of $D \to M$ since leaves of λ are incompressible. The second region, a folded leaf, cannot occur in an essential lamination. And finally the third region, a half disk mapped into $\mathcal{G}$ does occur, and thus we arrive at

Conclusion 1. Regions of $f^{-1}(D)$ with positive Euler characteristic contribute at least ε to the length $f(\partial D)$.

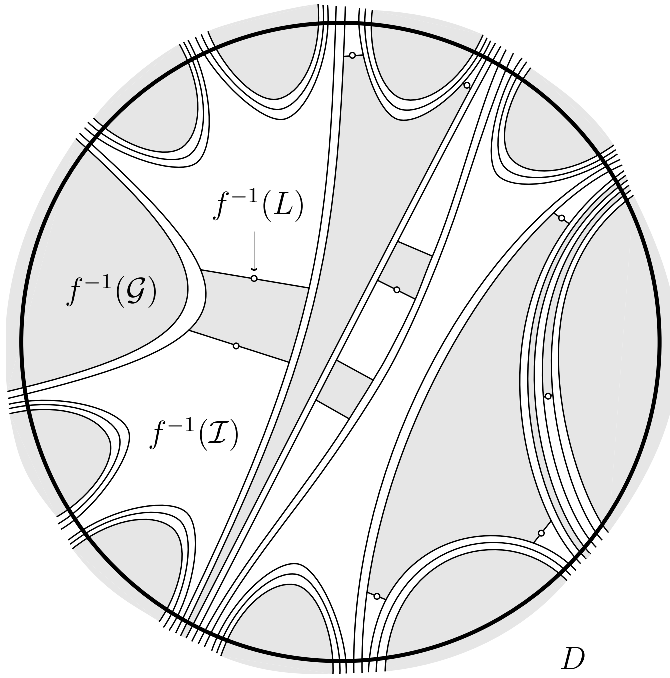


Figure 54. The pullback of $\lambda, \mathcal{G}$, and $\mathcal{I}$ to D.

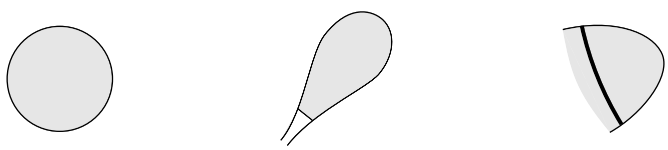


Figure 55. Possible regions of D with positive Euler characteristic.

Figure 56 shows typical regions of $f^{-1}(D)$ which contain points of $f^{-1}(L)$. The first figure, a cusped triangle, has negative Euler characteristic. The second region shown has Euler characteristic zero and is a product disk in $\mathcal{G}$. This cannot exist by the definition of $\mathcal{G}$. The third region also has Euler characteristic zero, but it contains an arc that is mapped into $\mathcal{G}$, thus it contributes at least ε to length$(f(\partial D))$. After removing the middle regions that cannot exist we reach

Conclusion 2. Points of $f^{-1}(L)$ either show up in regions of negative Euler chacateristic or they contribute at least ε to length$(f(\partial D))$. smallskip We can now complete the proof. We have a disk D such that

$$\text{area}(f(D)) < 2C \cdot |f(D) \cap L|,$$

thus a large area disk gives many points of $f^{-1}(L)$. By Conclusion 2, these points either directly contribute to the length of $f(\partial D)$, or they show up in regions of negative Euler characteristic. But $\chi(D) = 1$, thus the existence of regions with negative Euler characteristic implies the existence of regions with positive Euler characteristic. By Conclusion 1, these in turn contribute even more to the length

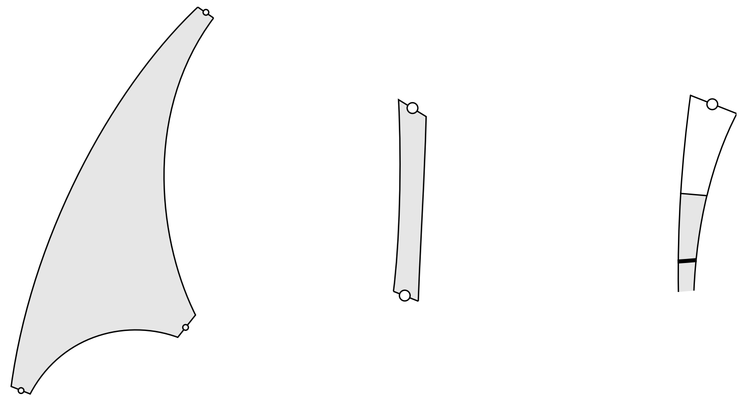


Figure 56. Regions of D which contain points of $f^{-1}(L)$.

of $f(\partial D)$. Thus we conclude

$$\operatorname{area}(f(D)) < 2C \cdot |f(D) \cap L| < C' \cdot \operatorname{length}(f(\partial D)).$$

□

A key feature of this proof that does not have an obvious analogue in contact topology is the crude notion of length given by pulling back $\mathcal{G}$ to ∂D.

We would like to end up by pointing out that there are no clear connections between tight contact structures on M and the fundamental group of M. For instance, it is not known if a homotopy 3-sphere supports a tight contact structure whether it must be S^3.

By way of contrast, there are many $\pi_1(M)$ actions that can be constructed from foliations and laminations. The *leaf space* is the quotient of the universal cover by leaves and complementary regions. The quotient is an *order tree*, and there is always an action of $\pi_1(M)$ on it.

Bestvina and Mess [2] show that if M is group negatively curved then there is an action of $\pi_1(M)$ on S^2. This can be applied to the manifolds of Theorem (6.6), and indeed by Calegari's work [3], there are far more manifolds in this collection than originally realized.

Palmeira's Theorem [27] is generalized to laminations in [16], and it follows that the universal cover $(\tilde{M}, \tilde{\lambda})$ is always homeomorphic to a product $(\mathbb{R}^2, \kappa) \times \mathbb{R}$ where κ is a lamination of the plane. Calegari and Dunfield [5] point out that $(\mathbb{R}^2, \kappa)$ can be thought of as $(\mathbb{H}^2, \kappa)$ and from this they can sometimes produce an action on S^1_∞. Calegari and Dunfield [5] have more general results. They generalize Thurston's work on the universal circle, and using Candel's theorem [4], they identify leaves of λ with $\mathbb{H}^2$, and they identify all S^1_∞'s coming from the $\mathbb{H}^2$'s to get a $\pi_1(M)$ action on S^1_{univ}. This works for taut foliations and some genuine laminations.

Acknowledgements

This article is based on a series of talks given at the Tokyo Institue of Technology from June 3 to June 7, 2002. I would like to thank Professor Matsumoto for inviting me and Professor Kojima for hosting my visit. I would particularly like to thank Ko Honda for his lectures during an informal seminar at the University of Georgia that introduced me to his point of view in contact topology. I would also like to thank Ko Honda and the referee for pointing out many corrections to the original manuscript.

Received March 14, 2003

Final version received November 26, 2003

University of Georgia
Athens
GA 30602
USA
will@math.uga.edu
http://www.math.uga.edu/˜will

References

[1] D. Bennequin, *Entrelacements et équations de Pfaff*, Astérisque **107–108** (1983), 87–161.
[2] M. Bestvina, G. Mess, *The boundary of negatively curved groups*, J. Amer. Math. Soc. **4** (3), (1991), 469–481.
[3] D. Calegari, *The geometry of $\mathbb{R}$-covered foliations*, Geom. Topol. **4** (2000), 457-515.
[4] A. Candel, *Uniformization of surface laminations*, Ann. Sci. École Norm. Sup. (4) **26** (4), (1993), 489–516.
[5] D. Calegari, N. Dunfield, *Laminations and groups of homeomorphisms of the circle*, Invent. Math. **152** (1), (2003), 149–204.
[6] T. Cofer, *A Class of Tight Contact Structures on $\Sigma_2 \times I$*, Ph.D. thesis, University of Georgia, (2002).
[7] V. Colin, *Recollement de variétés de contact tendues*, Bull. Soc. Math. France **127** (1), (1999), 43–69.
[8] V. Colin, *Sur la torsion des structures de contact tendues*, to appear in Ann. Sci. Éc. Norm. Sup.(4) **34** (2), (2001), 267–286.
[9] V. Colin, it Une infinité de structures de contact tendues sur les variétés toroïodales, Comment. Math. Helv. **76** (2001), 353–372.
[10] V. Colin, E. Giroux and K. Honda, *On the coarse classification of tight contact structures*, (preprint)
[11] Y. Eliashberg, *Classification of overtwisted contact structures on 3-manifolds*, Invent. Math. **98** (1989), 623–637.
[12] Y. Eliashberg, *Contact 3-manifolds twenty years since J. Martinet's work*, Ann. Inst. Fourier **42** (1992), 165–192.
[13] Y. Eliashberg and W. Thurston, *Confoliations*, University Lecture Series **13**, Amer. Math. Soc., Providence (1998).
[14] D. Gabai, *Foliations and the topology of* 3*-manifolds*, J. Diff. Geom. **18** (1983), 445–503.
[15] D. Gabai, *Quasi-minimal semi-Euclidean laminations in* 3*-manifolds*, Surveys in differential geometry, Vol. III Cambridge, MA, 1996), 195–242, *Int. Press. Boston, MA, 1998.*
[16] D. Gabai, W. Kazez, *Homotopy, isotopy and genuine laminations of* 3*-manifolds*, Geometric Topology, **1**, *AMS/IP* (1997), 123–138.
[17] D. Gabai, W. Kazez, *Group negative curvature for* 3*-manifolds with genuine laminations*, Geom. Topol. **2** (1998), 65–77.
[18] E. Giroux, *Convexité en topologie de contact*, Comment. Math. Helv. **66** (1991), 637–677.

[19] E. Giroux, *Structures de contact sur les variétés fibrées en cercles audessus d'une surface*, Comment. Math. Helv.**76** (2), (2001), 218–262.
[20] K. Honda, *On the classification of tight contact structures I*, Geom. Topol. **4** (2000), 309–368.
[21] K. Honda, *Gluing tight contact structures*, Duke Math. J. **115** (3), (2002), 435–478.
[22] K. Honda, W. Kazez and G. Matić, *Convex decomposition theory*, Internat. Math. Res. Notices (2002), 55–88.
[23] K. Honda, W. Kazez and G. Matić, *Tight contact structures on fibered hyperbolic 3-manifolds*, J. Differential Geom. **64** (2003), 305–358.
[24] K. Honda, W. Kazez and G. Matić, *On the Gabai-Eliashberg-Thurston Theorem*, Comment. Math. Helv. **79** (2004), 502–515.
[25] Y. Kanda, *The classification of tight contact structures on the 3-torus*, Comm. in Anal. and Geom. **5** (1997), 413–438.
[26] S. Makar-Limanov, *Tight contact structures on solid tori*, Trans. Amer. Math. Soc. **350** (1998), 1013-1044.
[27] C. Palmeira, *Open manifolds foliated by planes*, Ann. Math. (2) **107** (1), (1978), 109–131.
[28] W. Thurston, *A norm for the homology of 3-manifolds*, Mem. Amer. Math. Soc. **59** (339), (1986), 99–130.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

ON THREE-MANIFOLDS WITH BOUNDED GEOMETRY

MICHEL BOILEAU AND DARYL COOPER

ABSTRACT. In this note we combine some of Cheeger-Gromov's results [CG1,CG2,CG3] from the geometry of collapses of Riemannian 3-manifolds together with some three-dimensional topology to prove results which say that there are at most finitely many diffeomorphism classes of prime non-geometrizable three-manifolds which admit a metric of *bounded geometry* (i.e. with bounded sectional curvatures and bounded volume).

0. Introduction

Definition. A compact orientable 3-manifold is *geometrizable* if it has a splitting along a finite collection of disjoint essential spheres and tori into finitely many compact 3-manifolds whose interiors each admit a complete homogeneous riemannian metric (after capping off their boundary spheres by balls).

Thurston's geometrization conjecture states that all 3-manifolds are geometrizable.

There are eight homogeneous riemannian metric, which are locally modelled on the following 3-dimensional geometries: S^3, E^3, H^3, $S^2 \times E^1$, $H^2 \times E^1$, Nil, $\widetilde{SL_2(\mathbf{R})}$ and Sol.

A 3-manifold M is:

• **prime** if it is not the connect sum of two 3-manifolds neither of which is S^3.

• **irreducible** if every smoothly embedded sphere in M bounds a ball M.

• **∂-irreducible** if for every smooth properly embedded disc D in M there is a ball $B \subset M$ and a disc $D' \subset \partial M$ such that $\partial B = D \cup D'$.

• **atoroidal** if every $\mathbf{Z}^2$ subgroup in $\pi_1 M$ is conjugate into $\pi_1 \partial M$ and in addition $\pi_1 M$ does not contain the fundamental group of the klein bottle.

A prime orientable 3-manifold which is not irreducible is homeomorphic to $S^2 \times S^1$, and hence geometric. An irreducible orientable 3-manifold such that every $\mathbf{Z}^2$ subgroup of $\pi_1 M$ is conjugate into $\pi_1 \partial M$ is either atoroidal, or else the orientable I-bundle over the Klein bottle which is geometric.

By Thurston's hyperbolization theorem [Th2](cf. [Ka], [Ot1,2]) and the Torus theorem ([CJ],[Ga]), a non-geometrisable prime 3-manifold is irreducible, atoroidal and does not contain any embedded, incompressible, orientable surface. In particular it has an empty boundary.

2000 *Mathematics Subject Classification*: 57M50, 53C23.

Keywords and phrases: 3-manifold, geometric structure, graph manifold, F-structure, Dehn filling.

Research partially supported by the NSF.

Given a positive real number $v > 0$, let $\mathcal{M}(v)$ be the set of diffeomorphism classes of closed orientable 3-manifolds which admit a Riemannian metric g with bounded sectional curvature $|K_g| \leq 1$ and bounded volume $vol(M, g) \leq v$.

There are infinitely many geometrizable 3-manifold in $\mathcal{M}(v)$. In fact by the work of Cheeger and Gromov [CG1,CG2] all closed graph 3-manifolds (to be defined in §1) belong to $\mathcal{M}(v)$ for any $v > 0$, and this characterizes graph 3-manifolds. More precisely, there is a constant $v_0 > 0$ such that: $\forall v \leq v_0, \mathcal{M}(v) = \mathcal{M}(v_0)$ is the set of closed graph 3-manifolds.

By the work of Jørgensen and Thurston, for v sufficiently large (eg. bigger or equal to the hyperbolic volume of the figure eight knot complement), there are infinitely many closed hyperbolic 3-manifolds in $\mathcal{M}(v)$.

The main result of this note is the following finiteness result concerning non-geometrizable prime summands of 3-manifolds in $\mathcal{M}(v)$.

Theorem (0.1). *Given $v > 0$ there is only a finite set $\mathcal{NG}(v)$ of orientable, non-geometrizable 3-manifolds that may occur as a prime summand in the connected sum decomposition of a 3-manifold in $\mathcal{M}(v)$. Moreover for every 3-manifold in $\mathcal{M}(v)$ the number of such non-geometrizable prime summands is bounded above by a number $p(v)$ depending only on v.*

As straightforward corollaries we obtain:

Corollary (0.2). *There is a constant $n(v)$ depending only on v such that $\mathcal{M}(v)$ contains at most $n(v)$ prime 3-manifolds which are not geometrizable.*

Corollary (0.3). *There is a constant $s(v)$ depending only on v such that $\mathcal{M}(v)$ contains at most $s(v)$ homotopy spheres.*

Definition. For a compact orientable 3-manifold M, let $\text{Minvol}(M) = \inf\{vol(M, g)\}$ where g runs over all Riemannian metrics on $int(M)$ with bounded curvature $|K_g| \leq 1$.

Let $\mathcal{A}$ denote the set of compact orientable irreducible and atoroidal 3-manifolds, with zero Euler characteristic, and which do not admit a spherical metric (such a manifold is not a graph manifold). We denote by $\mathcal{H} \subset \mathcal{A}$ the subset of 3-manifolds which admit a complete hyperbolic structure of finite volume. By Thurston's hyperbolization theorem for Haken 3-manifolds, a manifold with non empty boundary in $\mathcal{A}$ belongs to $\mathcal{H}$. Thurston's geometrisation conjecture states that $\mathcal{H} = \mathcal{A}$.

When M admits a complete hyperbolic structure of finite volume g_0, a deep result, due to Besson-Courtois-Gallot [BCG] in the closed case and to Boland-Connel-Souto [BCS] in the cusp case, shows that the hyperbolic metric realizes the Minvol i.e. $\text{Minvol}(M) = Vol(M, g_0)$.

Since there is no graph 3-manifold in $\mathcal{A}$, it follows from Cheeger-Gromov's work [CG1,CG2] that a 3-manifold in $\mathcal{A}$ has a strictly positive Minvol. By Corollary (0.2), for a given value $v > 0$, the set $\{\text{Minvol} \leq v\} \cap \mathcal{A}$ contains at most finitely many prime, non-geometrizable 3-manifolds since they belong to $\mathcal{M}(v + 1)$.

Since the geometrizable 3-manifolds in $\mathcal{A}$ are exactly the subset $\mathcal{H}$ of hyperbolic 3-manifolds, the following result is a direct consequence of [BCG], [BCS], [Th,chap.5] and of Corollary (0.2). It shows that the set of values of the Minvol for manifolds in $\mathcal{A}$ behave like the set of volumes of hyperbolic manifolds.

Corollary (0.4). *The map* Minvol : $\mathcal{A} \to (0, +\infty)$ *is finite to one and the set of values* Minvol$(\mathcal{A}\backslash\mathcal{H})$ *is discrete. In particular the set of values* Minvol$(\mathcal{A})$ *is a well-ordered subset of* $\mathbf{R}_+$ *whose limit points coincide with the limit points of the subset* Minvol$(\mathcal{H})$.

There are two parts in the proof of Theorem (0.1). The first part (cf. §1) follows from Cheeger-Gromov's theory of collapses for riemannian manifolds with bounded sectional curvature. The second part (cf. §2) is a generalization of Thurston's hyperbolic Dehn filling theorem to the case of *graph-fillings*

1. Thick parts of Riemannian manifolds with bounded volume

A phenomenon which has received much attention in all dimensions from geometers is the notion of *collapse* : we say that a family of Riemannian metrics on a manifold collapses *with bounded geometry* if all the sectional curvatures remain bounded while the injectivity radius goes uniformly everywhere to zero.

For example any flat torus T^n collapses to any small dimensional torus T^k with $k < n$ by rescaling the metric on some of the S^1 factors.

Cheeger and Gromov [CG1,CG2] have proved that a necessary and sufficient condition for a manifold to have such a collapse with bounded geometry is the existence of a "generalized torus action" which they call an **F-structure**. F stands for "flat" in this terminology.

Intuitively an F-structure corresponds to different tori of varying dimension acting locally on finite coverings of open subsets of the manifold. Certain compatibility conditions on these local actions on intersections of these open subsets will insure that the manifold is partitioned into disjoint orbits of positive dimension. A precise definition of an F-structure can be given using the notion of sheaf of local groups actions, but we will not need it here.

A compact orientable 3-manifold M with an F-structure admits a partition into orbits which are circles and tori, such that each orbit has a saturated subset. A 3-manifold M has a *graph structure* in the sense of Waldhausen [Wa] and is a *graph manifold* if it can be obtained by glueing Seifert fiber spaces together along torus boundary components. These tori are not required to be incompressible. It follows from the definition of F-structure that such a partition corresponds to a *graph structure* on M (see [Ro1,§3]).

Another description of the family of all graph manifolds is that they are precisely those compact three manifolds which can be obtained, starting with the family of compact geometric *non-hyperbolic* three-manifolds, by the operations of connect sum and of glueing boundary tori together. Thus they arise naturally in both the Geometrization conjecture and in Riemannian geometry.

The aim of this section is to prove the following proposition which is true in any dimension:

PROPOSITION (1.1). *Let M be a closed Riemannian n-manifold with $|K_g| \leq 1$ and $vol(M, g) \leq v$. Then M has a decomposition $M = N \cup G$ into two compact n-submanifolds such that:*

• *G admits an F-structure such that $\partial N = \partial G$ is an union of orbits.*

• *N belongs, up to diffeomorphism, to a finite set $\mathcal{N}(n, v)$ of smooth, compact, orientable n-manifolds.*

Here is a straightforward corollary in dimension 3:

COROLLARY (1.2). *Every manifold $M \in \mathcal{M}(v)$ has a decomposition $M = N \cup G$ into two compact (maybe not connected) 3-submanifolds such that:*

• *G is a (maybe empty) graph manifold.*

• *N belongs, up to homeomorphism, to a finite set $\mathcal{N}(v)$ of compact orientable 3-manifolds with zero Euler characteristic.* □

Riemannian geometry takes an important part in the proof of Proposition (1.1). This proposition is the analogue in bounded variable curvature of Jørgensen's finiteness theorem [Thm 5.12], which states that all complete hyperbolic 3-manifolds of bounded volume can be obtained by surgery on a finite number of cusped hyperbolic 3-manifolds. The finiteness of hyperbolic manifolds with volume bounded above and injectivity radius bounded below is a precursor to Gromov's compactness theorem, while the Margulis lemma takes the place of the Cheeger-gromov thick/thin decomposition [CG2, Thm.0.1].

The following theorem is a precise version of Cheeger-Gromov's thick/thin decomposition (see [CFG, Thm.1.3 and 1.7] for a proof). We recall that the ε-thin part of a Riemannian n-manifold (M, g) is the set of points $\mathcal{F}(\varepsilon) = \{x \in M\, , inj(x, g) < \varepsilon\}$

THEOREM (1.3). *For each n, there is a constant μ_n, depending only on the dimension n, such that for any $0 < \varepsilon \leq \mu_n$ and any complete Riemannian n-manifold (M, g) with $|K_g| \leq 1$, there exists a Riemannian metric g_ε on M such that:*

(1) *The ε-thin part $\mathcal{F}(\varepsilon)$ of (M, g_ϵ) admits an F-structure compatible with the metric g_ε, whose orbits are all compact tori of dimension ≥ 1 and with diameter $< \varepsilon$.*

(2) *The Riemannian metric g_ε is ε-quasi-isometric to g and has bounded covariant derivatives of curvature, i.e. it verifies the following properties:*

• $e^{-\varepsilon} g_\varepsilon \leq g \leq e^{\varepsilon} g_\varepsilon$.

• $\|\nabla^g - \nabla^{g_\varepsilon}\| \leq \varepsilon$, *where ∇ and ∇^{g_ε} are the Levi-Civita connections of g and g_ε respectively.*

• $\|(\nabla^{g_\varepsilon})^k R_{g_\varepsilon}\| \leq C(n, k, \varepsilon)$, *where the constant C depends only on ε, the dimension n and the order of derivative k.*

Using Cheeger-Gromov's chopping theorem [CG3, Thm.0.1] one can prove the following:

Proposition (1.4). *For each integer $n \geq 2$, there are constants $\mu_n > 0$, $\Lambda_n > 0$, $\delta_n > 0$ and $c_n > 0$, depending only on n, such that for any closed Riemannian n-manifold (M, g) with $|K_g| \leq 1$, there is a metric g_n which is μ_n-quasi-isometric to g, with $|K_{g_n}| \leq \Lambda_n$ and a decomposition $M = N \cup G$ where:*

- *G is a compact n-submanifold which admits an F-structure compatible with g_n and $\partial N = \partial G$ is saturated.*
- *The injectivity radius for g_n at every point $x \in N$ verifies $inj(x, g_n) \geq \delta_n$.*
- *The second fundamental form of ∂N for the metric induced by g_n is bounded: $\|II^{g_n}_{\partial N}\| \leq c_n$.*
- *The volume $vol(\partial N, g_n) \leq c_n \cdot vol(M, g_n)$.*

Proof. We apply theorem (1.3) with the constant $\varepsilon = \mu_n$. So there is a metric g_n which is μ_n-quasi-isometric to g and such that $M = \mathcal{B}(\mu_n) \cup \mathcal{F}(\mu_n)$, where $\mathcal{B}(\mu_n) = \{x \in M\,, inj(x, g) \geq \mu_n\}$ and the μ_n-thin part $\mathcal{F}(\mu_n)$ admits an F-structure compatible with g_n. Moreover, since the covariant derivatives of the curvature of g_n have bounded norm by theorem (1.3), it follows that there is a constant $\Lambda_n > 0$ such that $|K_{g_n}| \leq \Lambda_n$. Therefore by the uniform decay of injectivity radius [GLP, Prop.8.22], there is a universal function $\phi_n(-,-)$, depending only on n, such that: $\forall x, x' \in M\,, inj(x', g_n) \geq \phi_n(inj(x, g_n), d_n(x, x'))$.

If $\mathcal{B}(\mu_n) = \emptyset$, we take $N = \emptyset$ and $G = M$.

We assume for the rest of the proof that $\mathcal{B}(\mu_n) \neq \emptyset$. We denote by d_n the distance on M associated with the metric g_n. Let $X \subset \mathcal{F}(\mu_n)$ be the set of points: $X = \{x \in \mathcal{F}(\mu_n)\,, d_n(x, \partial(\mathcal{B}(\mu_n)) \geq 1 + 2\mu_n\}$.

If $X = \emptyset$, then every point of M is at distance less than $2(1 + \mu_n)$ from a point of $\mathcal{B}(\mu_n)$. It follows from the uniform decay of injectivity radius that $inj(x, g_n) \geq \phi_n(\mu_n, 2(1 + \mu_n)) = \delta_n$ for every point $x \in M$. So we take $N = M$ and $G = \emptyset$.

If $X \neq \emptyset$, let $\mathcal{F}(X)$ be the union of all the orbits of points in X for the F-structure on $\mathcal{F}(\mu_n)$, compatible with g_n. It is a compact saturated subset of $\mathcal{F}(\mu_n)$. Since the diameter of the orbits of the F-structure is at most μ_n, it follows that $d_n(y, \partial(\mathcal{B}(\mu_n)) > 1$ for all point $y \in \mathcal{F}(X)$. In particular the closed tubular neighborhood of radius 1 around $\mathcal{F}(X)$, $T_1(\mathcal{F}(X))$, is contained in $\mathcal{F}(\mu_n)$. Since the local torus groups act by isometries, the equivariant form of Cheeger-Gromov's chopping theorem [CG3, Thm.0.1] (see also [Ro2, Thm.2.1]), shows that there is a compact n-submanifold $U \subset M$ with smooth boundary ∂U such that for some constant $c_n > 0$ depending only on n:

- $\mathcal{F}(X) \subset U \subset T_1(\mathcal{F}(X)) \subset \mathcal{F}(\mu_n)$ and U is saturated for the F-structure;
- $\|II^{g_n}_{\partial U}\| \leq c_n$;
- $vol(\partial U, g_n) \leq c_n \cdot vol(T_1(\mathcal{F}(X)), g_n) \leq c_n \cdot vol(M, g_n)$.

We set $G = U$ and $N = M \backslash \mathrm{int}(U)$. Since $X \subset U$, for every point $x \in N$ we have $d_n(x, \partial(\mathcal{B}(\mu_n)) \leq 2(1 + \mu_n)$. By the uniform decay of injectivity radius [GLP, Prop.8.22], we obtain as above that $inj(x, g_n) \geq \delta_n$ for every point $x \in N$. ∎

Proof of Proposition (1.1). By proposition (1.4), for some constants $\mu_n > 0$, $\Lambda_n > 0$, $\delta_n > 0$ and $c_n > 0$, depending only on n there is a metric g_n on M, which is μ_n-quasi-isometric to g, with $|K_{g_n}| \leq \Lambda_n$ and a decomposition $M = N \cup G$ such that:

• G is a compact n-submanifold which admits an F-structure compatible with g_n and $\partial N = \partial G$ is saturated.
• The injectivity radius for g_n at every point $x \in N$ verifies $inj(x, g_n) \geq \delta_n$.
• The second fundamental form of ∂N for the metric induced by g_n is bounded: $\|II^{g_n}_{\partial N}\| \leq c_n$.
• The volume $vol(\partial N, g_n) \leq c_n \cdot vol(M, g_n)$.

In particular, the volume of (M, g_n) verifies: $vol(M, g_n) \leq V(n, v)$ for a constant $V(n, v)$ depending only on μ_n and v, and thus only on n and v. Since $|K_{g_n}| \leq \Lambda_n$ and $inj(x, g_n) \geq \delta_n$ for every point $x \in N$, the diameter of N verifies: $diam(N, g_n) \leq D(n, v)$, where the constant $D(n, v)$ depends only on $v(n, v), \delta_n$ and Λ_n, and hence only on n and v.

To show that N belongs , up to diffeomorphism, to a finite set $\mathcal{N}(n, v)$ of smooth, compact, orientable n-manifolds, we use S. Kodani's extension [Ko] of Gromov's convergence theorem to some classes of Riemannian manifolds with boundary.

Let i_∂ be the infimum of inward normal injectivity radii of the boundary points of N. Then i_∂ is the infimum of the focal radius of ∂N and of half the length of a shortest geodesic which orthogonally intersects ∂N at the end points. (cf. [Ko, Lemma 6.3]). Let i_N be the minimum of i_∂ and the infimum of the injectivity radii of points at distance greater than i_∂ from ∂N. If $i_N < i_\partial$, then i_N is the infimum of the conjugate radii and of half the lengths of geodesic loops with base points at distance at least i_∂ from ∂N. In order to apply Kodani's results we need to have a lower bound on i_N, therefore we need to control the inward normal injectivity radius to ∂N. To do so the idea is to add a collar to ∂N. The following construction has been pointed out by J. Porti.

Since ∂N is a hypersurface in M, the uniform bounds $|K_{g_n}| \leq \Lambda_n$ and $\|II^{g_n}_{\partial N}\| \leq c_n$ imply that the focal radius of ∂N in M is bounded below by a constant $r_n = \frac{1}{\sqrt{\Lambda_n}} \arctan(\frac{\sqrt{\Lambda_n}}{c_n})$. Therefore the exponential map $exp : \nu^-_{\frac{r_n}{2}}(\partial N) \to M$ is a smooth immersion, where $\nu^-_{\frac{r_n}{2}}(\partial N)$ is the subspace of the normal bundle of ∂N which consists of normal vectors of length smaller or equal to $\frac{r_n}{2}$ and pointing ouside N. We use the exponential map to pull back the Riemannian metric g_n of M onto the collar $\nu^-_{\frac{r_n}{2}}(\partial N)$ of ∂N. We glue this collar to N along ∂N to get a Riemannian manifold N' with the same topological type as N and endowed with the metric g'_n which coincides with g_n on N and with the pull back metric on the collar $\nu^-_{\frac{r_n}{2}}(\partial N)$.

By [KO, Lemmas 3.1 and 3.2], see also [BZ, Chap. 6], the norm of the jacobian of the exponential map is uniformly bounded on $\nu^-_{\frac{r_n}{2}}(\partial N)$ above by a constant b_n and below by a constant $a_n > 0$, which depend only on Λ_n and c_n. It follows that the Riemannian metric (N', g'_n) has the following properties:
• $|K_{g'_n}| \leq \Lambda'_n$ where Λ'_n depends only on Λ_n, c_n, a_n, b_n, hence only on n.
• $\|II^{g'_n}_{\partial N'}\| \leq c'_n$, where the constant c'_n depends only on Λ_n, c_n, a_n, b_n by [KO, Lemma 3.1].
• $vol(N', g_n) \leq (1 + (a_n)^{-n}) vol(M, g_n) \leq (1 + (a_n)^{-n}) V(n, v) = V'(n, v)$.
• $i_{N'} \geq \delta'_n$, where δ'_n depends only on n, Λ'_n, c'_n, b_n and r_n, thus only on n . This follows from the uniform decay of injectivity radius in M, the uniform upper

bound on the jacobian of the exponential map and the uniform lower bounds on the conjugate radius of N' and focal radius of $\partial N'$.

• $diam(N', g'_n) \leq D'(n, v)$, since the volume of N' is bounded above by a constant $V'(n, v)$ and the injectivity radius of N' is bounded below by a constant δ'_n.

Therefore (N', g'_n) belongs to the class of n-dimensional compact Riemannian manifolds with bounded sectional curvature $|K_{g'_n}| \leq \Lambda'_n$ and a lower bound on the injectivity radius $i_{N'} \geq \delta'_n$. Moreover, if $\partial N' \neq \emptyset$, $\|II^{g'_n}_{\partial N'}\| \leq c'_n$. It follows from [GLP, Prop.7.5] and [Ko, Thm.A] in the case with boundary, that the Gromov-Hausdorff and the Lipschitz topology coincide for this class of manifolds. Furthermore $vol(N', g_n) \leq V'(n, v)$ and $diam(N', g_n) \leq D'(n, v)$, so the Riemannian manifold (N', g'_n) belongs to a class of riemannian manifolds which is precompact for the Gromov-Hausdorff topology by [GLP, Prop.5.2], and thus for the bilipschitz topology. It follows from the definition of the bilipschitz topology that there are, up to diffeomorphism, only finitely many manifolds in a precompact family with respect to this topology. Therefore there are, up to diffeomorphism, only finitely many manifolds N' and hence only finitely many manifolds N. □

2. Graph-fillings

Definition. A *graph-filling* of a compact orientable 3-manifold N is the operation of gluing a compact orientable (maybe not connected) graph 3-manifold G to N by identifying some toral components of ∂N with some toral components of ∂G.

A graph-filling is a generalization of a Dehn filling where each connected component of G is a solid torus.

Corollary (1.2) implies that every $M \in \mathcal{M}(v)$ either is a graph manifold, or belongs to $\mathcal{N}(v)$, or is obtained from a manifold in $\mathcal{N}(v)$ by a graph filling. Hence Theorem (0.1) is a straightforward consequence of Corollary (1.2) and the following result:

Proposition (2.1). *Let M be a compact orientable 3-manifold with non empty boundary a collection of tori. There is only a finite set $\mathcal{NG}(M)$ of compact, orientable, non-geometrizable 3-manifolds that may occur as prime factors of the connected sum decompositions of all the compact, orientable 3-manifolds obtained by graph fillings of M. Moreover the number of such prime factors (counted with multiplicity) is also bounded above by a constant depending only on M.*

The purpose of this section is to prove Proposition (2.1). Before starting the proof we give some definitions.

Definition. Let M be a compact orientable 3-manifold and let $T \subset \partial M$ be a boundary torus. A *slope* $\alpha \in H_1(T, \mathbf{Z})$ is a homology class corresponding to an essential simple closed curve on T. We denote by $M(\alpha)$ the compact orientable 3-manifold obtained by Dehn filling T with slope α i.e. by gluing a solid torus

$S^1 \times D^2$ along T in such way that the boundary of a meridian disk $\{*\} \times \partial D^2$ has slope α on T. By convention ∞ will denote the empty slope, so $M(\infty)$ means that no Dehn filling occurred along T.

Definition. Let V be a soli torus, a *cable space* is the complement of an open tubular neighborhood of a (r, s)-cable of the core of V, where r, s are coprime integers with $s \geq 2$. It has a Seifert fibration over an annulus with one single cone point.

Definition. A compact orientable 3-manifold H is *hyperbolicabled* if there is a finite (maybe empty) set of disjoint compact cable subspaces $C_1, \dots, C_k$ in H such that $C_i \cap \partial H$ is a torus component of ∂C_i, for $i = 1, \dots, k$, and that $H_0 = H \backslash \cup_{i=1}^k C_i$ is not empty and admits a complete hyperbolic metric of finite volume on its interior. When the family of cable subspaces $\{C_i\}_{i=1,\dots,k}$ is empty, the manifold H is said to be *hyperbolic*. Observe that a hyperbolicabled manifold is geometrizable.

The following lemma is a straightforward extension of Thurston's hyperbolic Dehn filling Theorem [Th1, Chap 5] :

Lemma (2.2). *Let H be a compact, orientable, hyperbolicabled 3-manifold , with q toral boundary components $T_1, \dots, T_q$. Then on each torus component $T_i \subset \partial H$ there is a finite exceptional set of slopes $\mathcal{S}_i$ such that for any collection of slopes $(\alpha_1, \dots, \alpha_q) \in (H_1(T_1, \mathbf{Z}) \cup \{\infty\} \backslash \mathcal{S}_1) \times \dots (H_1(T_q, \mathbf{Z}) \cup \{\infty\} \backslash \mathcal{S}_q)$, the 3-manifold $H(\alpha_1, \dots, \alpha_q)$ obtained by Dehn filling of H is irreducible, $\partial-$irreducible and geometrizable.*

Proof. Let $H_0 = H \backslash \cup_{i=1}^k C_i$ be the hyperbolic part of H, with $k \leq q$. By Thurston's hyperbolic Dehn filling theorem [Th1, Chap. 5], on each torus component $T'_i \subset \partial H_0$, $i = 1, \dots, q$, there is a finite exceptional set of slopes $\mathcal{S}'_i$ such that for any collection of slopes $(\beta_1, \dots, \beta_q) \in (H_1(T'_1, \mathbf{Z}) \cup \{\infty\} \backslash \mathcal{S}'_1) \times \dots (H_1(T'_q, \mathbf{Z}) \cup \{\infty\} \backslash \mathcal{S}'_q)$, the 3-manifold $H_0(\beta_1, \dots, \beta_q)$ obtained by Dehn filling of H_0 admits a complete hyperbolic structure of finite volume on its interior.

Let $T_i \subset \partial H$ be a boudary component. If $T_i = T'_i \subset \partial H_0$, then the exceptional set of slopes $\mathcal{S}_i = \mathcal{S}'_i$. Otherwise $T_i \subset \partial C_i$, where C_i is a cable subspace of H and $T'_i = \partial C_i \backslash T_i \subset H_0$.

If intersection number of the slope $\alpha \subset T_i$ with the fibre $f \subset T_i$ of the Seifert fibration of C_i is $|\Delta(\alpha, f)| \geq 2$, then the Dehn filled 3-manifold $C_i(\alpha)$ is a Seifert manifold over a disk, with two exceptional fibres and incompressible boundary. Hence gluing $C_i(\alpha)$ to a boundary component of an hyperbolic 3-manifold still yields an irreducible, $\partial-$irreducible and geometrizable 3-manifold.

If $|\Delta(\alpha, f)| = 1$, then $C_i(\alpha)$ is a solid torus. A homological calculation shows that the intersection numbers of two slopes β and β' on T'_i corresponding to the boundaries of meridian disks of $C_i(\alpha)$ and $C_i(\alpha')$ verifies: $|\Delta(\beta, \beta')| = s_i^2 |\Delta(\alpha, \alpha')|$, where $s_i \geq 2$ is the order of the exceptional fibre of C_i (cf. [Go, Lemma 3.3]). Then the existence of a finite exceptional set of slopes $\mathcal{S}'_i$ on $T'_i \subset \partial H_0$ implies the existence of a finite exceptional set of slopes $\mathcal{S}_i$ on T_i. $\square$

Let M be a compact irreducible and ∂-irreducible, orientable 3-manifold with non-empty boundary a finite collection of tori. Using the *JSJ-decomposition* it is easy to show that M contains a finite (possibly empty) minimal collection $\mathcal{T}$ of disjoint essential tori such that the closure of each component of $M \backslash \mathcal{T}$ is either a graph or a hyperbolicabled 3-manifold each of whose cable subspaces contains a boundary component of M. It is a subcollection of the *JSJ-family* of tori of M. One calls $\mathcal{T}$ the *reduced JSJ-family* of tori.

Let $T \subset \partial M$ be a torus component and let W_T be the closure of the connected component of $M \backslash \mathcal{T}$ containing T in its boundary.

Definition. A *bad* slope $\alpha \subset T$ is a slope such that either:
• W_T is a graph manifold and $W_T(\alpha)$ is either reducible, or ∂-compressible, or
• W_T is hyperbolicabled and α belongs to the exceptional set of slopes $\mathcal{S} \subset T$ given by the lemma (2.2).

The following is a generalization of the previous lemma (2.2).

LEMMA (2.3). *Let M be a compact, connected, orientable, irreducible and ∂ irreducible 3-manifold with non-empty boundary a finite collection of tori. Suppose also that M is not a cable space. Then on each torus component $T \subset \partial M$ there are only finitely many bad slopes.*

Proof. Let $\mathcal{T} \subset M$ be the reduced JSJ-family of tori and let W_T be the closure of the connected component of $M \backslash \mathcal{T}$ containing T.

We claim that W_T is not a cable-space. To see this, suppose that W_T is a cable space. Then $\partial W_T = T \cup T'$. If $T' \subset \partial M$ then since M is connected we have $M = W_T$, which contradicts our hypothesis. Otherwise T' is also a boundary component of some other component, C, of the reduced JSJ decomposition. By definition of reduced JSJ decomposition we see that C is not hyperbolic. Thus C is a graph manifold. But then $C \cup W_T$ is also a graph manifold which contradicts the minimallity of the collection $\mathcal{T}$ of tori in the reduced JSJ decomposition. This proves the claim. Thus if W_T is a graph manifold it is not a cable space hence by [CGLS,§2] there are only finitely many bad slopes on T.

Otherwise, when W_T is hyperbolicabled the set of bad slopes on T is finite by Lemma (2.2). □

Proof of Proposition (2.1). Every graph filling of a graph manifold is a graph manifold and hence has a geometric decomposition. Thus if M is a graph manifold the set $\mathcal{N}G(M)$ is empty. Hence we may assume that M is not a graph manifold. By considering the connected sum decomposition of M in prime factors, one reduces the proof of Theorem (2.1) to the case where M is irreducible and not a graph manifold. In particular M is not a solid torus and is ∂-irreducible.

Since any connected sum factor of a graph manifold is a graph manifold, we have only to consider graph fillings by irreducible graph manifolds. Moreover M is geometrizable because it is irreducible and $\partial M \neq \emptyset$, hence graph fillings by irreducible and ∂-irreducible, orientable graph manifolds always yield geometrizable 3-manifolds. Therefore we have only to deal with Dehn fillings by solid tori, because an orientable, irreducible 3-manifold with a compressible torus in its boundary is a solid torus.

Now we argue by induction on the number of boundary components of M.

If there is only one boundary component since M is irreducible and ∂-irreducible, Lemma (2.3) shows that except for finitely many bad slopes $\alpha \subset \partial M$ the Dehn filled 3-manifold $M(\alpha)$ is irreducible and geometrizable. This proves Theorem (2.1) in this case.

Let $T_1, \ldots, T_q$ be the boundary components of ∂M. By Lemma (2.3), except for a finite set of bad slopes $\mathcal{S}_i \subset T_i$ on each boundary torus, any collection of slopes $(\alpha_1, \ldots, \alpha_q) \in (H_1(T_1, \mathbf{Z}) \cup \{\infty\} \backslash \mathcal{S}_1) \times \ldots (H_1(T_q, \mathbf{Z}) \cup \{\infty\} \backslash \mathcal{S}_q)$, yields an irreducible and ∂-irreducible 3-manifold $M(\alpha_1, \ldots, \alpha_q)$ which is geometrizable.

For any bad slope $\beta_i \in \mathcal{S}_i \subset T_i$, the Dehn filled manifold $M(\beta_i) = M(\infty, \ldots, \alpha_i, \ldots, \infty)$ is compact orientable with strictly less boundary tori than M. From the discussion above, clearly $\mathcal{NG}(M) \subset \cup \mathcal{NG}(M(\beta_i))$, where the union is taken over the finite set of all bad slopes in $\cup_{i=1}^{q} \mathcal{S}_i$. Then $\mathcal{NG}(M)$ is finite since by the induction hypothesis the sets $\mathcal{NG}(M(\beta_i))$ are finite. In the same way the number of non-geometrizable prime factors for any graph filling of M is bounded above by the maximum of non-geometrizable prime factors for the graph fillings of the manifolds $M(\beta_i)$ where β_i runs over all bad slopes in $\cup_{i=1}^{q} \mathcal{S}_i$. $\square$

We can now prove the main theorem (0.1). By (1.2) there is a finite set $\mathcal{N}(v)$ of compact orientable 3-manifolds such that every $M \in \mathcal{M}(v)$ can be decomposed as $M = N \cup G$ with $N \in \mathcal{N}(v)$ and G a graph manifold. Then by (2.1) the set $\mathcal{NG}(N)$ is finite for each $N \in \mathcal{N}(v)$. The union of these finite sets as N varies over the finite set $\mathcal{N}(v)$ is $\mathcal{NG}(v)$ and is therefore finite. Furthermore the number of non-geometrizable prime summands is bounded by the maximum of the number of such summands that appear for any graph filling of any $N \in \mathcal{N}(v)$. Thus this bound, $p(v)$, depends only on the volume bound v. $\square$

Received May 28, 2003

Final version received April 02, 2004

Michel Boileau
Laboratoire Émile Picard
CNRS UMR 5580
Université Paul Sabatier
118 Route de Narbonne
F-31062 Toulouse
France
boileau@picard.ups-tlse.fr

Daryl Cooper
Mathematics Departement
University of California Santa Barbara
CA 93106-3080
USA
cooper@math.ucsb.edu

References

[BCG] G. Besson, G. Courtois, S. Gallot, *Entropies et rigidités des espaces localement symétriques de courbures strictement négative*, *Geom. Funct. Anal.* **5** (1995), 731–799.

[BCS] J. Boland, C. Connell, J. Souto, Volume rigidity for finite volume manifolds, Preprint (2002).

[BZ] Y.D. Burago and V.A. Zalgaller, *Geometric Inequalities*, A Series of Comprehensive Studies in Mathematics **285**, Springer-Verlag, Berlin, 1988.

[CJ] A. Casson and D. Jungreis, *Convergence groups and Seifert fibered* 3*-manifolds*, Invent. Math. **118** (1994), 441–456.

[CFG] J. Cheeger, K. Fukaya, M. Gromov, *Nilpotent structures and invariant metrics on collapsed manifolds*, J. Amer. Math. Soc. **5** (1992), 327–372.

[CG1] J. Cheeger, M. Gromov, *Collapsing Riemannian manifolds while keeping their curvature bounded*, part I: J. Differential Geom. **23** (1986), 309–346.

[CG2] J. Cheeger, M. Gromov, *Collapsing Riemannian manifolds while keeping their curvature bounded*, part II: J. Differential Geom. **32** (1990), 269–298.

[CG3] J. Cheeger, M. Gromov, Chopping Riemannian manifolds, Pitman Monographs Surveys Pure Appl. Math. Longman Sci. Tech., Harlow, **52** (1991), 85–94.

[CGLS] M. Culler, C. McA. Gordon, J. Luecke and P. B. Shalen, *Dehn surgeries on knots*, Ann. of Math. (2) **125** (1987), 237–300.

[Ga] D. Gabai, *Convergence groups are Fuchsian groups*, Ann. Math. **136** (1992), 447–510.

[GLP] M. Gromov, J. Lafontaine, P. Pansu, *Structures métriques pour les variétés riemanniennes*, CEDIC, Paris, 1981.

[Go] C. McA. Gordon, *Dehn surgery and satellite knots*, *Trans. Amer. Math. Soc.* **275** (1983), 687–708.

[Ka] M. Kapovich, *Hyperbolic manifolds and discrete groups*, Progress in Mathematics, **183**, Birkhauser Boston, Inc., Boston, MA, 2001.

[Ko] S. Kodani, *Convergence theorem for Riemannian manifolds with boundary*, Compositio Math. **75** (1990), 171–190.

[Ot1] J. P. Otal, *Le théorème d'hyperbolisation pour les variétés fibrées de dimension 3*, *Astérisque* **235** (1996).

[Ot2] J. P. Otal, *Thurston's hyperbolization theorem of Haken manifolds*, Preprint (1997).

[Ro1] X. Rong, *The limiting eta invariant of collapsed 3-manifolds*, J. Diff. Geom. **37** (1993), 535–568.

[Ro2] X. Rong, *Rationality of geometric signatures of complete 4-manifolds*, Invent. Math. **120** (1995), 513–554.

[Th1] W. P. Thurston, *The geometry and topology of* 3*-manifolds*, Princeton Math. Dept., 1979.

[Th2] W. P. Thurston, *Hyperbolic structures on 3-manifolds, I: Deformations of acylindrical manifolds*, Ann. of Math. (2) **124** (1986), 203–246.

[Wa] F. Waldhausen, *Eine Klasse von 3-dimensionalen Mannigfaltigkeiten*, *Invent. Math.* **3–4** (1967), 308–333, 87–117.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

CONWAY POLYNOMIALS OF THE CLOSURES OF ORIENTED 3-STRING TANGLES

HUGO CABRERA-IBARRA

Abstract. Given a certain type of oriented 3-string tangle, we consider five different ways for closing it to obtain knots or links, and give formulas for calculating the Conway polynomials of the closures of the composition of two such 3-tangles. We also give a certain relation among these polynomials.

1. Introduction

Tangles were introduced by Conway [2] as basic building blocks for the construction of knots. In this article we analyze the relation between the Conway polynomials associated to the closure of a skein element s of a 3-room (which is a connected domain with three ingoing and three outgoing strands, a skein element in a 3-room can be viewed as a 3-tangle with orientation on its strands) and the ones associated to the composition $s_1 \cdot s_2$ of two of such skein elements.

We assign to s the 2×2 matrix $M_\nabla(s)$ whose entries are the Conway polynomials of certain closures of s; in each of those closures no other crossings are added. It satisfies $M_\nabla(s_1 \cdot s_2) = M_\nabla(s_1)M_\nabla(s_2)$.

In [6], Giller made analogous computations in the case of 2-rooms and formulas to compute the Conway polynomial of the numerator and denominator of the composition of two 2-string oriented tangles were obtained; Giller pointed out that similar computations could be made in the case 3-rooms.

In [1], another matrix M associated to 3-string tangles without orientation was obtained; certain relations between these two matrices suggest that an analysis of $M_\nabla(s)$ will give insight into the comprehension of 3-tangles. Some attempts to classify the set of 3-string tangles have been made in [1], [3].

The study of 3-string tangles will be useful to analyze certain enzymes called recombinases, as it was in the case of 2-string tangles [4], [5], [8]. In DNA site-specific recombination, a recombination enzyme attaches to a pair of DNA sites, breaks both sites, and recombines the sites to different ends. Electron micrographs of recombinases bound to DNA show the enzyme as a blob with 2 or 3 loops of DNA sticking out of this blob. In the case of the Gin enzyme there are three loops of DNA, thus the mathematics of 3-tangles can be useful in the study of this enzyme.

2000 *Mathematics Subject Classification*: Primary 57M25; Secondary 57M27.

Keywords and phrases: knots, Tangles, Conway Polynomial.

2. Preliminaries

Remember that the Conway polynomial of an oriented link is computed by the following recursive formulas:

i) $\nabla_{L_l}(z) = \nabla_{L_r}(z) + z\nabla_{L_s}(z)$

ii) $\nabla_{\bigcirc}(z) = 1$

where (L_l, L_r, L_s) is a skein triple of oriented knots or links that are the same, except in a crossing neighborhood where they look as in Fig. (1).

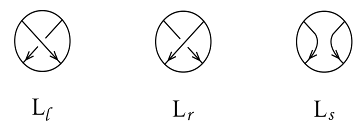


Figure 1. Skein triple

An n-room is a connected domain (usually a rectangle) in $\mathbb{R}^2$, with n ingoing and n outgoing strings; the room may contain oriented simple closed curves. We will only deal with 3-rooms. In Fig. (2) a) there is an example of a room.

The skein of a room R, $S(R)$, is the set of all collections of strands in the room which connect ingoing to outgoing strands; Fig. (2) b) shows an example of a skein element of the room in Fig. (2) a); this element can also be seen as a 3-string tangle with orientations on its strands. In this discussion we will take R to be the room in Fig. (2) c). Given $s_1, s_2 \in S(R)$, we say that $s_1 = s_2$ if there exists an ambient isotopy which carries s_1 into s_2.

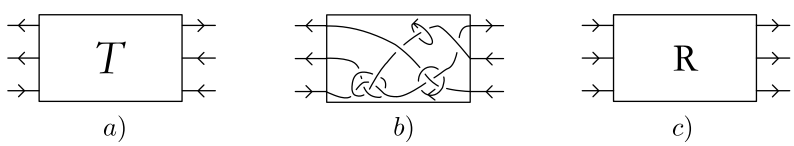


Figure 2. a) An example of a room, b) a skein element, c) the room R.

Let S_3 denote the full symmetric group on 3 letters, given a skein element s, we assign a permutation $\pi(s) \in S_3$ by numbering the strands of R as in Fig. (3).

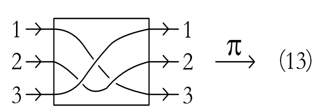


Figure 3. The action of π.

Now pick an ordering $\{\pi_i\}_{i=1,\dots,6}$ of S_3 and fix a choice of skein elements $\{s_{\pi_i}\}_{i=1,\dots,6}$, as in Fig. (4), such that $\pi(s_{\pi_i}) = \pi_i$, and such that s_{π_i} contains no free components. Define on $S = S(R)$ a binary operation by juxtaposition as in Fig. (5); note that $\pi(s_1 \cdot s_2) = \pi(s_1)\pi(s_2) \in S_3$.

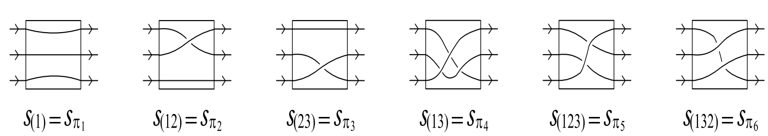


Figure 4. The selected skein elements.

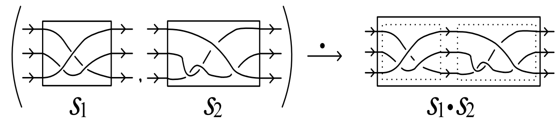


Figure 5. The operation of juxtaposition.

We will call (s_l, s_r, s_s) a skein triple, where s_l is a skein element in which a left-handed crossing appears, s_r (respectively s_s) is the same element with the same crossing changed to a right-handed (respectively smoothed) crossing.

Let F be the quotient field of $\mathbb{Z}[z]$, $V(S)$ the vector space generated by S over F, and $N(S)$ the vector space generated by $\{s_l - s_r - zs_s \,|\, (s_l, s_r, s_s) \text{ a skein triple}\}$ over F. Define the vector space $L(S) = V(S)/N(S)$; moreover, $L(S)$ is an algebra under the extension of $\cdot$ to $L(S)$ given by $s_1 \cdot (s_2 + \alpha s_3) = s_1 \cdot s_2 + \alpha s_1 \cdot s_3$, $\alpha \in F$, which preserves the relations in $N(S)$.

In [6], the following two results have been proved.

Theorem (2.1). *The set* $\{s_{\pi_j} \mid j = 1, \dots, 6\}$ *is a basis for* $L(S)$.

Corollary (2.2). *Any skein element* $s \in L(S)$ *can be expressed uniquely as a linear combination of the* s_{π_j}, *and therefore* $\dim_F L(S) = 6$.

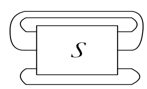


Figure 6. The closure N of s.

For $s \in S$, define $N(s)$ as the knot (or link) obtained by closing s as in Fig. (6). Denote by s^N the Conway polynomial of $N(s)$. Given elements s and x in S we define $s^*(x) = \nabla(N(x \cdot s)) = (x \cdot s)^N$, as before, s^* may be extended linearly to all of $L(S)$ obtaining the dual $s^* : L(S) \to \mathbb{Z}[z] \subset F$. Note that s^* preserves skein moves: $s^*(s_r + zs_s) = s^*(s_r) + zs^*(s_s) = s^*(s_l)$.

Let M be the 6×6 matrix defined by $M_{ij} = s_{\pi_i}^*(s_{\pi_j})$; then we have that

$$M = \begin{pmatrix} 0 & 0 & 0 & z & 1 & 1 \\ 0 & 0 & 1 & 1+z^2 & z & z \\ 0 & 1 & 0 & 1+z^2 & z & z \\ z & 1+z^2 & 1+z^2 & 3z^2+z^4 & 2z+z^3 & 2z+z^3 \\ 1 & z & z & 2z+z^3 & 1+z^2 & z^2 \\ 1 & z & z & 2z+z^3 & z^2 & 1+z^2 \end{pmatrix}.$$

It can be seen that $\det M = -(z^2+4) \neq 0$, and therefore $\{s_{\pi_i}^*\}$ is a basis for $L^*(S)$.

Let us define a bilinear form $\varphi : L(S) \times L(S) \to F$ by $\varphi(s_1, s_2) = (s_1 \cdot s_2)^N$. Since $\{s_{\pi_i}^*\}$ is a basis for $L^*(S)$ there exist $a_{kl} \in F$ such that $\varphi = \sum a_{kl} s_{\pi_k}^* \otimes s_{\pi_l}^*$.

Then

$$\begin{aligned}\varphi(s_{\pi_i}, s_{\pi_j}) &= \sum a_{kl} s_{\pi_k}^{\;*}(s_{\pi_i}) s_{\pi_l}^{\;*}(s_{\pi_j}) = \sum a_{kl} M_{ki} M_{lj} \\ &= \sum_l \left(\sum_k a_{kl} M_{ki} \right) M_{lj} = \left((A^T M)^T M \right)_{ij} = (M^T A M)_{ij}\,,\end{aligned}$$

where $A_{kl} = a_{kl}$. Since $\varphi(s_{\pi_i}, s_{\pi_j}) = M_{ji}$ we have that $M^T = M^T A M$, and therefore $\varphi = A = M^{-1}$.

Given $\sigma_1, \sigma_2 \in L(S)$, define $u_j = \left(s_{\pi_1}^{\;*}(\sigma_j) \;\; s_{\pi_2}^{\;*}(\sigma_j) \; ... \; s_{\pi_6}^{\;*}(\sigma_j) \right)$, $j = 1, 2$; then we have

$$\varphi(\sigma_1, \sigma_2) = u_1 M^{-1} u_2^T\,, \tag{2.3}$$

where

$$M^{-1} = \frac{-1}{z^2+4} \begin{pmatrix} -z^4 - 3z^2 + 2 & z^3 + 3z & z^3 + 3z & z & -z^2 - 2 & -z^2 - 2 \\ z^3 + 3z & 2 & -z^2 - 2 & -2 & z & z \\ z^3 + 3z & -z^2 - 2 & 2 & -2 & z & z \\ z & -2 & -2 & 2 & -z & -z \\ -z^2 - 2 & z & z & -z & -2 & z^2 + 2 \\ -z^2 - 2 & z & z & -z & z^2 + 2 & -2 \end{pmatrix}.$$

3. Computations for another room

Similar formulas can be derived for any room. For example, let us make similar computations for the room in Fig. (7), which we will denote by R'. Let s_1 and s_2 be elements of the skein of this room; as before we define $\cdot$ as in Fig. (8) and a bilinear form $\psi(s_1, s_2) = \nabla(N(s_1 \cdot s_2))$. Now, for s_1 and s_2 define, respectively, skein elements σ_1 and σ_2 of our previous room R as it is shown in Fig. (9). It is easy to see that $\psi(s_1, s_2) = \varphi(\sigma_1, \sigma_2)$.

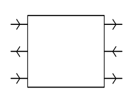


Fig. 7. The room R'.

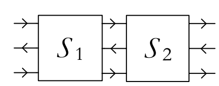
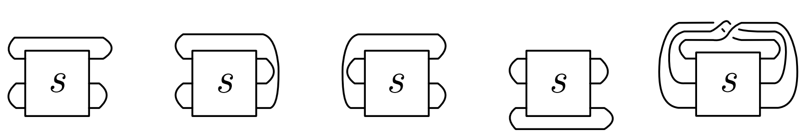


Fig. 8. The operation $s_1 \cdot s_2$.

σ_1 σ_2

Fig. 9. The associated skein elements in R.

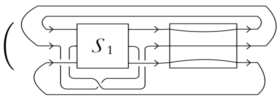
We define $A(s)$, $B(s)$, $C(s)$, $D(s)$, and $E(s)$ to be the knots (or links) obtained by closing $s \in S(R')$ as in Fig. (10). We will denote by s^A, s^B, s^C, s^D, and s^E the Conway polynomials associated to $A(s), B(s), C(s), D(s)$, and $E(s)$ respectively.



$A(s)$ $B(s)$ $C(s)$ $D(s)$ $E(s)$

Figure 10. The closures A, B, C, D, and E of s.

By drawing links one can see that

$s^*_{\pi_1}(\sigma_1) = \nabla\Big($  $\Big) = s_1^N;$

in a similar way, the following formulas are obtained:

$$u_j = v_j V_j, \qquad\qquad j = 1, 2\,,$$

where u_j is as in Eq (2.3), $v_j = \big(s_j^A \;\, s_j^B \;\, s_j^C \;\, s_j^D \;\, s_j^E \;\, s_j^N\big)$, and

$$V_1 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & z \\ 0 & 0 & 0 & 2z & 1 & -z^2 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & z \\ 0 & 0 & 0 & 1 & 0 & -z \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad V_2 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & z \\ 0 & 0 & 0 & 2z+z^3 & 1+z^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & z \\ 0 & 0 & 0 & 1 & & \\ z^2 & z & 0 & & & \\ 1 & z & z & 0 & 0 & z^2 \end{pmatrix}.$$

Then, by the formula obtained in Eq (2.3), we have that

$$\psi(s_1, s_2) = \varphi(\sigma_1, \sigma_2) = u_1 M^{-1} u_2^T.$$

Since $u_j = v_j V_j$ then

$$\psi(s_1, s_2) = u_1 M^{-1} u_2^T = v_1 \; V_1 M^{-1} V_2^T \; v_2^T.$$

This lead us to the following proposition.

Proposition (3.1). *For any $s_1, s_2 \in S(R')$,*

$$(s_1 \cdot s_2)^N = v_1 \big(V_1 M^{-1} V_2^T\big) v_2^T\,, \tag{3.2}$$

where v_j, V_j, for $j = 1, 2$, are defined as above.

In order to compute $(s_1 \cdot s_2)^A$, $(s_1 \cdot s_2)^B$, $(s_1 \cdot s_2)^C$, and $(s_1 \cdot s_2)^D$ we proceed as follows. Since $(s_1 \cdot s_2)^D$ can be expressed as $(t_1 \cdot t_2)^N$, for t_1 and t_2 as in Fig. (11), then $t_1 = s_1$, $t_2^A = s_2^B$, $t_2^B = 0$, $t_2^C = s_2^D$, $t_2^D = 0$, $t_2^E = s_2^B$, and $t_2^N = s_2^S$. Using Proposition (3.1) we obtain:

$$(s_1 \cdot s_2)^D = (t_1 \cdot t_2)^N = v_1 \big(V_1 M^{-1} V_2^T\big) \left(s_2^B \;\; 0 \;\; s_2^D \;\; 0 \;\; s_2^B \;\; s_2^D\right)^T = s_1^D s_2^B + s_1^C s_2^D.$$

Figure 11. Expressing $(s_1 \cdot s_2)^D$ as $(t_1 \cdot t_2)^N$.

Similarly, we have that

$$(s_1 \cdot s_2)^A = s_1^B s_2^A + s_1^A s_2^C, \quad (s_1 \cdot s_2)^B = s_1^A s_2^D + s_1^B s_2^B, \quad (s_1 \cdot s_2)^C = s_1^C s_2^C + s_1^D s_2^A.$$

Note that the equations obtained for $(s_1 \cdot s_2)^A$, $(s_1 \cdot s_2)^B$, $(s_1 \cdot s_2)^C$, and $(s_1 \cdot s_2)^D$ can also be obtained by using a formula of Giller in Proposition 15 of [6].

Given an element $s \in S(R')$ we will assign to it the following matrix, which is an invariant for s because of the properties of the Conway polynomial.

$$M_\nabla(s) = \begin{pmatrix} s^C & s^D \\ s^A & s^B \end{pmatrix} .$$

Then we obtain the following.

PROPOSITION (3.3). *If $s_1, s_2 \in S(R')$, then $M_\nabla(s_1 \cdot s_2) = M_\nabla(s_1)M_\nabla(s_2)$.*

Proof. Since

$$M_\nabla(s_1)M_\nabla(s_2)=\begin{pmatrix} s_1^C & s_1^D \\ s_1^A & s_1^B \end{pmatrix}\begin{pmatrix} s_2^C & s_2^D \\ s_2^A & s_2^B \end{pmatrix}=\begin{pmatrix} s_1^C s_2^C+s_1^D s_2^A & s_1^D s_2^B+s_1^C s_2^D \\ s_1^B s_2^A+s_1^A s_2^C & s_1^A s_2^D+s_1^B s_2^B \end{pmatrix} ,$$

we obtain $M_\nabla(s_1)M_\nabla(s_2)=M_\nabla(s_1 \cdot s_2)$. ☐

Remark. In [1], a matrix that is associated to a 3-string tangle without orientation has been given. Let us sketch how to obtain it.

The Kauffman bracket can be seen as a function from 3-string tangle diagrams without orientation on its strands to Laurent polynomials; it is characterized in [7], where it is shown that the bracket polynomial is an invariant of links under regular isotopy. Let T be a tangle diagram; we define the *bracket of* T as follows:

$$\langle T\rangle = \alpha(T)\langle\hat\alpha\rangle + \beta(T)\langle\hat\beta\rangle + \delta(T)\langle\hat\delta\rangle + \chi(T)\langle\hat\chi\rangle + \psi(T)\langle\hat\psi\rangle \,, \tag{3.4}$$

where $\langle T\rangle$ is obtained by applying to the diagram T the formulas which define the bracket polynomial repeatedly, until only the five tangles given in Eq. (3.4) are left. Here $\hat\alpha$, $\hat\beta$, $\hat\delta$, $\hat\chi$, and $\hat\psi$ denote the tangles shown in Fig. (12), the coefficients $\alpha(T)$, $\beta(T)$, $\delta(T)$, $\chi(T)$, and $\psi(T)$ are polynomials in a and a^{-1} that are invariant under regular isotopy, and the brackets $\langle\hat\alpha\rangle$, $\langle\hat\beta\rangle$, $\langle\hat\delta\rangle$, $\langle\hat\chi\rangle$, and $\langle\hat\psi\rangle$ are place holders for the result of bracket's computation restricted to the tangle diagram T.

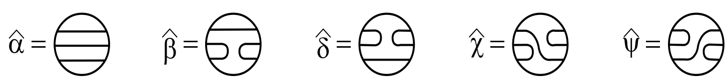


Figure 12. Special tangles

We assign to T the matrix

$$M(T)(a,a^{-1})=\begin{pmatrix} \alpha(T)+\chi(T) & \beta(T) \\ \delta(T) & \alpha(T)+\psi(T) \end{pmatrix} ; \tag{3.5}$$

since the bracket polynomial is an invariant under regular isotopy, $M(D)(a,a^{-1})$ possesses the same property.

Given A and B two 2×2 matrices, we define the equivalence relation: $A \sim B$ if and only if $A = (-a^{-3})^k B$ for some $k \in \mathbb{Z}$. With this relation, the equivalence class $[M(T)(a,a^{-1})]$, which we will write as $M(T)$, is an invariant of the tangle T.

For this matrix we have

THEOREM (3.6). *Given two tangle diagrams T_1 and T_2 we have that*

$$M(T_1 \cdot T_2) = M(T_1)M(T_2) + d\begin{pmatrix} \chi_1 & \beta_1 \\ \delta_1 & \psi_1 \end{pmatrix}\begin{pmatrix} \delta_2 & \psi_2 \\ \chi_2 & \beta_2 \end{pmatrix} ,$$

where $d = -(a^2 + a^{-2})$ *and* $M(T_j) = \begin{pmatrix} \alpha_j + \chi_j & \beta_j \\ \delta_j & \alpha_j + \psi_j \end{pmatrix}$ *for* $j = 1, 2$.

Evaluate $M(T)$ at $a = \sqrt{i}$ and denote this matrix by $M_1(T)$; then we obtain that

$$M_1(T_1 \cdot T_2) = M_1(T_1)M_1(T_2)\,. \tag{3.7}$$

Note that if, as before, we close our 3-string tangle T in four different ways then

$$\langle T^A \rangle = \delta, \qquad \langle T^B \rangle = \alpha + \psi, \qquad \langle T^C \rangle = \alpha + \chi, \qquad \langle T^D \rangle = \beta\,.$$

Where $\langle T^A \rangle$ (respectively $\langle T^B \rangle$, $\langle T^C \rangle$, and $\langle T^D \rangle$) is the bracket polynomial associated to the knot or link T^A (respectively T^B, T^C, and T^D).

We will assign to T a new matrix M_K, which involves the bracket polynomials of some closures of T,

$$M_K(T) = \begin{pmatrix} \langle T^C \rangle & \langle T^D \rangle \\ \langle T^A \rangle & \langle T^B \rangle \end{pmatrix}\,.$$

By Eq. (3.7), $M_K(T_1 \cdot T_2) = M_K(T_1)M_K(T_2)$; compare this with the formula for $M_\nabla(s_1 \cdot s_2)$.

Note that, although we have this relation, if we define $\alpha(T) = \frac{\langle T^B \rangle + \langle T^C \rangle - \langle T^N \rangle}{2}$, from the proof of Theorem 2.1 in [1], it follows that $\alpha(T_1 \cdot T_2) = \alpha(T_1)\alpha(T_2)$. However, we do not have an analogous relation for the Conway polynomial, i.e., if we define $\omega(s) = \frac{s^B + s^C - s^N}{2}$ we do not have that $\omega(s_1 \cdot s_2) = \omega(s_1)\omega(s_2)$.

A generalization of this work could be to find similar relations in the case of 4-rooms, which could give us an approach of the case n-rooms and, as we point out in our remark, an alternative way to obtain the classification of rational 3-tangles.

The author wishes to thank the anonymous reviewers for their contribution to readability and clarity. This paper was partially supported by CONACYT under grant 39579-F.

Received July 01, 2003

Final version received February 02, 2004

Departamento de Matemáticas Aplicadas y
Sistemas Computacionales
IPICYT
Apdo. Postal 3-74
Tangamanga
San Luis Potosí, S.L.P.
México
cabrera@ipicyt.edu.mx

References

[1] H. Cabrera-Ibarra, *On the classification of rational 3-tangles*, J. Knot Theory Ramifications **12** (2003), 921–946.

[2] J. Conway, An enumeration of knots and links and some of their related properties, in Computational Problems in Abstract Algebra, Proc. Conf., Oxford (1967), 329–358. New York: Pergamon Press, (1970).

[3] J. Emert, C. Ernst, *N-string tangles*, J. Knot Theory Ramifications **9** (2000), 987–1004.

[4] C. Ernst, D. Sumners, *A calculus for rational tangles: applications to DNA recombination*, Math. Proc. Cambridge Philos. Soc. **108** (1990), 489–515.

[5] C. Ernst, D. Sumners, *Solving tangle equations arising in a DNA recombination model*, Math. Proc. Cambridge Philos. Soc. **126** (1999), 23–36.

[6] C. A. Giller, *A family of links and the Conway calculus*, Trans. Amer. Math. Soc. **270** (1982), 75–109.

[7] J. R. Goldman, L. H. Kauffman, *Rational tangles*, Adv. in Appl. Math. **18** (1997), 300–332.

[8] D. W. Sumners, C. Ernst, S. Spengler, N. Cozzarelli, *Analysis of the mechanisms of DNA recombination using tangles*, Q. Rev. Biophys. **28** (1995), 253–313.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

ARTIN PRESENTATIONS OF COMPLEX SURFACES

J. S. CALCUT AND H. E. WINKELNKEMPER

Abstract. We construct Artin Presentations of infinitely many complex surfaces. Namely, for all elliptic surfaces $E(n)$, in particular for the Kummer surface $K3$. Thus, not only does AP Theory contain an analogue of Donaldson's Theorem, but also a purely group-theoretic theory of Donaldson and Seiberg-Witten invariants.

Not surprisingly, our explicit Artin presentations for the Kummer surface are approachable with a computer using, say, MAGMA and provide a plethora of interesting examples pertaining to knot theory in $\mathbb{Z}-$homology $3-$spheres.

1. Introduction

In the purely group-theoretic theory of Artin Presentations, a smooth, compact, connected, simply-connected 4-manifold $W^4(r)$ with a connected boundary $\partial W^4(r) = M^3(r)$ is already determined, and can be reconstituted, from a certain presentation (an Artin Presentation) of the fundamental group of its boundary [W1]. If the boundary is S^3 then of course the Artin Presentation presents the trivial group. Even in this case the Artin Presentation already encodes *all of the smooth structure* of the 4-manifold. Thus, it makes sense to ask whether an *arbitrary*, smooth, *closed*, connected, simply-connected 4-manifold is given by an Artin Presentation.

We extend important work of Harer, Kas and Kirby [HKK] and show that all elliptic surfaces $E(n)$ admit Artin Presentations. *This gives the first bridge between AP theory and algebraic geometry.* These Artin Presentations are of special interest due to the fact that complex algebraic surfaces possess nontrivial Donaldson invariants. In particular, this augments the remarkable fact (Theorem 1 of [W1], [R] p.621) that Donaldson's Theorem, despite being proved with gauge theory/connections (i.e. the smooth continuum), persists and survives the radical, discrete, purely group theoretic holography of AP Theory.

The following illustrates the AP theory program concerning the computation of Seiberg-Witten and Donaldson invariants and shows that the group theoretic AP encoding goes much deeper than e.g. the mere encoding of a group through its presentation:

Recall González-Acuña's formula, [CS] p.66, for the Rohlin invariant of a $\mathbb{Z}$-homology 3-sphere $\Sigma^3(r)$ given by an Artin Presentation $r \in \mathcal{R}_n$ (for clarity we consider here only the case where $A(r)$ is the identity matrix, see section 2.1 for

2000 *Mathematics Subject Classification*: Primary 14J27, 57M05, 57M07, 57R57; Secondary 57R60, 57M27.

Keywords and phrases: Artin presentation, elliptic surface, Donaldson invariants.

notation):

$$\mu\left(\Sigma^3(r)\right) = \frac{d^2-1}{8} \bmod 2,$$

where $d = \triangle(-1)$, $\triangle$ being the Alexander polynomial of the associated presentation:

$$\langle x_1, \ldots, x_n \mid x_1 r_1 = r_1 x_2, x_2 r_2 = r_2 x_3, \ldots, x_{n-1} r_{n-1} = r_{n-1} x_n \rangle,$$

where the group obviously abelianizes to $\mathbb{Z}$.

This remarkable formula is entirely from the discrete theory of finitely presented groups: there is no need to mention cobordisms, spin structures, skein methods, Heegaard decompositions, representations into $SU(2)$, Riemannian metrics, infinite dimensional or moduli spaces, or indeed even the smooth continuum, nor do any metric dependence, wall crossing, or word problems arise here.

We remark that González-Acuña's formula already shows that an analogue of Floer theory should also appear in AP theory since the Rohlin invariant is the Euler characteristic (mod 2) in Floer theory. In fact, we suspect that 'the 8 of González-Acuña is the 8 of Floer'.

Concerning the importance of relating Donaldson and Floer theory, both mathematically and physically, see [D] p.63 and [Wi1] p.352.

Consider the more general problem concerning the relative Donaldson invariants [TB],[Wi1] of $W^4(r)$ which, when $A(r)$ is unimodular, take values in the Floer homology of $\partial W^4(r) = \Sigma^3(r)$.

The computational program of AP theory can be stated as: these invariants and others should be computed solely in function of the Artin Presentation r in the discrete theory of finitely presented groups, just as, with González-Acuña's formula, this was done for the Rohlin invariant of $\Sigma^3(r)$.

This is entirely in the purely group-theoretic spirit of the Princeton School of Artin, Fox, Lyndon, Papakyriakopoulos, Stallings, et al. and extends their approach, as far as $3D/4D$ manifold theory is concerned, to its natural metamathematical boundary.

Immediate natural, important general questions arise (both mathematical and physical):

1. *Since AP theory dispenses not only with metrics but even topology, what becomes of Witten's celebrated Feynmanian formulation of Donaldson's invariants as correlation functions/expectation values* [D] *p.53,* [Wi2], [Wi3], [AJ], [Di] *pp.36,39? What is the* topologically independent *(i.e. purely AP theoretical) analogue of Witten's* metric independent *Lagrangian for the Casson theory* [AJ] *p.121? What does González-Acuña's formula for the Rohlin invariant suggest? Is the mysterious question about the relationship between the Donaldson invariants of oppositely oriented X^4 related to the purely group-theoretic one of finding the inverse in $\mathcal{R}_n$ of an Artin Presentation?*

2. *In the absence of moduli spaces, etc., is Witten's "mass-gap" discussion regarding Donaldson theory,* [Wi3] *pp.289-291, still relevant in AP theory?*

3. *Is the Denjoy-like inequivalence between Seiberg-Witten theory and Donaldson theory detectable in AP theory? Recall that Seiberg-Witten theory requires*

spinors and the Dirac operator, i.e. an underlying C^1 structure, whereas Donaldson's theory is valid on the wider class of Lipschitz manifolds [D] *p.69,* [S], [DS].

4. *In general, the word problem obstructs the study of arbitrary smooth 4-manifolds. Although 4-manifolds in AP Theory are simply connected, we can still ask whether the group-theoretical physical questions of Geroch-Hartle* [GH] *(see also* [F]*) are still relevant when transferred to the group theory of 3-manifolds. Theorem I of* [W1] *seems to illustrate a purely group-theoretic Bohm-Aharonov phenomenon.*

5. *AP Theory does not just dispense with the smooth continuum, but also dispenses with integer (co)homology/intersection theory since all of this information is already given simply by the symmetric integer matrix $A\,(r)$. Hence, should e.g. the Kronheimer-Mrowka canonical basic class of $W^4\,(r)$,* when $\partial W^4\,(r) = S^3$, [D] p.52, [K], [St], *be already determined with Number Theory, à la Elkies* [E], [D] *p.67 and Borcherds* [B]*, thus explaining the persistence of invariants constructed with the aid of a complex structure when this structure does not exist? For the same reason, difficult 'minimal genus' and 'simple type' problems,* [D] *p.68,* [St] *p.156, should be studied in this, their ultimate natural context, where artificial complications caused by the use of the smooth continuum are absent.*

It does not seem surprising, due to the basic nature of the $K3$ complex surface (e.g. it is the only 4D, closed, simply connected Calabi-Yau manifold and its quadratic form is the first even non-Donaldson form), that our Artin Presentations lead to several interesting and instructive examples (section 3 ahead) which complement and extend to the 'softer' non-Donaldson case those examples obtained from such matrices as E_8, ϕ_{4n}, and the Coxeter-Todd extremal duodenary matrix $2D_{12}^2$ [W1].

2. The Artin Presentations

The purpose of this section is to construct Artin Presentations for all elliptic surfaces $E\,(n)$. This is carried out completely for $E\,(2)$, which is diffeomorphic to the Kummer surface $K3$ [GS], p.74, and follows mutatis mutandis for the others. The organization runs as follows: 2.1 is a brief discussion of Artin Presentations and framed pure braids, in 2.2 we obtain a surgery diagram for $E\,(n)$ that is a framed pure braid, 2.3 provides an explicit algorithm (fixing all conventions) for obtaining an Artin Presentation from a framed pure braid, and 2.4 combines everything obtaining the desired Artin Presentation for $K3$.

(2.1) Artin Presentations and Pure Braids. We begin by reviewing some of the fundamentals of AP theory. For a rigorous introduction to AP theory, proofs of the statements made below and a thorough bibliography we refer the reader to [W1].

Let $F_n = \langle x_1, \ldots, x_n \rangle$ be the free group on n-generators. An Artin Presentation r is a balanced presentation $r = \langle x_1, \ldots, x_n \mid r_1, \ldots, r_n \rangle$ satisfying the equation:

$$\text{(AC)} \qquad x_1 x_2 \cdots x_n = \left(r_1^{-1} x_1 r_1\right)\left(r_2^{-1} x_2 r_2\right) \cdots \left(r_n^{-1} x_n r_n\right),$$

in F_n, which we will refer to as the Artin Condition. The set of all Artin Presentations on $n-$generators is denoted $\mathcal{R}_n$ and forms a group. By Ω_n we mean the compact 2−disk with n−holes and boundary $\partial\Omega_n$ equal to the disjoint union of $\partial_0, \partial_1, \ldots, \partial_n$ (see [W1] p.225). An Artin Presentation $r \in \mathcal{R}_n$ determines, among other things, the following:

$$
\begin{array}{rcl}
\pi(r) & : & \text{the group presented by } r, \\
M^3(r) & : & \text{a closed orientable 3-manifold,} \\
W^4(r) & : & \text{a smooth compact connected} \\
 & & \text{simply-connected 4-manifold,} \\
A(r) & : & \text{an } n \times n \text{ symmetric integer matrix,} \\
h(r) & : & \text{a self diffeomorphism of } \Omega_n \text{ unique} \\
 & & \text{up to isotopy fixing } \partial\Omega_n \text{ with} \\
 & & h|_{\partial\Omega_n} \text{ equal to the identity.}
\end{array}
$$

The relationships between these objects are canonical. The manifold $M^3(r)$ bounds $W^4(r)$, has fundamental group isomorphic to $\pi(r)$, and is the open book defined by $h(r)$. The symmetric matrix $A(r)$ is the exponent sum matrix of r and also represents the intersection form of $W^4(r)$. The manifold $M^3(r)$ is a $\mathbb{Z}$−homology 3−sphere if and only if $\det A(r) = \pm 1$, and in this case we write $\Sigma^3(r)$ instead of $M^3(r)$.

An Artin Presentation $r \in \mathcal{R}_n$ also determines an automorphism of F_n by the mapping $x_i \mapsto r_i^{-1} x_i r_i$. Namely, this is the automorphism $h_\# : \pi_1(\Omega_n, p_0) \to \pi_1(\Omega_n, p_0)$ where p_0 is a distinguished point in $\partial_0 \subset \partial\Omega_n$ and $x_1, \ldots, x_n$ represent the canonical generators (see Figure 9 ahead and [W1] p.225 and p.244). This view will prove useful when composing Artin Presentations.

As pointed out in [W1], $\mathcal{R}_n$ is canonically isomorphic to $P_n \times \mathbb{Z}^n$, the framed pure braid group, where P_n is the pure braid group on n−strands. To see this, notice that $r \in \mathcal{R}_n$ determines $h = h(r)$ and h can be realized concretely in $\mathbb{R}^3$ by taking $\Omega_n \times I$ (I denotes the closed unit interval), suitably braiding the inner boundary tubes with one another, and twisting the inner boundary tubes by some integer numbers of complete revolutions (see [W1] p.245). Twisting the inner tubes can be accomplished by elementary Dehn twists about the ∂_i and these Dehn twists commute with all others. This braiding/twisting of the inner boundary tubes is easily seen to be equivalent to specifying both a pure braid (pure as $h|_{\partial\Omega_n} = id$) and an integer (the 'framing coefficient') for each strand.

Let $r \in \mathcal{R}_n$. The manifold $W^4(r)$ is defined in [W1] p.250 as follows. Embed Ω_n in S^2 and extend h to all of S^2 by the identity. Then, extend this map to a self diffeomorphism of all of D^3, calling the result $H = H(r)$ (which is unique up to isotopy). Letting $W(H)$ be the mapping torus of H, $W^4(r)$ is defined to be $W(H)$ union $(n+1)$ 2−handles attached canonically. Notice that $W(H)$ is diffeomorphic to $D^3 \times S^1$ (= 0−handle ∪ 1−handle) as all orientation preserving self diffeomorphisms of D^3 are smoothly isotopic to the identity. We wish to examine this construction more closely. The self diffeomorphism h of Ω_n can be realized, as described in the previous paragraph, in $\mathbb{R}^3$ as $\Omega_n \times I$ with the inner boundary tubes braided and twisted; the map h of Ω_n is then obtained by bending the twisted $\Omega_n \times I$ around and sticking the ends $\Omega_n \times 0$ and $\Omega_n \times 1$ together in the canonical way, exactly as one does to close a braid. To construct H, one can first extend h to D^2 by taking the twisted $\Omega_n \times I$ and filling in the

n inner boundary tubes with n copies of $D^2 \times I$. One must take some care here. For each boundary tube $\partial_i \times I$, $i = 1, \ldots, n$, let p_i be a distinguished point (see Figure 9 ahead and [W1] p.225). Let $*$ be a distinguished point in ∂D^2. Then, when filling the i^{th} boundary tube $\partial_i \times I$ with $D^2 \times I$ one must attach $* \times I$ to $p_i \times I$ and fill with the identity at the ends $\partial_i \times 0$ and $\partial_i \times 1$. Now, h has been extended to D^2 and is concretely realized as $D^2 \times I$ by sticking the ends together as when closing a braid; call this intermittent mapping torus $M(h)$ which is diffeomorphic to $D^2 \times S^1$. Now, extending the map to D^3 is trivial (again, $h|_{\partial \Omega_n} = id$) and one immediately sees that the $2-$handle attached corresponding to ∂_0 cancels the $1-$handle from the open book construction. Moreover, this cancellation occurs without disturbing the rest of the boundary of $W(H)$. Thus, we are left with a $0-$handle (i.e. D^4) with boundary S^3 containing a very nice copy of $M(h)$. To obtain $W^4(r)$ we now attach the remaining n $2-$handles to D^4 along the copies of $D^2 \times S^1$ in $M(h)$ in the canonical way.

Summarizing the previous two paragraphs, an Artin Presentation r determines a framed pure braid β in $\mathbb{R}^3$ (which is the same as in S^3) and $W^4(r)$ is obtained from D^4 by attaching $2-$handles according to β. In the language of the Kirby calculus, all $W^4(r)\,s$ are '$2-$handlebodies' ([GS], p.124). For more on the manifolds $W^4(r)$ see section 4.

Remark (2.1.1). One subtle but important distinction that must be made here between an $r \in \mathcal{R}_n$ and a framed pure braid in $S^3 = \partial D^4$ is that in an Artin Presentation the framings are canonically included (they are not 'put in by hand' as in the Kirby calculus) thus, e.g. avoiding serious self-linking problems [Wi1], p.363. In fact, a moment of reflection by the reader should reveal that without this 'canonicity' one would not obtain the purely group theoretic analogue of Donaldson's theorem [W1], p.240 Theorem 1, and its important consequences. See also [W1], p.241 and [W3].

Hence, one tack to obtain an Artin Presentation for a specific $4-$manifold is to obtain a surgery diagram for the manifold that is a framed pure braid in S^3 and then determine the corresponding Artin Presentation from this framed pure braid. Of course, saying an Artin Presentation r gives a *closed* $4-$manifold X^4 means that $M^3(r) = S^3$ and $W^4(r) \cup D^4 = X^4$ (i.e. close up with a $4-$handle). We pursue this tack in sections 2.2-2.4 below. We abuse notation and say an Artin Presentation or a surgery diagram gives a closed $4-$manifold when it actually presents the closed manifold minus the interior of a $4-$handle (which can only be attached in one way, so there is no ambiguity).

We close this section by recalling useful knot theoretic structures in AP Theory. The simplicity of these structures allows us to avoid doing surgery 'by hand', avoids self-linking problems, etc. by use of a computer algebra system such as MAGMA and significantly adds to the power of AP Theory. We point out that, as usual, everything is group theoretic.

Fix $r \in \mathcal{R}_n$, $r = \langle x_1, \ldots, x_n \mid r_1, \ldots, r_n \rangle$, with $\det A(r) = \pm 1$, in particular $\Sigma^3(r)$ is a $\mathbb{Z}-$homology $3-$sphere. There are $n+1$ distinguished knots in $\Sigma^3(r)$ that are defined by the boundary circles $\partial_0, \ldots, \partial_n$ of Ω_n and we denote these knots by $k_0, \ldots, k_n$. Let c_i denote the complement of k_i in $\Sigma^3(r)$ and let G_i denote the fundamental group of c_i. Since $A(r)$ is unimodular, $A(r)^{-1}$ is also a symmetric integer matrix and, in fact, is the linking matrix of the knots k_i,

$i = 1, \ldots, n$. We let b_{ij} denote the ij^{th} entry of $A(r)^{-1}$ (abbreviating b_{ii} to just b_i) and let $s = \sum_{ij} b_{ij}$. In $\Sigma^3(r)$, the self linking number of k_0 is s and of k_i, $i \neq 0$, is b_i. We let m_i, l_i denote the peripheral structure of the knot k_i, which consists of two special commuting elements in G_i, where m_i is a meridian of k_i and l_i is homologically trivial in the complement of k_i. Then, we have:

$$\begin{aligned} G_0 &= \langle x_1, \ldots, x_n \mid r_1 = r_2 = \cdots = r_n \rangle, \\ m_0 &= \text{any } r_i, \\ l_0 &= x_1 x_2 \cdots x_n m_0^{-s}, \end{aligned}$$

and for $i = 1, \ldots, n$ we have:

$$\begin{aligned} G_i &= \langle x_1, \ldots, x_n \mid r_1, r_2, \ldots, r_{i-1}, r_{i+1}, \ldots, r_n \rangle, \\ m_i &= r_i, \\ l_i &= x_i m_i^{-b_i}. \end{aligned}$$

Two remarks are in order. First of all, we get *all* knots and links in any arbitrary closed, orientable 3−manifold this way (González-Acuña unpublished). Second, the definition given here of G_i for $i \neq 0$ appears to be slightly different from that given in [W1], p.227, but in fact the two are equivalent (this was pointed out to the second author by González-Acuña). This follows since the Artin Condition (AC) implies that in G_i (definition given here) we have:

$$x_1 x_2 \cdots x_n = x_1 x_2 \cdots x_{i-1} \left(r_i^{-1} x_i r_i\right) x_{i+1} \cdots x_n,$$

which immediately implies that $x_i = r_i^{-1} x_i r_i$ in G_i. That is, $(x_i, r_i) = 1$ in G_i (where (a, b) is MAGMA notation for the commutator $a^{-1}b^{-1}ab$), showing the two definitions are equivalent. In fact, for $i \neq 0$, m_i and l_i commuting in G_i is equivalent to x_i and r_i commuting in G_i.

(2.2) Pure Braid for $E(n)$. Our starting point is the framed link diagram in [HKK], p.66 (see also [GS], p.305) that presents a 2−handlebody with boundary S^3 and gives $E(n)$ upon closing up with a 4−handle. (As mentioned earlier, we abuse notation and say this diagram presents $E(n)$ where no confusion should arise.) By straightforward isotopy of the outer strand (the trefoil) we obtain Figure 1. The two large bands both represent $6n - 2$ strands, each strand with framing -2. A box containing '-1' represents a twist of all strands (as when twisting ribbon) in the direction corresponding to a negative crossing in our orientation convention in Figure 8. We refer to the trefoil in Figure 1 as T and to the small circle linking it as S, which have framings 0 and $-n$ respectively.

All circles formed by closing a pure braid are individually not knotted, so the first step is to unknot the trefoil T. To accomplish this, one performs a 2−handle slide on T; in practice this corresponds to performing a band sum of T with a parallel curve to another knot K representing the framing on K (see [GS], pp.141-143). Here we slide T over the innermost circle in the left large band using the trivial band as in Figure 2. One checks that the curve in Figure 2 that T is being band summed with is a parallel curve to the innermost strand and has linking number -2 with it (don't forget the '-1' box!). Let T' denote the result of 2-handle sliding T. Figure 3 is obtained from Figure 2 by isotopy, in particular grab the part of T' in Figure 2 that hangs below the two large bands and swing it back and then up (other minor changes by isotopy here should

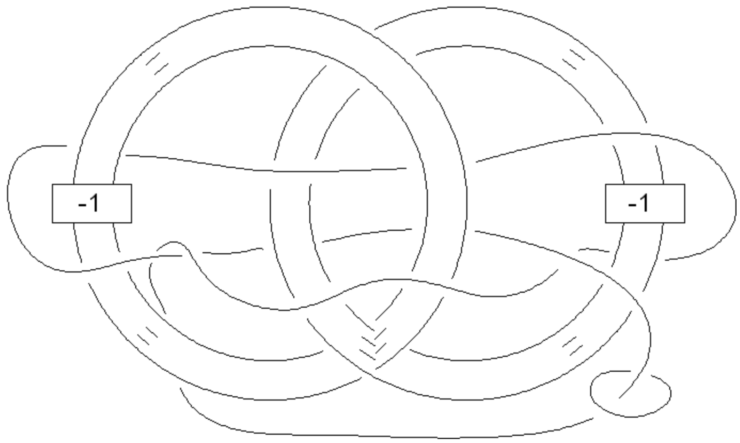


Figure 1. Surgery diagram for $E(n)$. The large bands represent $6n-2$ strands and all framings equal -2, except the trefoil T with framing 0 and the small circle S linking it with framing $-n$.

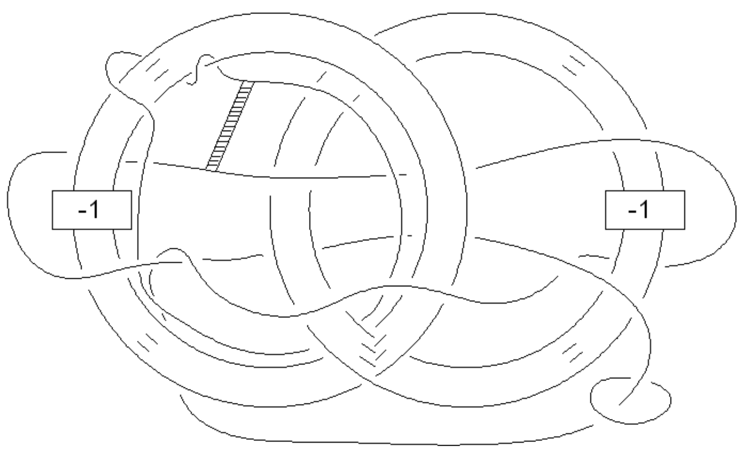


Figure 2. A $2-$handle slide of T over the innermost curve in the left large band using the indicated parallel curve and dashed band.

be obvious). Straightforward isotopy of Figure 3 produces Figure 4 where it is apparent that T' is not knotted.

It does not seem possible to isotop Figure 4 to a pure braid, so we perform another $2-$handle slide. This time, slide T' over the outermost strand in the right large band (again using a trivial band to band sum with) as shown in Figure 5. After a little isotopy one obtains Figure 6 (ignoring the hatched rectangle for the moment). Let T'' denote the result in Figure 6 of sliding T' (S is unchanged).

Now, Figure 6 isotops nicely to a pure braid. To see this, take the hatched rectangle in Figure 6, grab its upper left long boundary edge and pull it around, making a rather large (ambient) expansion of the hatched rectangle into a large backwards 'C' shape (the short dimension of the hatched rectangle extends and

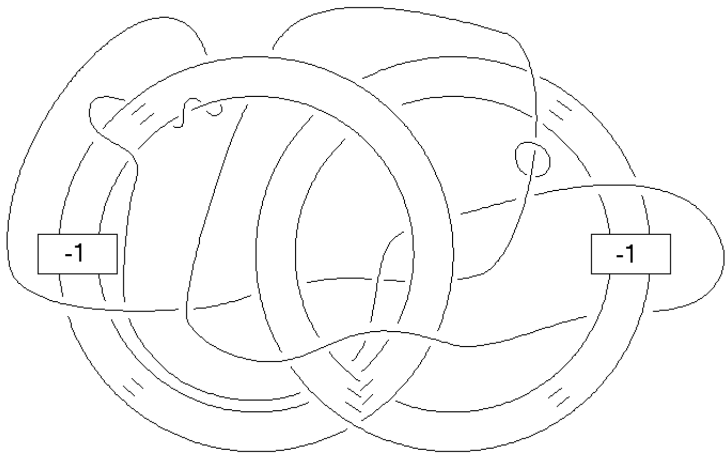


Figure 3. The result T' of 2-handle sliding T.

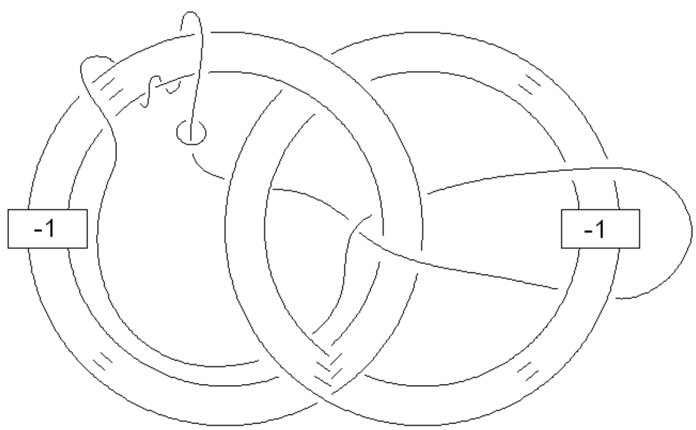


Figure 4. The result of isotoping T' (and S), which is not knotted.

bends around). Except for S, one now has a pure braid. A little more straightforward isotopy produces Figure 7, which is a pure braid for $E(n)$. The hatched rectangle does not appear in Figure 7, but one imagines it bending around on the right-hand side to close the braid. Figure 7 contains a total of $12n-2$ strands: the two large bands each represent $6n-2$ strands (each strand therein has framing -2), the $(12n-3)^{\text{rd}}$ strand (second from the right) is T'', and the $(12n-2)^{\text{nd}}$ strand (right-most) is S with framing $-n$.

It remains to determine the framing on T'' (this is the only one that changed), which is calculated using the formula in [GS] p.142. The first $2-$handle slide results in T' with framing -2 since the relevant (signed, according to handle addition or subtraction) linking number is 0. The second $2-$handle slide results in T'' with framing still -2 since in this case the relevant signed linking number (whose overall sign is independent of orientation choices) is equal to $+1$ implying

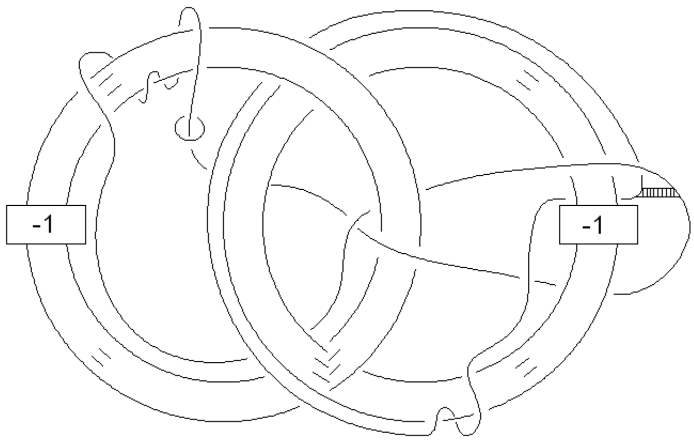


Figure 5. A $2-$handle slide of T' over the outermost curve in the right large band using the indicated parallel curve and dashed band.

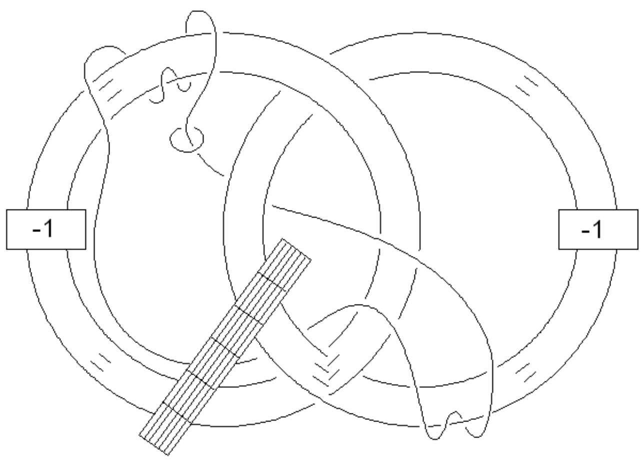


Figure 6. The result T'' of $2-$handle sliding T. The hatched rectangle will be used to isotop to a pure braid.

$\pm 2lk\left(\cdot,\cdot\right) = +2$. Thus, in the pure braid diagram for $E\left(n\right)$ in Figure 7 all framings equal -2 except for the right-most strand which has framing $-n$. In particular, for the Kummer surface $E\left(2\right)$ all framings equal -2.

Remark (2.2.1). In Figure 1, the two large bands together form the compactified Milnor fiber $M_c\left(2,3,6n-1\right)$ with boundary the Seifert fibered $\mathbb{Z}-$homology $3-$sphere $\Sigma\left(2,3,6n-1\right)$ and the trefoil union the small circle linking it form the Gompf nucleus $N\left(n\right)$ (see [GS], sec. 3.1, 6.3, 7.3 and 8.3). It is clear from the above that all Milnor fibers $M_c\left(2,3,6n-1\right)$ admit Artin Presentations.

(2.3) An Algorithm. Given a framed pure braid in $\mathbb{R}^3$, we wish to construct the corresponding Artin presentation. To make this explicit, we must fix some

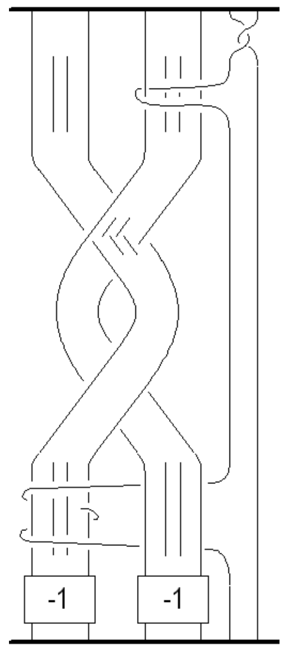


Figure 7. Pure braid for $E(n)$. The large bands represent $6n-2$ strands and all framings equal -2, except for the rightmost strand with framing $-n$.

conventions. We will use β to denote both a braid and a framed braid, where no confusion should arise. As usual, braids will be drawn as generic diagrams in the plane with the strands ordered $1, 2, \ldots, n$ from left to right. We read our braids upwards, especially when composing them. In particular, each strand is oriented up. For a pure braid β, C_i will denote the oriented circle consisting of the i^{th} strand and the trivial segment that would close that strand upon closing the braid (the orientation is inherited from that of the corresponding braid strand). Crossings in any oriented generic link diagram in the plane are assigned a sign as in Figure 8. If C_1 and C_2 are two oriented circles in a generic link diagram in the plane, then their linking number $lk(C_1, C_2)$ is defined to be the number of positive undercrossings of C_2 under C_1 minus the number of negative undercrossings of C_2 under C_1. The linking number is well defined and symmetric (see [GS] sec. 4.5). For an n-strand framed pure braid β the linking matrix $L(\beta)$ of β is the $n \times n$ symmetric integer matrix L where $L_{ij} = lk(C_i, C_j)$ for $i \neq j$ and equals the framing coefficient of C_i for $i = j$. Similarly, one can define the linking matrix of any ordered oriented framed generic link diagram in the plane.

Remark (2.3.1). If $r \in \mathcal{R}_n$ corresponds to β a framed pure braid then $A(r) = L(\beta)$. This follows from [W1], section 1 and [GS], p.125. We note that orientations/conventions fixed agree with both [W1] and [GS].

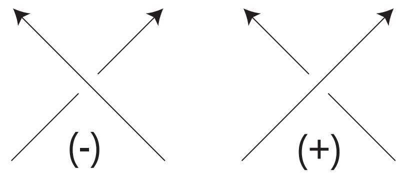


Figure 8. Crossing signs in an oriented link diagram.

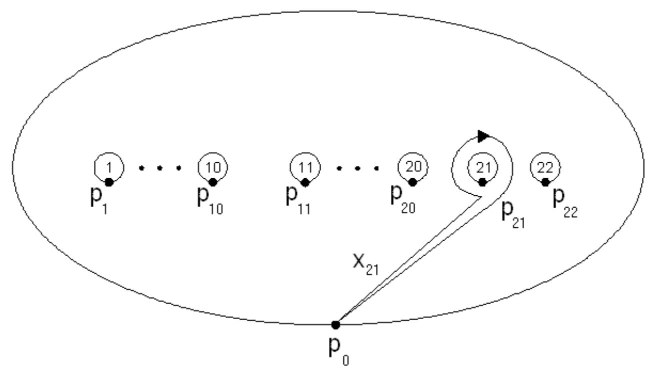


Figure 9. Ω_{22} with basepoints $p_0, \ldots, p_{22}$ on boundary components $\partial_0, \ldots, \partial_{22}$. Also indicated is a generator x_{21} of $\pi_1(\Omega_{22}, p_0)$.

Any pure braid $\beta \in P_n$ can be written as a product of Dehn twists about simple closed curves in Ω_n. Thus, we will need these three steps:

Step I. Given a pure braid β resulting from a single Dehn twist, determine the corresponding Artin Presentation.

Step II. Compose two Artin Presentations.

Step III. Correct Framings.

Remark (2.3.2). Again, Step III is necessary since when going from a framed pure braid (where framings are not canonically included) to an Artin Presentation (where framings are canonically included) an ad hoc framing correction must be made at some point.

We describe these in detail.

Step I. First, $\pi_1(\Omega_n, p_0)$ has canonical generators. Figure 9 shows Ω_{22} with basepoint p_0 and the generator x_{21} (the other generators are similar; see also [W1] p.225 and p.244). Also depicted in Figure 9 are basepoints on the boundary components $\partial_1, \ldots, \partial_{22}$ (as referred to in section 2.1).

We use two examples to illustrate this step. For the first example, take the Dehn twist depicted in Figure 10 about the oriented simple closed curve D_1 (for the moment ignore the small segment laid across D_1). Usually one would take a cylinder neighborhood $S^1 \times [-1, 1]$ of D_1 in Ω_{22} and replace it with a twisted version (often a cut along D_1 takes place) according to some fixed orientation convention (see, for example, [GS] p.295). Following the motivation setforth in section 2.1, we prefer to realize the Dehn twist canonically as an

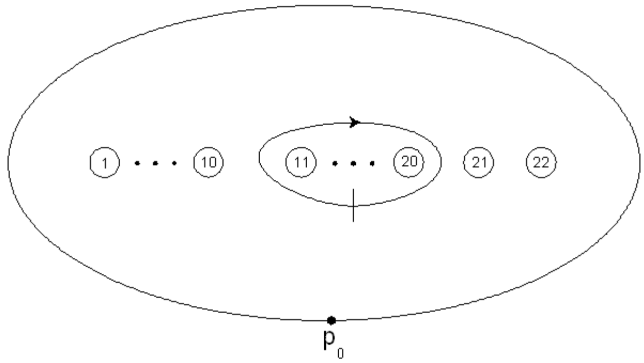


Figure 10. Ω_{22} with an oriented simple closed curve D_1 and a small segment laid across it.

isotopy of Ω_{22} in $\mathbb{R}^3$ as follows. Start with a copy of Ω_{22} (as in Figure 10) laying flat on the (possibly imaginary) table in front of you and a small cylinder neighborhood $N = S^1 \times [-1,1]$ of D_1 in Ω_{22}. The inner boundary curve of N bounds a compact disk with 10 holes denoted Ω'_{10}. Slowly raise Ω_{22} up off the table and while doing so grab Ω'_{10} and slowly rotate it clockwise about a central point (with the cylinder neighborhood N stretching like rubber) one complete revolution. If one pictures the paths traced out by the center points of the 22 punctures in Ω_{22} during this Dehn twist, one immediately sees the pure braid obtained from Figure 7 with $n = 2$ by just taking the '-1' box on strands $11-20$ and taking the remaining strands to be trivial. This Dehn twist, realized as an isotopy, gives a self diffeomorphism h of Ω_{22} that is fixed on $\partial\Omega_{22}$, namely the time 1 map of the isotopy. As discussed above in section 2.1 and [W1] pp.243-244, the automorphism $h_\#$ of $\pi_1(\Omega_{22}, p_0) \cong F_{22}$ induced by h is of the form $x_i \mapsto r_i^{-1} x_i r_i$ for some words r_i and $r = \langle x_1, \ldots, x_{22} \mid r_1, \ldots, r_{22}\rangle$ is our desired Artin presentation. The word r_i is nontrivial ($\neq 1$) only for $i = 11, \ldots, 20$ and these are all equal to one another. To compute r_{11}, say, lay a straight segment across D_1 in front of ∂_{11} as in Figure 10 and follow the segment through the isotopy above. After the isotopy, add two oriented edges to the isotoped segment: one from p_0 to the upper endpoint and one from the lower endpoint to p_0 as in Figure 11; the word in $\pi_1(\Omega_{22}, p_0)$ represented by this oriented loop is $r_{11} = x_{20}^{-1} x_{19}^{-1} \cdots x_{11}^{-1}$.

We note two important points concerning the above example. First, it conveyed the orientation convention of Dehn twists used here, namely grab the inner compact disk with holes and twist it in the direction of the arrow on the curve one is twisting about. Second, the small segment laid across D_1 formed the 'meat' of the relations and only crossed D_1 once. When computing r_i in general, one must choose this segment to traverse all occurrences of the curve one is twisting about between a nice path (usually a straight line segment or a small isotopy of one) from p_0 to p_i. This is shown in the following example.

For this example, take the Dehn twist depicted in Figure 12. The automorphism of F_{22} is clearly the identity on $x_1, \ldots, x_{10}, x_{22}$. Figure 13 shows the loop representing both words $r_{11} = r_{21} = x_{21}^{-1} x_{11}^{-1}$ (as the reader can verify using the

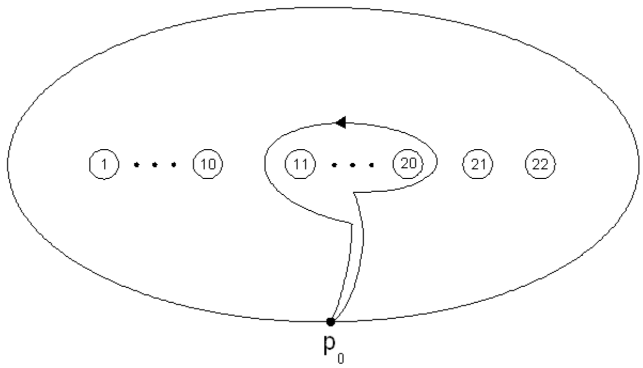


Figure 11. Ω_{22} with a loop representing $r_{11}, \ldots, r_{20}$.

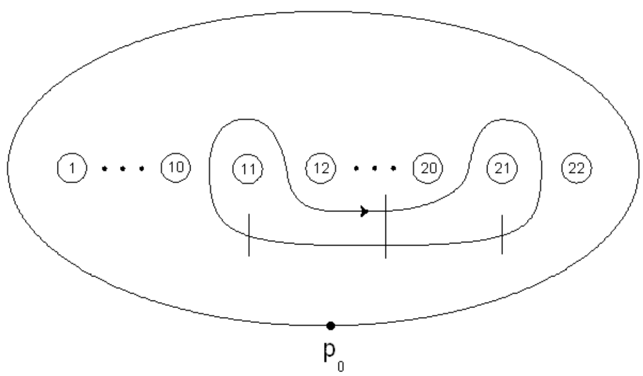


Figure 12. Ω_{22} with an oriented simple closed curve D_{24} and three small segments laid across it.

two small segments in Figure 12 that cross D_{24} once). The more interesting relations are $r_{12}, \ldots, r_{20}$, which are all equal to one another. To compute these one must use a segment that crosses D_{24} twice, such as the middle segment in Figure 12. The resulting loop is shown in Figure 14 and represents the word $x_{21}x_{11}x_{21}^{-1}x_{11}^{-1}$. This completes Step I.

Step II. Our data is two Artin presentations r, r' arising from Dehn twists about D, D' with corresponding h, h' and $h_{\#}, h'_{\#}$. Then, the composite Artin presentation $r'' = r' \circ r$ is obtained using the formula (see [W1], p.245):

$$r''_i = r'_i \cdot h'_{\#}(r_i) .$$

Step II is impractical by hand when the presentations are not small and use of a computer algebra system, such as MAGMA, is invaluable.

Step III. Our data now is a framed pure braid β and an Artin presentation r' resulting from repeated applications of Steps I and II. One also has the matrices $L(\beta)$ and $A(r')$ which differ only possibly on their diagonals. One corrects (see

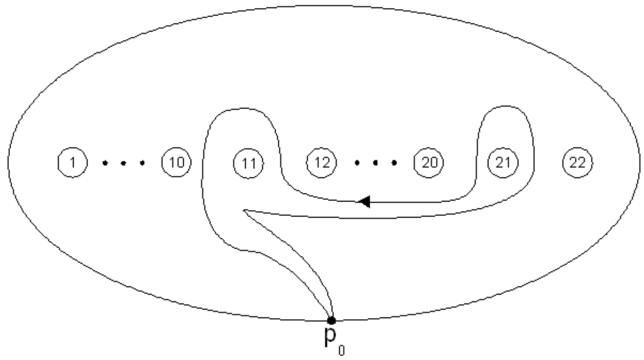


Figure 13. Ω_{22} with a loop representing r_{11} and r_{21}.

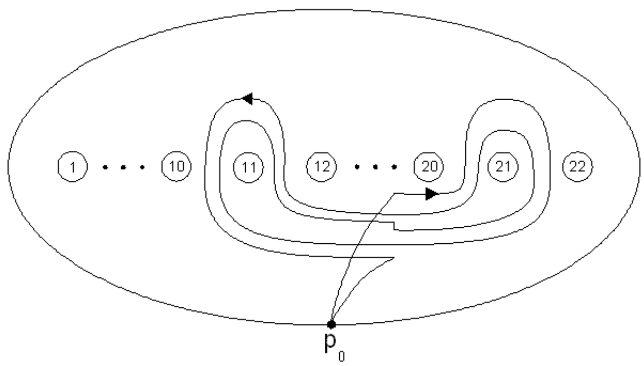


Figure 14. Ω_{22} with a loop representing $r_{12}, \ldots, r_{20}$.

Remark (2.3.2) and Section 2.1) using the simple rule:

$$\begin{aligned} \text{let } \delta_i &= L(\beta)_{ii} - A(r')_{ii}, \text{ and} \\ \text{let } r_i &= x_i^{\delta_i} \cdot r_i'. \end{aligned}$$

The result is the Artin presentation $r = \langle x_1, \ldots, x_n \mid r_1, \ldots, r_n \rangle$ and $A(r) = L(\beta)$. We point out that when correcting framings one must multiply on the left by the corresponding $x_i^{\delta_i}$, otherwise the resulting presentation is usually not Artin. This completes Step III.

(2.4) Artin Presentation of $K3$. Begin with the framed pure braid in Figure 7 with $n = 2$. Call this braid β and recall that all framings equal -2. We need a series of Dehn twists producing β (ignoring framings for the moment). To take care of β (reading up from the bottom) up until the point where the two large bands first cross each other, perform Dehn twists about D_1, D_2, D_3, and D_4 (in that order!) as in Figure 10 and Figures 15-17. (It may seem that the '-1' on the left band has been left off, but the reader should check that this is not the case.) Now we attack the brunt of β consisting of the 'Milnor fiber' where the two large bands cross each other and then intertwine. For

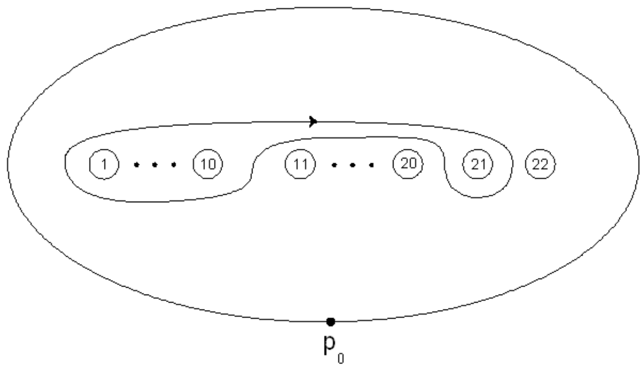


Figure 15. Ω_{22} with an oriented simple closed curve D_2.

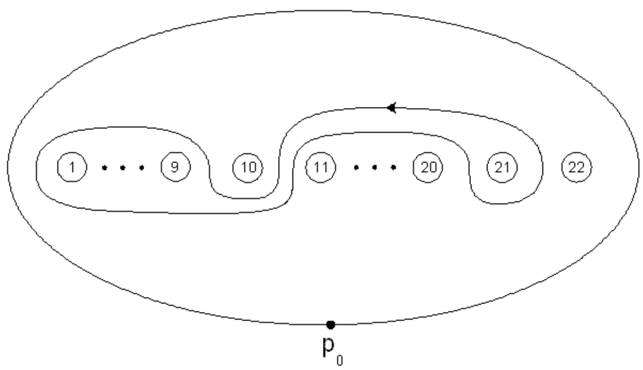


Figure 16. Ω_{22} with an oriented simple closed curve D_3.

this part we will need Figures 18 and 19 repeated in an alternating fashion. Figure 18 represents $D_5, D_7, D_9, \ldots, D_{23}$ where D_{5+2j}, $j = 0, 1, 2, \ldots, 9$, corresponds to Figure 18 with $k = j + 1$ and $k' = j + 11$. Figure 19 represents $D_6, D_8, D_{10}, \ldots, D_{22}$ where D_{6+2j}, $j = 0, 1, 2, \ldots, 8$, corresponds to Figure 19 with $k = j+1$. Then, one performs Dehn twists about the following ordered and oriented curves: $D_5, D_6, \ldots, D_{22}, D_{23}$. The reader should check that this series of Dehn twists performs as claimed. To finish up, one twists about D_{24} as in Figure 12 and then about D_{25} as in Figure 20. This series of Dehn twists gives β up to framings.

Now, using Step I from section 2.3, one writes down the Artin presentation corresponding to each of the Dehn twists in this series. We organize this data into a 25×22 array R of words in F_{22} where $R[i, \cdot]$ corresponds to D_i (i.e. $R[i, j]$ is the j^{th} relation of the i^{th} Artin presentation). Assume that R is initialized as the 25×22 array of identity elements in F_{22}. The nontrivial elements in R are as follows.

$R[1, i]$	
$i = 11, \ldots, 20$	$x_{20}^{-1} x_{19}^{-1} x_{18}^{-1} x_{17}^{-1} x_{16}^{-1} x_{15}^{-1} x_{14}^{-1} x_{13}^{-1} x_{12}^{-1} x_{11}^{-1}$

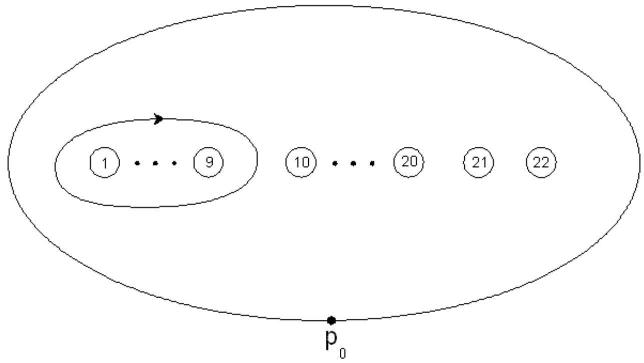


Figure 17. Ω_{22} with an oriented simple closed curve D_4.

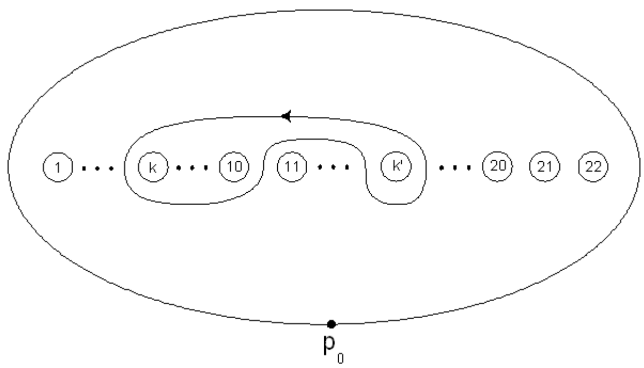


Figure 18. Ω_{22} with an oriented simple closed curve D_*.

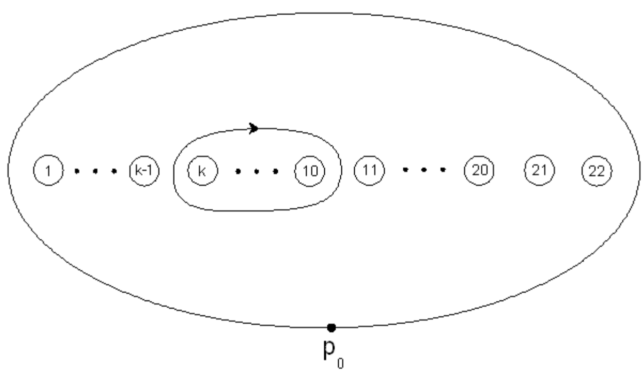


Figure 19. Ω_{22} with an oriented simple closed curve D_*.

$R[2,i]$	
$i = 1, \ldots, 10$	$x_{11}x_{12}x_{13}x_{14}x_{15}x_{16}x_{17}x_{18}x_{19}x_{20}$ $x_{21}^{-1}x_{20}^{-1}x_{19}^{-1}x_{18}^{-1}x_{17}^{-1}x_{16}^{-1}x_{15}^{-1}x_{14}^{-1}x_{13}^{-1}x_{12}^{-1}x_{11}^{-1}$ $x_{10}^{-1}x_9^{-1}x_8^{-1}x_7^{-1}x_6^{-1}x_5^{-1}x_4^{-1}x_3^{-1}x_2^{-1}x_1^{-1}$
$i = 21$	$x_{21}^{-1}x_{20}^{-1}x_{19}^{-1}x_{18}^{-1}x_{17}^{-1}x_{16}^{-1}x_{15}^{-1}x_{14}^{-1}x_{13}^{-1}x_{12}^{-1}x_{11}^{-1}$ $x_{10}^{-1}x_9^{-1}x_8^{-1}x_7^{-1}x_6^{-1}x_5^{-1}x_4^{-1}x_3^{-1}x_2^{-1}x_1^{-1}$ $x_{11}x_{12}x_{13}x_{14}x_{15}x_{16}x_{17}x_{18}x_{19}x_{20}$

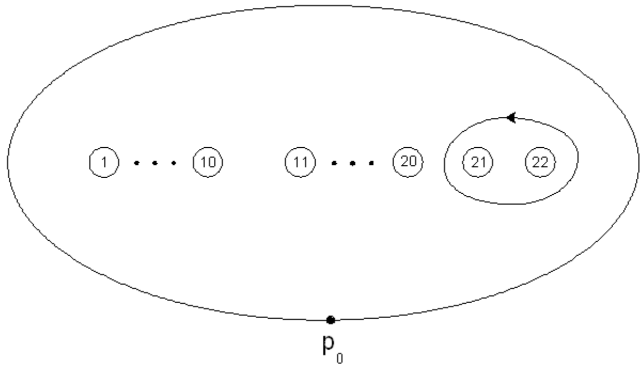


Figure 20. Ω_{22} with an oriented simple closed curve D_{25}.

$R[3,i]$	
$i=1,\ldots,9$	$x_1x_2x_3x_4x_5x_6x_7x_8x_9$ $x_{11}x_{12}x_{13}x_{14}x_{15}x_{16}x_{17}x_{18}x_{19}x_{20}x_{21}$ $x_{20}^{-1}x_{19}^{-1}x_{18}^{-1}x_{17}^{-1}x_{16}^{-1}x_{15}^{-1}x_{14}^{-1}x_{13}^{-1}x_{12}^{-1}x_{11}^{-1}$
$i=10$	$x_9^{-1}x_8^{-1}x_7^{-1}x_6^{-1}x_5^{-1}x_4^{-1}x_3^{-1}x_2^{-1}x_1^{-1}$ $x_{11}x_{12}x_{13}x_{14}x_{15}x_{16}x_{17}x_{18}x_{19}x_{20}$ $x_{21}^{-1}x_{20}^{-1}x_{19}^{-1}x_{18}^{-1}x_{17}^{-1}x_{16}^{-1}x_{15}^{-1}x_{14}^{-1}x_{13}^{-1}x_{12}^{-1}x_{11}^{-1}$ $x_1x_2x_3x_4x_5x_6x_7x_8x_9$ $x_{11}x_{12}x_{13}x_{14}x_{15}x_{16}x_{17}x_{18}x_{19}x_{20}x_{21}$ $x_{20}^{-1}x_{19}^{-1}x_{18}^{-1}x_{17}^{-1}x_{16}^{-1}x_{15}^{-1}x_{14}^{-1}x_{13}^{-1}x_{12}^{-1}x_{11}^{-1}$
$i=21$	$x_{20}^{-1}x_{19}^{-1}x_{18}^{-1}x_{17}^{-1}x_{16}^{-1}x_{15}^{-1}x_{14}^{-1}x_{13}^{-1}x_{12}^{-1}x_{11}^{-1}$ $x_1x_2x_3x_4x_5x_6x_7x_8x_9$ $x_{11}x_{12}x_{13}x_{14}x_{15}x_{16}x_{17}x_{18}x_{19}x_{20}x_{21}$

$R[4,i]$	
$i=1,\ldots,9$	$x_9^{-1}x_8^{-1}x_7^{-1}x_6^{-1}x_5^{-1}x_4^{-1}x_3^{-1}x_2^{-1}x_1^{-1}$

Now, the relations $R[5-23,i]$ lend themselves well to looping/shorthand (which we utilize especially when using MAGMA). Let $w=x_{19}^{-1}x_{18}^{-1}\cdots x_{11}^{-1}$ and let w_j denote the first j letters of w read from the right for $j=0,\ldots,9$. For example, $w_0=1$ (i.e. the identity in F_n) and $w_2=x_{12}^{-1}x_{11}^{-1}$. Then, $R[5,i]$, $R[7,i]$, , $R[23,i]$ are defined by the following where $j=0,1,\ldots,9$:

$R[5+2j,i]$	
$i=(j+1),\ldots,10$	$x_{j+1}x_{j+2}\cdots x_{11+j}w_j$
$i=11+j$	$w_jx_{j+1}x_{j+2}\cdots x_{11+j}$

Also, $R[6,i]$, $R[8,i]$, , $R[22,i]$ are defined by the following where $j=0,1,\ldots,8$:

$R[6+2j,i]$	
$i=(j+1),\ldots,10$	$x_{10}^{-1}x_9^{-1}\cdots x_{j+1}^{-1}$

And the last two Artin presentations:

$R[24,i]$	
$i = 11, 21$	$x_{21}^{-1}x_{11}^{-1}$
$i = 12, \ldots, 20$	$x_{21}x_{11}x_{21}^{-1}x_{11}^{-1}$

$R[25,i]$	
$i = 21, 22$	$x_{21}x_{22}$

The list of Artin Presentations corresponding to the series of Dehn twists given above is complete. Now, one simply composes these 25 presentations (with MAGMA!) using a loop statement and the formula from Step II in section 2.3. Call the result of this iterated composition r'. To correct the framings, one computes the exponent sum matrix of r' (again using MAGMA) and checks the diagonal of this matrix which is (starting from the upper left):

$$\left(\underbrace{-1,\ldots,-1}_{9\text{ times}},0,-1,\underbrace{0,\ldots,0}_{10\text{ times}},1\right)$$

To make these entries all equal -2, one corrects r' using Step III calling the result r. This is the desired Artin presentation for the Kummer surface $K3$.

After obtaining r with MAGMA, one immediately checks that the presentation is in fact Artin. To do so, simply prompt MAGMA to compute the right hand side of the Artin condition (AC). The result should be (and for our r is) the left hand side of (AC). This is an important test, but it is also a test that MAGMA can always carry out as the word problem in F_n is solved and MAGMA must only freely reduce.

By construction, $M^3(r)$ is S^3 and $W^4(r)$ is $K3$. Despite the length of the presentation r (which is given below) MAGMA readily verifies that $\pi(r) = 1$. To look at $W^4(r)$ one proceeds to $A(r)$ which appears in Figure 21. This matrix is even, unimodular, has 19 negative eigenvalues and 3 positive ones, hence is $\mathbb{Z}-$congruent to $2E_8 \oplus 3H$ as expected. One is now ready to reap the rewards of this work. The Artin presentation r can be easily and orderly investigated with MAGMA where nothing has to be done by hand and one doesn't need to worry about surgery diagrams, etc. Examples appear in the following section.

The inverse matrix of $A(r)$, which appears in Figure 22, provides the peripheral structure of the knots k_i, $i = 0, \ldots, 22$, described at the beginning of this section. Notice that the diagonal consists entirely of -2, 0, and 2, which as a consequence immediately again gives Artin presentations for the appropriate $(1, \pm 1)$ Dehn spheres. Further, notice that the total sum of $A(r)^{-1}$, denoted s, equals -6, another computational advantage.

The knots k_i are nontrivial only for $i = 0, 10, 11, 21, 22$; k_{10} and k_{11} are 5_2s, k_{22} is a trefoil, and k_{21}, with Alexander polynomial $\Delta = t^4 - t^2 + 1$, is a cable of the trefoil. However, k_0 has Alexander polynomial $\Delta = t^8 - 2t^7 - 5t^5 + 13t^4 - \ldots$ and is off the usual knot tables; its 2, 3, 4, 5 torsion is given by (29), $(13, 13)$, $(15, 435)$, $(251, 251)$.

It seems curious that here the only non-fibered knots are k_{10} and k_{11}, precisely where the pair of $3s$ appears off the diagonal in $A(r)^{-1}$ (Figure 22); see also the end of section 2.1.

As $\mathcal{R}_{22}$ is a group, one may wish to compute r^{-1}. To do so, one performs the same series of Dehn twists as for r but in the reverse order and with reverse

$$
\begin{array}{cccccccccccccccccccccc}
-2&-1&-1&-1&-1&-1&-1&-1&-1&-1&1&&&&&&&&&&&\\
-1&-2&-1&-1&-1&-1&-1&-1&-1&-1&1&1&&&&&&&&&&\\
-1&-1&-2&-1&-1&-1&-1&-1&-1&-1&1&1&1&&&&&&&&&\\
-1&-1&-1&-2&-1&-1&-1&-1&-1&-1&1&1&1&1&&&&&&&&\\
-1&-1&-1&-1&-2&-1&-1&-1&-1&-1&1&1&1&1&1&&&&&&&\\
-1&-1&-1&-1&-1&-2&-1&-1&-1&-1&1&1&1&1&1&1&&&&&&\\
-1&-1&-1&-1&-1&-1&-2&-1&-1&-1&1&1&1&1&1&1&1&&&&&\\
-1&-1&-1&-1&-1&-1&-1&-2&-1&-1&1&1&1&1&1&1&1&1&&&&\\
-1&-1&-1&-1&-1&-1&-1&-1&-2&-1&1&1&1&1&1&1&1&1&1&&&\\
-1&-1&-1&-1&-1&-1&-1&-1&-1&-2&1&1&1&1&1&1&1&1&1&1&-1&\\
1&1&1&1&1&1&1&1&1&1&-2&-1&-1&-1&-1&-1&-1&-1&-1&-1&-1&\\
&1&1&1&1&1&1&1&1&1&-1&-2&-1&-1&-1&-1&-1&-1&-1&-1&&\\
&&1&1&1&1&1&1&1&1&-1&-1&-2&-1&-1&-1&-1&-1&-1&-1&&\\
&&&1&1&1&1&1&1&1&-1&-1&-1&-2&-1&-1&-1&-1&-1&-1&&\\
&&&&1&1&1&1&1&1&-1&-1&-1&-1&-2&-1&-1&-1&-1&-1&&\\
&&&&&1&1&1&1&1&-1&-1&-1&-1&-1&-2&-1&-1&-1&-1&&\\
&&&&&&1&1&1&1&-1&-1&-1&-1&-1&-1&-2&-1&-1&-1&&\\
&&&&&&&1&1&1&-1&-1&-1&-1&-1&-1&-1&-2&-1&-1&&\\
&&&&&&&&1&1&-1&-1&-1&-1&-1&-1&-1&-1&-2&-1&&\\
&&&&&&&&&1&-1&-1&-1&-1&-1&-1&-1&-1&-1&-2&&\\
&&&&&&&&&-1&-1&&&&&&&&&&-2&1\\
&&&&&&&&&&&&&&&&&&&&1&-2
\end{array}
$$

Figure 21. $A\,(r)$ for r representing the Kummer surface.

orientation. One must repeat Step I for all of these Dehn twists and the work is equivalent to the work involved with getting r. After doing so, one compares the lengths of the relations in r and r^{-1} which appear below. (We use $\#r$ to denote the total length of all relations.) We note that shorter presentations are not necessarily more useful computationally, especially with MAGMA, as one quickly finds.

Relation	r	r^{-1}	Relation	r	r^{-1}
1	130	176	12	252	502
2	131	403	13	247	501
3	132	628	14	240	500
4	133	851	15	231	499
5	134	1072	16	220	498
6	135	1291	17	207	497
7	136	1508	18	192	496
8	137	1723	19	175	495
9	138	1936	20	156	494
10	644	2126	21	529	573
11	258	108	22	5	383

–2		1	1		–1	–1		1	1	–1	1	1		–1	–1		1	1			
	–2		1	1		–1	–1		1	–1	–1	1	1		–1	–1		1	1		
1		–2		1	1		–1	–1			–1	–1	1	1		–1	–1		1		
1	1		–2		1	1		–1	–1	1		–1	–1	1	1		–1	–1			
	1	1		–2		1	1		–1	1	1		–1	–1	1	1		–1	–1		
–1		1	1		–2		1	1			1	1		–1	–1	1	1		–1		
–1	–1		1	1		–2		1	1	–1		1	1		–1	–1	1	1			
	–1	–1		1	1		–2		1	–1	–1		1	1		–1	–1	1	1		
1		–1	–1		1	1		–2			–1	–1		1	1		–1	–1	1		
1	1		–1	–1		1	1			3		–1	–1		1	1		–1	–1	–2	–1
–1	–1		1	1		–1	–1		3			1	1		–1	–1		1	1	–2	–1
1	–1	–1		1	1		–1	–1			–2		1	1		–1	–1		1		
1	1	–1	–1		1	1		–1	–1	1		–2		1	1		–1	–1			
	1	1	–1	–1		1	1		–1	1	1		–2		1	1		–1	–1		
–1		1	1	–1	–1		1	1			1	1		–2		1	1		–1		
–1	–1		1	1	–1	–1		1	1	–1		1	1		–2		1	1			
	–1	–1		1	1	–1	–1		1	–1	–1		1	1		–2		1	1		
1		–1	–1		1	1	–1	–1			–1	–1		1	1		–2		1		
1	1		–1	–1		1	1	–1	–1	1		–1	–1		1	1		–2			
	1	1		–1	–1		1	1	–1	1	1		–1	–1		1	1		–2		
									–2	–2										2	1
									–1	–1										1	

Figure 22. The inverse matrix $A(r)^{-1}$ providing the peripheral structures of the knots $k_0, \ldots, k_{22}$.

Total Relator Length	
$\#r$	$\#r^{-1}$
4562	17260

In the following, we denote the just constructed r, r^{-1} by $k3, k3^{-1}$. Let t_1 be the Torelli of [W1] p.228. If we multiply $k3^{-1}$ "at 20 by t_1" [W1] p.227, i.e. if we take the Artin presentation $r \in \mathcal{R}_{22}$ where r_i equals 1 for $i < 20$ and equals t_1 written in the variables x_{20}, x_{21}, x_{22} for $i = 20, 21, 22$ and multiply it by $k3^{-1}$, we obtain an Artin presentation, which we denote by $k3^{-1}t1.20$, then π remains trivial and all knot groups stay the same except G_0 whose Alexander polynomial changes from $\Delta = t^8 - 2t^7 - 5t^5 + 13t^4 - \ldots$ to $\Delta = t^{10} - 8t^9 + 14t^8 - 2t^7 - 13t^6 + 15t^5 - \ldots$ (both polynomials are irreducible and the new 2, 3, 4, 5 torsions are given by (9), $(65, 65)$, $(3, 3, 9)$, $(899, 899)$). Assuming the latter homotopy 3$-$sphere is actually S^3, we have two a priori different smooth structures on the same underlying topological 4$-$manifold. (Recall that the Torelli preserve $A(r)$ and Freedman's theorem holds if the boundaries are the same).

Do these smooth structures differ due to, say, the arguments of Fintushel-Stern [FS]*?*

To obtain another Artin presentation for the $K3$ surface, which we denote by $\overline{k3}$ and with inverse $\overline{k3}^{-1}$, we take the pure braid in Figure 7 with $n = 2$

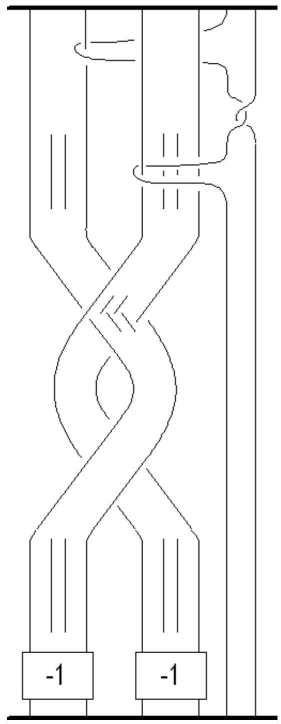


Figure 23. Modified pure braid for $E\,(n)$.

and modify it by an isotopy (the same modification applies to $E\,(n)$ in general). Take the portion of C_{21} that crosses under the right large band and intertwines with the left large band and simply slide it down to the bottom of the braid and then, using the (not drawn) trivial segments that close the braid, slide it around to the top of the braid. The result is shown in Figure 23. Of course, the framings for this braid are the same as before. Following Steps I-III above we obtain $\overline{k3}$. The isotopy of the braid preserved the order of the strands and hence the matrix $A\,(r)$ for this new presentation is exactly the same as before (Figure 21). For these Artin presentations we have $\#\overline{k3} = 6994$ and $\#\overline{k3}^{-1} = 4398$. We note that $\overline{k3}^{-1}$ is the shortest of the four Artin presentations given here for the Kummer surface.

3. Examples

Thanks to the computer friendly, simple presentations of knot groups and their peripheral structures in AP theory, examples therein need not be laboriously constructed: *they just need to be systematically discovered with MAGMA.* Due to the 'conical', universal structure of AP theory, at least in principle this can at least be done in a systematic, orderly, complete way. Thus, from the beginning AP theory, due to the fact, e.g. that framings need not be put in by hand, automatically and easily yields many of the known interesting examples of classical 3$-$manifold and knot theory: old and new. From the simplest definition of Poincaré's homology 3−sphere to examples pertaining to the Cabling conjecture [GAS]. Specifically, at the very beginning [W1] AP theory easily yields cosmetic surgery examples, Luft-Sjerve spheres with fixed point free involutions, failure of Property R in general for $\mathbb{Z}$−homology 3−spheres, in particular giving boundaries of Mazur manifolds, and nontrivial knots in homotopy 3−spheres with trivial Alexander polynomial, a phenomenon first discovered by Seifert in the early 1930s.

Using the just constructed Artin presentations of the $K3$ surface, we continue illustrating this natural, canonical flow of instructive examples.

If G is a group, by $ab\,(G,n)$ we denote the abelianizations of the subgroups of index $\leq n$ (up to conjugation) and we use MAGMA notation, e.g., $ab\,(G,4) = 1\,[0]\,,2\,[7,0]\,,4\,[2,2,0]\,,4\,[0,0]\,,$ means that G abelianizes to $\mathbb{Z}$ and has, up to conjugation, one subgroup of index 2 which abelianizes to $\mathbb{Z}_7 \times \mathbb{Z}$, no subgroups of index 3, and two subgroups of index 4 abelianizing to $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}$ and $\mathbb{Z} \times \mathbb{Z}$, respectively.

By, say, $k3^{-1}st24$, we denote the Artin presentation in $\mathcal{R}_{24}$ obtained by not changing r_i of $k3^{-1}$ for $i \leq 22$ and setting $r_{23} = x_{23}$ and $r_{24} = x_{24}$. It is clear (see end of previous section) what, say, $k3^{-1}st24t_3.22 \in \mathcal{R}_{24}$ should be. By $x_i^m r$ we denote the Artin presentation where r_i is changed to $x_i^m r_i$. The Torellis $t_1, t_2, t_3 \in \mathcal{R}_3$ and $T_4' \in \mathcal{R}_4$ are as in [W1] pp.228,229,231. Furthermore, Δ_i denotes the Alexander polynomial of k_i.

I. *Regarding the Cabling Conjecture* [GAS] *in general.* Consider $\Sigma^3\,(r)$ where $r = x_{22}^{-1}k3^{-1}st24t_3.22 \in \mathcal{R}_{24}$ ($\#r = 17301$); $\pi\,(r)$ has a balanced (non-Artin) presentation with just three generators:

$$\Big\langle a,b,c \mid c^2 = bcb,\;\; (cbc)^{-1}\,ab^6\,(cbc)^{-1}\,a^{-1}b^{-1}cbc = \\ b^{-6}ab^6 = b^{-2}\,(cbc)^{-1}\,a^{-1}cbcb^{-6}a\,(ba)^2\,cbc\Big\rangle,$$

and is therefore π−prime in the sense of [GAS], however, the $(1,-1)$ Dehn sphere of the knot k_{21} has fundamental group isomorphic to $I\,(120) * \pi_1\,(\Sigma\,(2,3,11))$.

Question: is this Dehn sphere homeomorphic to $\Sigma\,(2,3,5)\,\#\Sigma\,(2,3,11)$?

The knot k_{21} has the same Alexander polynomial as that of the granny knot in S^3, but their knot groups differ since they have different $ab\,(\;,5)$s.

The $(1,1)$ Dehn sphere of the knot k_3, where $\Delta_3 = t^2 - t + 1$, is simply connected and so $\Sigma^3\,(r)$ is a $(1,\pm1)$ Dehn sphere of a knot k in a homotopy 3−sphere with Alexander polynomial $\Delta = t^2 - t + 1$, but whose group G has a different $ab\,(\;,3)$ than that of the trefoil and is presented by:

$$G = \Big\langle a,b,c \mid bcb = cb^2c,\;\; b\left(a,\left(b^{-1}a\right)\,\hat{}\,\left(b^2\,(bc)^{-1}\,c\,(cb)^{-1}\right)\right)\Big\rangle.$$

Here, recall that in MAGMA notation $(x,y) = x^{-1}y^{-1}xy$ and $x\hat{}y = y^{-1}xy$. The homology sphere $\Sigma^3\,(r)$ is the quotient space of a free regular action of $I\,(120)$ on an M^3 with $H_1\left(M^3,\mathbb{Z}\right) = \mathbb{Z}_3^{12}$ and $ab\,(\pi\,(r)\,,15) = ab\,(I\,(120)\,,15)$, however their $ab\,(\;,20)$s differ. The Casson invariant, $\lambda\left(\Sigma^3\,(r)\right)$, of $\Sigma^3\,(r)$ is ±1.

Question: is G a knot group of S^3?

II. Tinkering with our Artin presentations for $K3$ seems to give an abundance of $\mathbb{Z}-$homology $3-$spheres with nontrivial knots where Property R fails, i.e. $G/\langle l\rangle = \mathbb{Z}$ where l is the longitude.

i) k_{10},k_{11} of $\Sigma^3(r)$ where $r = x_1^{-1}x_{22}^{-1}k3^{-1}t_2.1 \in \mathcal{R}_{22}$ ($\#r = 17916$).

ii) k_{20}, k_{22} of $\Sigma^3(r)$ where $r = x_{18}^{-1}k3^{-1}T_4'.19 \in \mathcal{R}_{22}$ ($\#r = 37009$).

iii) k_{15}, k_{22} of $\Sigma^3(r)$ where $r = x_{20}^{-1}k3^{-1}t_3.20 \in \mathcal{R}_{22}$ ($\#r = 44913$).

iv) k_{10}, k_{11} of $\Sigma^3(r)$ where $r = x_{18}^{-1}k3^{-1}t_1.9 \in \mathcal{R}_{22}$ ($\#r = 48643$). Here, $ab(G_{10}, 5) = 1[0], \ldots, 5[0],\ 5[0,0], 5[0,0,0], 5[2,0,0], 5[28371,0]$. The fundamental group of its $(1,1)$ Dehn sphere has one single subgroup of index 5 and it abelianizes to $\mathbb{Z}_{28371}$. Such large finite numbers have not appeared before in computations in AP theory. What does their appearance mean?

v) The simplest example seems to be k_{22} of $\Sigma^3(r)$ where $r = x_{20}^{-1}k3^{-1}st23t_3.21 \in \mathcal{R}_{23}$ ($\#r = 27628$). Here $\pi(r)$ and G_{22} are presented by:

$$\begin{aligned}\pi(r) &= \left\langle a, b \mid (aba)^3 = (bab)^2, (ba)^3 = \left(a^{-1}bab\right)^2 \right\rangle,\\ G_{22} &= \left\langle a, b \mid (aba)^3 = (ba)^2 (bab)^{-1} (ab)^2 \right\rangle.\end{aligned}$$

As is well known, the falsity of Property R, i.e. $G/\langle l\rangle = \mathbb{Z}$, implies that the Alexander polynomial is trivial; we also obtain an abundance of nontrivial knots with trivial Alexander polynomials in homotopy $3-$spheres (such examples were first discovered by Seifert in the early 1930s): let $r = x_{20}^{-1}k3^{-1}st24t_3.22 \in \mathcal{R}_{24}$ ($\#r = 17301$), then $\Sigma^3(r)$ is simply connected and $\Delta_{20} \equiv 1$ but $ab(G_{20}, 5) = 1[0], \ldots, 5[0]$, and $5[3,15,0]$ repeated 5 times; let $r = x_2\overline{k3}t_2.20 \in \mathcal{R}_{22}$ ($\#r = 11101$), then $\Sigma^3(r)$ is simply connected and $\Delta_1 \equiv \Delta_{12} \equiv 1$ but $ab(G_1, 5) = ab(G_{12}, 5) = 1[0], \ldots, 5[0],\ 5[0,0,0], 5[3,3,0]$. Here G_{12} is presented by:

$$\left\langle a, b, c \mid \left(a^{-1}, c\right)(c, b)(a, b)\, c = b = \left(c^{-1}, a^{-1}\right)\left(b, c^{-1}\right)(a, b)\left(c^{-1}, a^{-1}\right)\right\rangle.$$

Question: is G_{12} a knot group of S^3?

III. If $r = k3^{-1}t_3.20 \in \mathcal{R}_{22}$ ($\#r = 44550$), then $\Delta_1 \equiv 1$ and $\Delta_2 \equiv 1$ but G_1 and G_2 are not isomorphic since their $ab(\ ,5)$s differ. However, both of their $(1,1)$ Dehn spheres are simply connected. This illustrates in a different way the phenomenon that 'far away' knots in homotopy $3-$spheres can have homeomorphic $(1,1)$ Dehn spheres [Br].

Unlike with the Donaldson matrices E_8, φ_{4n}, etc., with $K3$ we obtain a much larger amount of knots with $\Delta \equiv 1$. *Is this related to the 'softness' of $K3$ as a Calabi-Yau manifold?*

4. The manifolds $W^4(r)$

We have answered in the affirmative whether all elliptic surfaces $E(n)$ appear as $W^4(r)$s. An open problem is whether every smooth, compact, connected, simply-connected 4-manifold X^4 with a connected, simply-connected boundary $\partial X^4 = M^3$ is a $W^4(r)$. (See [GS] p.344 for a related problem).

In dimension 3, AP theory obtains *all* closed, orientable, connected 3-manifolds and there seem to be no great conceptual difficulties on the horizon in obtaining all Seiberg-Witten invariants of $3-$manifolds in AP theory [L], [T] pp.viii,115. Unlike in the simplicial combinatorial case, in AP theory the *same* purely group-theoretic data that determines the $3-$manifold, namely r, also *canonically* and *holographically* determines the $4-$manifold. Hence, developing $3-$dimensional Seiberg-Witten theory in this, its correct, ultimate arena, holds greater promise in further developing also the outstanding open $4-$dimensional theory in AP theory.

Similar arguments hold for studying the smoothings of a $4-$manifold, à la Fintushel-Stern [FS], using the action of the Torelli, thus generalizing their important work. We remark that, if the 3D Poincaré conjecture were true, then by Freedman's theorem the relation between the Torelli action and smoothings would become even more direct, purely group-theoretic and pristine, perhaps too much so.

Relevant to all of the above is that although finitely presented group theory is considered a difficult subject, the undeniable metamathematical similarities of AP theory with braid theory, holographic dessins ďenfant theory, as well as numerous genuine analogies with Modern Physics, give hope for a definitive, realistic, computer approachable, holographic, and universal approach to X^4 theory [D] p.69, [W2], [W3].

Received January 09, 2003

Final version received August 18, 2003

Department of Mathematics
University of Maryland
College Park, MD 20742
USA
jsc3@math.umd.edu

hew@math.umd.edu

References

[AJ] M. Atiyah, and L. Jeffrey, *Topological Lagrangians and cohomology*, J. Geom. Phys. **7** (1990), 119–136.

[B] R.E. Borcherds, *Automorphic forms with singularities on Grassmannians*, Inv. Math. **132** (1998), 491–562.

[Br] W. Brakes, *Manifolds with multiple knot-surgery descriptions*, Math. Proc. Cambridge Phil. Soc. **87** (3), (1980), 443–448.

[CS] S. Cappell, and J. Shaneson, *Some new $4-$ manifolds*, Ann. Math. **104** (1976), 61–72.

[Di] R. Dijkgraaf, *Lectures on 4-manifolds and topological gauge theories*, Nucl. Phys. B (Proc. Suppl.) **45** B,C, (1996), 29–45.

[D] S.K. Donaldson, *The Seiberg-Witten Equations and 4-manifold Topology*, Bull. Amer. Math. Soc. **33** (1996), 45–70.

[DS] S. K. Donaldson, and D. Sullivan, *Quasiconformal 4-manifolds*, Acta Math. **163** (1989), 181–252.

[E] N. Elkies, *A characterization of the $\mathbb{Z}^n$ lattice*, Math. Res. Lett. **2** (1995), 321–326.

[F] H. Feshbach, *Group theory: Who needs it?*, Physics Today, August 1986.

[FS] R. Fintushel, and R. Stern, *Knots, links, and 4-manifolds*, Inv. Math. **134** (1998), 363–400.

[GH] R. Geroch, and J. Hartle, *Computability and physical theories*, Found. Phys. **16** (1986), 533–550.

[GS] R. Gompf, and A. Stipsicz, *4-manifolds and Kirby Calculus*, Grad. Stud. Math. **20**, Amer. Math. Soc. 1999.

[GAS] F. González-Acuña, and H. Short, *Knot surgery and primeness*, Math. Proc. Cambridge Phil. Soc. **99** (1986), 89–102.

[HKK] J. Harer, A. Kas, R. and Kirby, *Handlebody decompositions of complex surfaces*, Mem. Amer. Math. Soc. **62** (350), 1986.

[K] D. Kotschick, *Gauge theory is dead! Long live gauge theory!*, Notices Amer. Math. Soc. **42** (1995), 335–338.

[L] Y. Lim, *The equivalence of Seiberg-Witten and Casson invariants for homology* 3-*spheres*, Math. Res. Lett. **6** (1999), 631–643.

[R] A. Ranicki, *High-dimensional Knot Theory: Algebraic Surgery in Codimension 2*, Springer Monographs in Math., Springer-Verlag, 1998.

[St] R. Stern, *Computing Donaldson invariants*, IAS. Park City Math. Ser., **4**, Amer. Math. Soc. 1994, 145-194.

[S] D. Sullivan, Abstract of talk at the II AMS-SMM meeting, Guanajuato, 1995.

[TB] C. Taubes, and J. Bryan, *Donaldson-Floer theory in gauge theory and the topology of 4-manifolds*, IAS.Park City Math. Ser. **4** Amer. Math. Soc. (1994), 195–221.

[T] V. Turaev, *Introduction to Combinatorial Torsions*, Lectures in Math. ETH Zürich, 2001.

[W1] H. E. Winkelnkemper, *Artin Presentations I: Gauge Theory, 3+1 TQFTs and the Braid Groups*, J. Knot Theory Ramifications, **11** (2), (2002), 223-275.

[W2] H. E. Winkelnkemper, *Presentaciones Artinianas, Topología 4D y la Física Moderna*, Carta Informativa de la Soc. Mat. Mexicana, September 2002, 3–6.

[W3] H. E. Winkelnkemper, *Artin Presentations II: A Discrete, Purely Group-theoretic TOE*, to appear.

[Wi1] E. Witten, *Quantum Field Theory and the Jones Polynomial,* Comm. Math. Phys. **121** (1989), 351–399.

[Wi2] E. Witten, *Supersymmetric Yang-Mills theory on a 4-manifold*, J. Math. Phys. **35** (1994), 5101.

[Wi3] E. Witten, *Physical methods applied to Donaldson theory*, Functional Analysis on the Eve of the 21st Century, S. Gindikin, et al. editors, Progress in Math. **131**, Birkhäuser, 1995.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

REPRESENTING AND RECOGNIZING TORUS BUNDLES OVER $\mathbb{S}^1$

MARIA RITA CASALI

Abstract. As it is well-known, torus bundles over $\mathbb{S}^1$ are identified by means of regular integer matrices of order two (see [S]). In the present paper an algorithmic procedure is described which allows to construct, directly from any matrix $A \in GL(2;\mathbb{Z})$, an edge-coloured graph representing the torus bundle $TB(A)$ associated to A. As a consequence, five topologically undetected elements of Lins's catalogue of orientable 3-manifolds (see [L]) are finally recognized as torus bundles over $\mathbb{S}^1$.

1. Introduction

The present paper gives an approach to the study of fiber bundles with base space $\mathbb{S}^1$ and fiber T (the bidimensional torus) via edge-coloured graphs as a combinatorial PL-manifolds representation tool. In particular, an algorithmic procedure is described which allows to construct, directly from any matrix $A \in GL(2;\mathbb{Z})$, a pseudosimplicial triangulation (and, hence, the edge-coloured graph $\Gamma(A)$ visualizing it) of the torus bundle $TB(A)$ associated to A, i.e. of the quotient

$$TB(A) = \frac{T \times [0,1]}{\sim_A},$$

where the equivalence relation $\sim_A$ is given by

$$(x,0) \sim_A (\tilde{\phi}_A(x),1), \qquad \forall x \in T,$$

$\tilde{\phi}_A$ being the punctured homeomorphism $(T,x_0) \to (T,x_0)$ ($x_0 \in T$) having A as an associated matrix.

As a consequence, since edge-coloured graphs give rise to an n-dimensional combinatorial invariant for PL-manifolds - called *regular genus* (see [G] for its definition and, for example, [CG] and [BCG] for subsequent related results) -, coinciding with Heegaard genus in the 3-dimensional setting, torus bundles over $\mathbb{S}^1$ are combinatorially proved to have Heegaard genus less than or equal to three (as already obtained, via handle-decomposition, in [TO]).

On the other hand, the described construction is applied in order to topologically recognize all torus bundles belonging to Lins's catalogue (see [L]) of

2000 *Mathematics Subject Classification*: 55R10, 57N12, 57M15.

Keywords and phrases: 3-manifold, fiber-bundle, torus-bundle, monodromy, Heegaard genus, edge-coloured graph .

Work performed under the auspices of the G.N.S.A.G.A. of the C.N.R. (National Research Council of Italy) and financially supported by M.U.R.S.T. of Italy (project "Strutture geometriche delle varietà reali e complesse") and by Università degli Studi di Modena e Reggio Emilia (project "Strutture finite e modelli discreti di strutture geometriche continue").

closed connected orientable 3-manifolds represented by edge-coloured graphs up to 28 vertices (i.e. admitting a pseudosimplicial triangulation consisting of at most 28 tetrahedra). In fact, Lins's catalogue contains exactly five undetected manifolds whose fundamental group coincides with the fundamental group of a torus bundle, that is a semidirect product between $\mathbb{Z}$ and $\mathbb{Z} \times \mathbb{Z}$, induced by a matrix of $GL(2;\mathbb{Z})$. Here, the 4-coloured graph $\Gamma(A)$ associated to each one of these matrices is constructed and simplified by suitable combinatorial moves not affecting the homeomorphism class of the represented manifold, until obtaining an element of Lins's catalogue which encode exactly the undetected manifold with fundamental group $Z \cdot A$.

We point out that the combinatorial nature of the representing tools, together with the algorithmic feature of the described construction, allow to imagine a suitable implementation of the whole process; actually, a Visual Basic program[1] has been produced, automatically yielding the 4-coloured graph $\Gamma(A)$, directly from a matrix $A \in GL(2;\mathbb{Z})$, in the case that A contains a null element (which is the case occurring for each torus bundles in Lins's catalogue).

Since torus bundles frequently appear in existing catalogues of 3-manifolds (see, for example, [M] and [MP]), the author hopes that the construction obtained in the present paper will be of use in order to perform interesting comparisons between different 3-manifold complexity notions.

2. Basic notions on torus bundles over $\mathbb{S}^1$

As it is well-known, the homeotopy group of bidimensional torus T, i.e. the mapping class group of punctured homeomorphisms $(T, x_0) \to (T, x_0)$ $(x_0 \in T)$, is isomorphic to the group of automorphisms of $\pi_1(T)$, i.e. to $GL(2;\mathbb{Z})$ (see [ZVC]; Theorem 5.15.5).

This implies that any matrix $A = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix} \in GL(2;\mathbb{Z})$ induces, up to isotopy, a homeomorphism $\tilde{\phi}_A : T \to T$; if c_0 and c_1 denote, respectively, a meridian and a longitude of torus T, oriented so that their intersection number is $+1$, then $\tilde{\phi}_A$ maps c_0 (resp. c_1) into the curve $c_0' = a_{00}c_0 + a_{01}c_1$ (resp. $c_1' = a_{10}c_0 + a_{11}c_1$), fixing the (unique) intersection point $c_0 \cap c_1$.

According to [S]; sections 3.2 and 18.1, the homeomorphism $\tilde{\phi}_A$ uniquely determines a fiber bundle (with base space $\mathbb{S}^1$ and fiber T), defined as the quotient

$$TB(A) = \frac{T \times [0,1]}{\sim_A}$$

where the equivalence relation $\sim_A$ is given by

$$(x, 0) \sim_A (\tilde{\phi}_A(x), 1), \qquad \forall x \in T.$$

Note that $A \in GL(2;\mathbb{Z})$ directly implies that $\det(A) \in \{\pm 1\}$; more precisely, the torus bundle $TB(A)$ is orientable (resp. non-orientable) if and only if $\det(A) = +1$ (resp. $\det(A) = -1$). Moreover, by the classification theorem of

[1]Even if it pursues an autonomous aim, this program has been thought of in order to become a part of a wider C++ program, called *DUKE III*, which is devoted to automatic analysis, manipulation and recognition of PL-manifolds via edge-coloured graphs. Both programs are available by request to author's address.

fiber spaces (see [S]; Theorem 18.5 or [TO]; Proposition 2), two torus bundles $TB(A)$ and $TB(A')$ turn out to be equivalent if and only if A' is conjugate to either A or A^{-1} in $GL(2;\mathbb{Z})$.

The following technical lemma will be useful for our purposes, in order to restrict the class of matrices effectively inducing torus bundles.

Lemma (2.1). *Let M be a torus bundle over $\mathbb{S}^1$. Then M is equivalent to $TB(A)$ for some matrix $A = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix} \in GL(2;\mathbb{Z})$ such that*

$$a_{i0} \cdot a_{i1} \geq 0 \quad \forall i \in \{0,1\} \qquad \textit{and} \qquad A_{0j} \geq A_{1j} \quad \forall j \in \{0,1\},$$

where $A_{ij} = \max\{|a_{ij}|, 1\},\ \forall i,j \in \{0,1\}$.

Proof. First of all, note that $\det(A) \in \{\pm 1\}$ excludes the possibility that a row of A consists of concordant elements and the other one consists of discordant elements. Thus, the existence, for any matrix in $GL(2;\mathbb{Z})$, of a conjugate matrix $A = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix}$ satisfying $a_{i0} \cdot a_{i1} \geq 0 \ \ \forall i \in \{0,1\}$ directly follows from the conjugation between $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\begin{pmatrix} a & -b \\ -c & d \end{pmatrix}$:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} a & b \\ -c & -d \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} a & -b \\ -c & d \end{pmatrix}$$

On the other hand, it is easy to check that if a matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2;\mathbb{Z})$ contains all non-null elements, the case $|a| > |c|$ and $|b| < |d|$ (resp. the case $|a| < |c|$ and $|b| > |d|$) is excluded by the condition $\det(A) \in \{\pm 1\}$. Thus, the existence, for any matrix in $GL(2;\mathbb{Z})$, of a conjugate matrix $A = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix}$ satisfying $A_{0j} \geq A_{1j} \ \ \forall j \in \{0,1\}$ (in addition to $a_{i0} \cdot a_{i1} \geq 0 \ \ \forall i \in \{0,1\}$) directly follows from the conjugation between $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\begin{pmatrix} d & c \\ b & a \end{pmatrix}$:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} c & d \\ a & b \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} d & c \\ b & a \end{pmatrix}$$

Finally, note that the last conjugation allows us to assume that both conditions of the statement hold also in the case of a matrix $A = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix} \in GL(2;\mathbb{Z})$ containing null elements: in fact, if $a_{i\bar{j}} = 0$, $\det(A) \in \{\pm 1\}$ surely yields $|a_{ij'}| = |a_{i'\bar{j}}| = 1$, with $\{i, i'\} = \{\bar{j}, j'\} = \{0,1\}$, and so $A_{0\bar{j}} = A_{1\bar{j}}$. □

Definition (2.2). A matrix $A \in GL(2;\mathbb{Z})$ will be said to be *in normalized shape* if it satisfies both conditions of Lemma (2.1). The subset of $GL(2;\mathbb{Z})$ consisting of regular integer matrices of order two in normalized shape will be denoted by the symbol $\tilde{GL}(2;\mathbb{Z})$.

The following theorem collects results about the *Heegaard genus* $\mathcal{H}(M)$ of a torus bundle M over $\mathbb{S}^1$, originally obtained in [TO]; Theorem 1 and [TO]; Proposition 3, respectively; in the present paper (section 4), they will be proved as consequences of the algorithmic procedure to construct an edge-coloured graph representing $TB(A)$, for any matrix $A \in \tilde{GL}(2;\mathbb{Z})$.

PROPOSITION (2.3). a) $\mathcal{H}(M) \leq 3$ *for any torus bundle M over* $\mathbb{S}^1$.

b) *If M is a torus bundle over $\mathbb{S}^1$ with associated monodromy* $A = \begin{pmatrix} m & \epsilon \\ 1 & 0 \end{pmatrix}$ $(\epsilon \in \{1, -1\})$, *then* $\mathcal{H}(M) \leq 2$.

3. Representation of PL-manifolds by means of edge-coloured graphs

The representation theory for PL-manifolds [2] of arbitrary dimension via edge-coloured graphs has its origin within the italian school of Mario Pezzana (see [FGG] or [BCG] for details), but quickly developed with contribution of researchers from different countries (see, for example, [BM], [CV], [LM], [L], [V]).

If M^n is a compact PL n-manifold, a *coloured triangulation* of M^n may be defined as a pair $(\bar{K}, \xi)$, where:
- $\bar{K}$ is a pseudocomplex [3] triangulating M^n, with vertex set $S_0(\bar{K})$;
- $\xi : S_0(\bar{K}) \to \Delta_n = \{0, 1, \dots, n\}$ is a map (*vertex-labelling*) which is injective on the vertex-set of each n-simplex of $\bar{K}$.[4]

An $(n+1)$-*coloured graph* (Γ, γ) representing M^n is nothing but a combinatorial tool used to visualize $(\bar{K}, \xi)$. In fact, the underlying multigraph $\Gamma = \Gamma(K)$ coincides with the dual graph of K, i.e. the 1-skeleton of the ball-complex dual to K, while the edge-colouring $\gamma : E(\Gamma) \to \Delta_n$ is induced by the vertex-labelling of K: (Γ, γ) has a vertex $v(\sigma)$ for each (labelled) n-simplex $\sigma \in K$, and an i-coloured edge ($i \in \Delta_n$) connecting $v(\sigma)$ and $v(\tau)$ for every pair σ, τ of n-simplices of K sharing the $(n-1)$-face opposite to i-labelled vertex.

It is very easy to check that edge-coloured graphs are a universal tool to represent PL-manifolds: in fact, for every PL n-manifold M^n, the existence of a coloured triangulation (and, hence, an edge-coloured graph) representing M^n may be directly proved by considering the first baricentric subdivision of any simplicial triangulation of M^n, and by labelling every vertex by the dimension of the simplex whose barycenter it is.

Of course, for any fixed n-manifold M^n, many edge-coloured graphs exist, which represent M^n. In particular, edge-coloured graphs which coincide up to permutations of the vertex set and/or of the colour set (i.e. the so called *colour-isomorphic graphs*) do obviously represent the same manifold. In [L] and [CG], an alphanumerical *code* $c(\Gamma)$ is defined for any coloured graph Γ, which allows to effectively recognize colour-isomorphic graphs.[5]

---

[2] For basic notions on piecewise-linear (PL) category, we refer to [RS]. Throughout the present paper, all PL-manifolds are assumed to be closed and connected, unless otherwise stated.

[3] Remember that - according to [HW] - a pseudocomplex is a ball-complex which differs from a simplicial complex because its "h-simplices" may intersect in more than one face.

[4] In case $\partial M^n \neq \emptyset$, it is also required every n-labelled vertex to be internal in $\bar{K}$.

[5] Note that code computation may be easily implemented; for example, DUKE III program contains a suitable "code computation" function.

Within the representation theory for PL-manifolds by edge-coloured graphs, a combinatorial invariant - called *regular genus* - has been introduced and deeply investigated (see [G], [FG$_1$], [CG] and their bibliography). It may be thought of as a natural extension of the notion of *genus* for surfaces and of *Heegaard genus* for 3-manifolds; in particular, for any 3-dimensional manifold M^3, the equality $\mathcal{G}(M^3) = \mathcal{H}(M^3)$ holds, where $\mathcal{G}(M^3)$ denotes the regular genus of M^3 (and $\mathcal{H}(M^3)$ denotes - as already recalled - its Heegaard genus).

According to [G], it is known that, for every $(n+1)$-coloured graph (Γ, γ) representing a closed orientable (resp. non-orientable) n-manifold M^n and for every cyclic permutation $\epsilon = (\epsilon_0, \epsilon_1, \ldots, \epsilon_n = n)$ of Δ_n, there exists a regular embedding of (Γ, γ) into an orientable (resp. non-orientable) surface F_ϵ; furthermore, the genus $\rho_\epsilon(\Gamma)$ of F_ϵ (resp. half the genus $\rho_\epsilon(\Gamma)$ of F_ϵ) may be easily computed by the following formula, where $g_{i,j}$ denotes the number of $\{\epsilon_i, \epsilon_j\}$-coloured cycles of (Γ, γ) and p is the number of vertices of (Γ, γ):

$$\sum_{j \in \mathbb{Z}_{n+1}} g_{j,j+1} + (1-n) \cdot \frac{p}{2} \;=\; 2 - 2\rho_\epsilon(\Gamma)$$

In particular, if $n = 3$, $\rho_\epsilon(\Gamma)$ may be obtained through a simpler formula, too:

$$\rho_\epsilon(\Gamma) \;=\; g_{0,2} - g_{\hat{1}} - g_{\hat{3}} + 1$$

where $g_{\hat{j}}$ $(j \in \Delta_3)$ is the number of connected components of $\Gamma_{\hat{\epsilon}_j} = (V(\Gamma), \gamma^{-1}$ $(\Delta_3 - \{j\}))$.

The regular genus $\rho(\Gamma)$ of an $(n+1)$-coloured graph (Γ, γ) is, by definition, the minimum value of $\rho_\epsilon(\Gamma)$ over all cyclic permutations ϵ of Δ_n. Finally, the *regular genus* of a PL-manifold M^n is defined as:

$$\mathcal{G}(M^n) = \min\{\rho(\Gamma) \;/ (\Gamma, \gamma) \text{ is an (n+1)-coloured graph representing } M^n\}.$$

4. From matrices to 4-coloured graphs representing torus bundles

The following paragraph will be entirely devoted to show how to construct edge coloured graphs representing torus bundles, directly from integer matrices inducing them.

THEOREM (4.1). *Let $A \in \tilde{GL}(2;\mathbb{Z})$. An algorithmic procedure exists, which allows to directly construct a 4-coloured graph $\Gamma(A)$ representing the torus bundle $TB(A)$ with monodromy induced by A.*

Proof. The statement is directly proved by construction, via the following steps.

First step: We construct two cell-complexes $\bar{K}_0$ and $\bar{K}_1$ triangulating the torus T, so that a bijective cell-map $\bar{\Phi}_A : \bar{K}_0 \to \bar{K}_1$ exists, with $|\bar{\Phi}_A| = \tilde{\phi}_A$.

In order to obtain $\bar{K}_1$, it is sufficient to consider on $\frac{I \times I}{\sim}$ the geometrical realization of c'_i $(i \in \{0,1\})$ consisting of the $A_{i0} + A_{i1} - 1$ edges, parallel to $\vec{v} \equiv (a_{i0}, a_{i1})$, having as end-points the $A_{i1}+1$ vertices on $I \times \{0\} = I \times \{1\}$ of first coordinate $\frac{h}{A_{i1}}$, $h \in \{0, \ldots, A_{i1}\}$ and the $A_{i0}+1$ vertices on $\{0\} \times I = \{1\} \times I$ of second coordinate $\frac{k}{A_{i0}}$, $k \in \{0, \ldots, A_{i0}\}$.[6]

[6]Note that, in case $a_{ij} = 0$, the geometrical realization of c'_i simply coincides with the canonically identified edges $I \times \{0\} = I \times \{1\}$ (if $j = 1$) or $\{0\} \times I = \{1\} \times I$ (if $j = 0$).

On the other hand, let $\bar{K}_0$ be the cellular subdivision of $\frac{I\times I}{\sim}$ constructed in the following way:

• let us consider the $A_{00}+A_{01}$ vertices on $I\times\{0\}=I\times\{1\}$ of first coordinate $\frac{r}{A_{00}+A_{01}-1}$, $r\in\{0,\dots,A_{00}+A_{01}-1\}$ and the $A_{10}+A_{11}$ vertices on $\{0\}\times I=\{1\}\times I$ of second coordinate $\frac{s}{A_{10}+A_{11}-1}$, $s\in\{0,\dots,A_{10}+A_{11}-1\}$; [7]

• let us consider, for every vertex on $\{0\}\times I=\{1\}\times I$, an edge internal to $I\times I$ parallel to $\vec{w}\equiv(A_{10}+A_{11}-1,\mu(A)(A_{00}+A_{01}-1))$, where $\mu(A)=(-1)^{a_{0j}\cdot a_{1j}}$; [8]

• finally, in case $a_{i,j}\neq 0\ \forall i,j\in\mathbb{N}_2$, let us also consider the $A_{00}+A_{01}-A_{10}-A_{11}$ edges internal to $I\times I$, having both the end-points on $I\times\{0\}=I\times\{1\}$, parallel to $\vec{w}'\equiv(A_{10}+A_{11},\mu(A)(A_{00}+A_{01}-1))$.

Note that both $\bar{K}_0$ and $\bar{K}_1$ consist of $\sum_{i,j\in\{0,1\}}|a_{ij}|-1$ cells, among which $4-2n_0(A)$ are triangular cells and $\sum_{i,j\in\{0,1\}}|a_{ij}|-5+2n_0(A)$ are quadrangular ones, $n_0(A)$ being the number of null elements in A; moreover, the required bijective cell-map $\bar{\Phi}_A:\bar{K}_0\to\bar{K}_1$, with $|\bar{\Phi}_A|=\tilde{\phi}_A$, is easily induced by $\tilde{\phi}_A(c_i)=c_i'$ (with correct orientations).

Second step: We construct two coloured triangulations K_0 and K_1 of the torus T, so that a bijective coloured simplicial map $\Phi_A:K_0\to K_1$ exists, with $|\Phi_A|=\tilde{\phi}_A$.

K_0 (resp. K_1) is simply obtained from $\bar{K}_0$ (resp. $\bar{K}_1$) by performing a baricentric subdivision and by labelling every vertex of K_0 (resp. K_1) with the dimension of the corresponding cell of $\bar{K}_0$ (resp. $\bar{K}_1$). Hence, the bijective cell-map $\bar{\Phi}_A:\bar{K}_0\to\bar{K}_1$ canonically induces a bijective coloured simplicial map $\Phi_A:K_0\to K_1$, with the property $|\Phi_A|=\tilde{\phi}_A$.

Third step: We construct a coloured triangulation $\bar{K}$ of the product $T\times I$, so that $\bar{K}|_{T\times\{0\}}=K_0$ and $\bar{K}|_{T\times\{1\}}=K_1$.

Let $\bar{H}_0$ (resp. $\bar{H}_1$) denote the cellular subdivision of $I\times I$ inducing, via canonical boundary identification, the cellular subdivision $\bar{K}_0$ (resp. $\bar{K}_1$) of $T=\frac{I\times I}{\sim}$. Then, let $\bar{H}$ be the cellular subdivision of the cube $I\times I\times I$ consisting of exactly one 3-cell, coinciding with $\bar{H}_0$ (resp. $\bar{H}_1$) on $I\times I\times\{0\}$ (resp. $I\times I\times\{1\}$) and containing exactly one 2-cell for each other face of $I\times I\times I$. Then, let H be the coloured simplicial triangulation of $I\times I\times I$ obtained from $\bar{K}$ by performing a baricentric subdivision and by labelling every vertex with the dimension of the cell of $\bar{K}$ whose baricenter the given vertex is. It is now easy to check that $\bar{K}$ is simply obtained from H by canonical identification of opposite faces $\{0\}\times I\times I$ and $\{1\}\times I\times I$ (resp. $I\times\{0\}\times I$ and $I\times\{1\}\times I$).

Fourth step: A coloured triangulation K_A of the torus bundle $TB(A)$ is obtained from $\bar{K}$ by identifying faces $\bar{K}|_{T\times\{0\}}$ and $\bar{K}|_{T\times\{1\}}$ according to Φ_A.

In order to complete the algorithmic construction, it is now sufficient to consider the edge coloured graph $\Gamma(A)$ such that $\Gamma(A)=\Gamma(K_A)$ (as described in the previous paragraph). □

---

[7] Note that the edge $I\times\{0\}=I\times\{1\}$ (resp. $\{0\}\times I=\{1\}\times I$) of $\bar{K}_0$ results to be subdivided into $A_{00}+A_{01}-1$ (resp. $A_{10}+A_{11}-1$) edges, as well as the geometrical realization of c_0' (resp. c_1') in $\bar{K}_1$.

[8] The assumption $det(A)\in\{+1,-1\}$ ensures that either $a_{0j}\cdot a_{1j}\geq 0\ \forall j\in\{0,1\}$ or $a_{0j}\cdot a_{1j}\leq 0\ \forall j\in\{0,1\}$ surely holds.

Example (I). If $A = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \in GL(2;\mathbb{Z})$, then $TB(A)$ is equivalent to $TB(A')$, with $A' = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \in \tilde{GL}(2;\mathbb{Z})$ (see Lemma (2.1)). Then, the first step of the described algorithm yields the cell-complexes $\bar{K}_0$ and $\bar{K}_1$ triangulating the torus T depicted in Figure 1(a) (where the bijective cell-map $\bar{\Phi}_{A'} : \bar{K}_0 \to \bar{K}_1$ with $|\bar{\Phi}_{A'}| = \tilde{\phi}_{A'}$ is visualized by labelling pairs of corresponding cells by equal symbols (for example, (x, x')). Furthermore, Figure 1(b) illustrates the coloured triangulations K_0 and K_1 of the torus T obtained in the second step (where equally labelled simplices are assumed to correspond each other in the bijective coloured simplicial map $\Phi_{A'} : K_0 \to K_1$, with $|\Phi_{A'}| = \tilde{\phi}_{A'}$). Finally, in Figure 1(c) the boundary of the coloured simplicial triangulation H of $I \times I \times I$ obtained in the third step is depicted, and equally labelled 2-simplices indicate boundary identifications necessary to yield $K_{A'}$ from $H = v * \partial H$, v being the unique inner 3-coloured vertex of H (forth step). The resulting edge-coloured graph $\Gamma(A') = \Gamma(K_{A'})$ is shown in Figure 1(d), where 3-coloured edges are understood through equal labelling of pairs of 3-adjacent vertices.

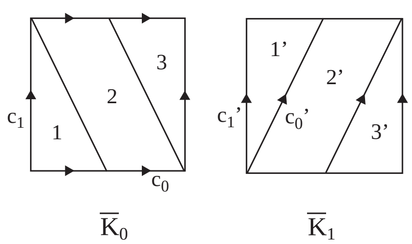


Figure 1(a)

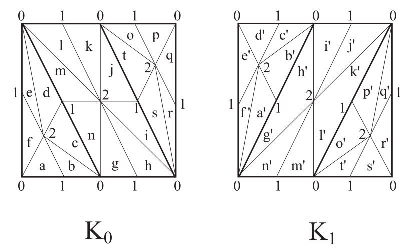


Figure 1(b)

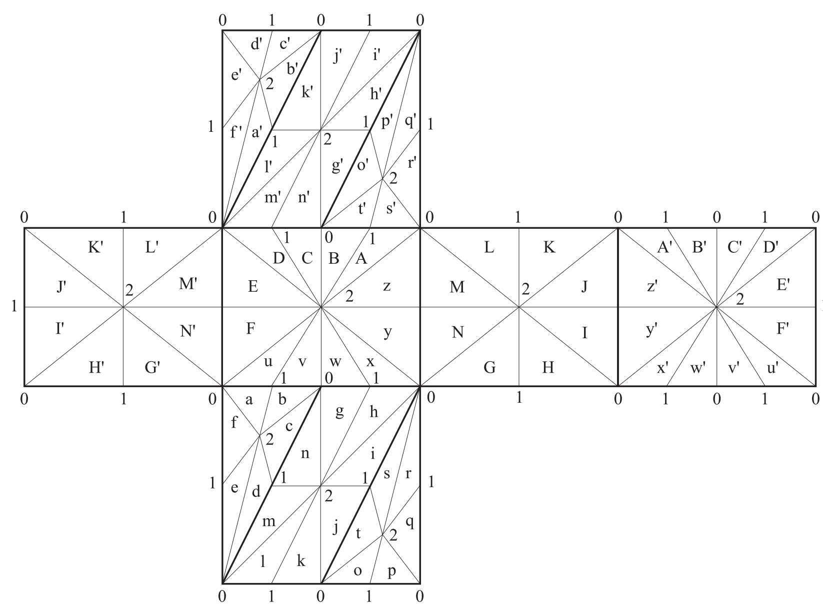


Figure 1(c)

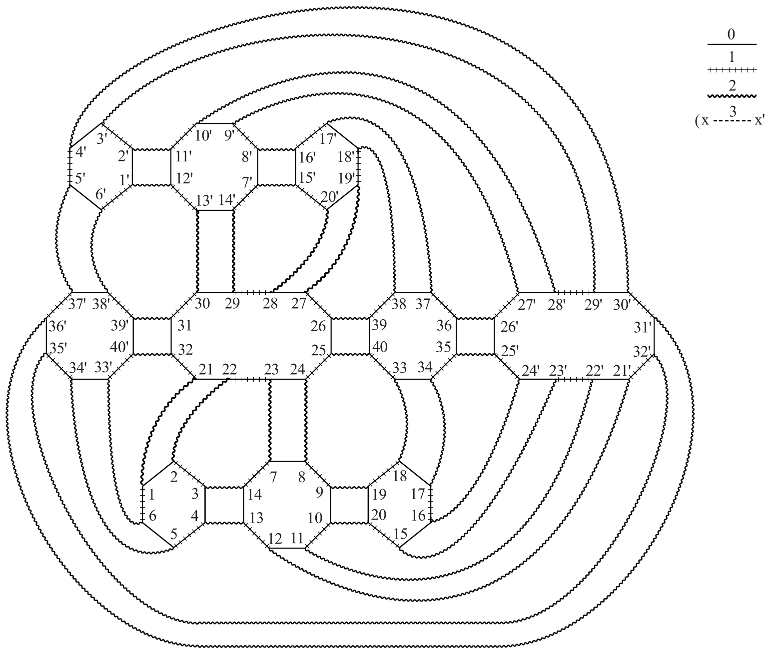


Figure 1(d)

Example (II). If $A = \begin{pmatrix} -2 & 1 \\ -3 & 2 \end{pmatrix} \in GL(2;\mathbb{Z})$, then $TB(A)$ is equivalent to $TB(A')$, with $A' = \begin{pmatrix} 2 & 3 \\ -1 & -2 \end{pmatrix} \in \tilde{GL}(2;\mathbb{Z})$ (see Lemma (2.1)). Then, the first step of the described algorithm yields the cell-complexes $\bar{K}_0$ and $\bar{K}_1$ triangulating the torus T depicted in Figure 2(a) (where the bijective cell-map $\bar{\Phi}_{A'} : \bar{K}_0 \to \bar{K}_1$ with $|\bar{\Phi}_{A'}| = \tilde{\phi}_{A'}$ is visualized by equally labelling of corresponding cells). Further, Figure 2(b) illustrates the coloured triangulations K_0 and K_1 of the torus T obtained in the second step (where equally labelled 2-simplices are assumed to correspond each other in the bijective coloured simplicial map $\Phi_{A'} : K_0 \to K_1$, with $|\Phi_{A'}| = \tilde{\phi}_{A'}$). Subsequent steps of the described algorithm follow as in Example (I).

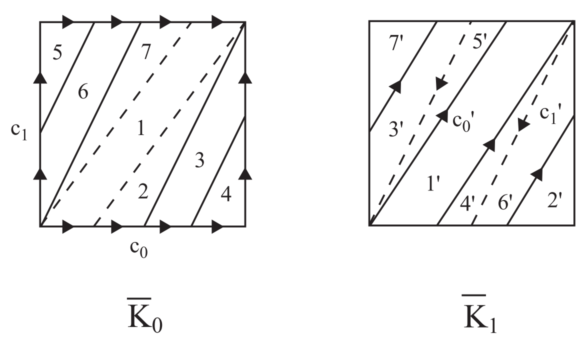
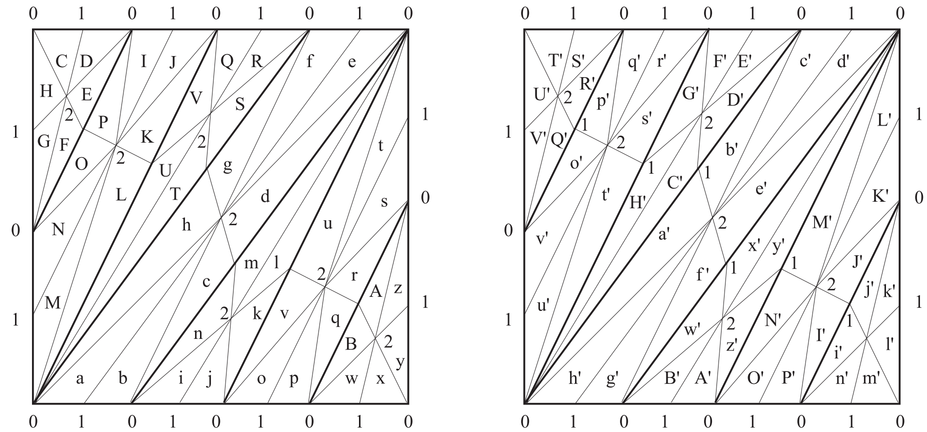


Figure 2(a)



K_0 K_1

Figure 2(b)

Remark (A). Note that, by virtue of Lemma (2.1), any torus bundle over $\mathbb{S}^1$ turns out to admit a 4-coloured graph $\Gamma(A)$ (obtained as an output of the algorithmic procedure of Theorem (4.1), for a suitable $A \in \tilde{GL}(2;\mathbb{Z})$) representing it.

Remark (B). For every $A \in \tilde{GL}(2;\mathbb{Z})$), the 4-coloured graph $\Gamma(A)$ representing $TB(A)$ enjoys the following combinatorial features (which may be easily checked by direct computation via the corresponding geometrical properties of the coloured triangulation K_A):

$2p = \#V(\Gamma(A)) = 8(3\sum_{i,j\in\{0,1\}}|a_{ij}| - 4 + 2n_0(A))$;
$g_{01} = 2\sum_{i,j\in\{0,1\}}|a_{ij}| + 2$;
$g_{02} = g_{03} = g_{13} = 2(3\sum_{i,j\in\{0,1\}}|a_{ij}| - 4 + 2n_0(A))$;
$g_{12} = 4(\sum_{i,j\in\{0,1\}}|a_{ij}| - 2 + n_0(A))$;
$g_{23} = 2(2\sum_{i,j\in\{0,1\}}|a_{ij}| - 3 + n_0(A))$;
$g_{\hat{0}} = \sum_{i,j\in\{0,1\}}|a_{ij}| - 3 + n_0(A)$;
$g_{\hat{1}} = 2\sum_{i,j\in\{0,1\}}|a_{ij}| - 3 + n_0(A)$;
$g_{\hat{2}} = 1 + \sum_{i,j\in\{0,1\}}|a_{ij}|$;
$g_{\hat{3}} = 1$.

Proposition (4.2). *For every $A \in \tilde{GL}(2;\mathbb{Z})$, the following relations hold:*

a) $\rho(\Gamma(A)) = \sum_{i,j\in\{0,1\}}|a_{i,j}| + 1$;
b) $\mathcal{G}(TB(A)) \leq 3$.

Furthermore, if $A \in \tilde{GL}(2;\mathbb{Z})$ is such that $a_{11} = 0$, then

c) $\mathcal{G}(TB(A)) \leq 2$.

Proof. If $\bar{\epsilon} = (0,2,1,3)$, a direct computation yields:

$$\rho_{\bar{\epsilon}}(\Gamma(A)) = g_{01} - g_{\hat{2}} - g_{\hat{3}} + 1 = 2\sum_{i,j\in\{0,1\}}|a_{ij}| + 2 - (1 + \sum_{i,j\in\{0,1\}}|a_{ij}|) - 1 + 1 =$$

$$= \sum_{i,j\in\{0,1\}}|a_{ij}| + 1$$

On the other hand, it is easy to check that, for any permutation ϵ' of Δ_3, $\rho_{\epsilon'}(\Gamma(A)) \geq \rho_{\bar{\epsilon}}(\Gamma(A))$ holds; thus, statement (a) follows.

In order to prove statement (b), it is necessary to note that, if σ, σ' are two cells of the cellular triangulation $\bar{H}_0$ (resp. $\bar{H}_1$) of $I \times I$, sharing a common boundary edge e, with $e \notin \partial(I \times I)$, then the $\{0,1\}$-coloured cycle of $\Gamma(A)$, dual to the $\{2,3\}$-labelled edge of K_A having the baricenter of σ (resp. σ') as an end-point, has exactly one common vertex with any $\{2,3\}$-coloured cycle of $\Gamma(A)$, dual to one of the ($\{0,1\}$-labelled) edges of K_A subdividing e: this means that any such common vertex identifies a so called *generalized dipole*, which is a combinatorial structure that may be easily eliminated by a finite sequence of elementary moves on edge-coloured graphs, yielding a new graph, with one less $\{0,1\}$-coloured cycles, representing the same 3-manifold (see [FG$_2$] for details). It is not difficult to check that, since $\bar{H}_0$ (resp. $\bar{H}_1$) contains $\sum_{i,j\in\{0,1\}} |a_{ij}| - 2$ edges not belonging to $\partial(I \times I)$, the combinatorial structure of K_A allows to perform in $\Gamma(A)$, for every $A \in \tilde{GL}(2;\mathbb{Z})$, a finite sequence of $\sum_{i,j\in\{0,1\}} |a_{ij}| - 2$ "independent" generalized dipole eliminations, giving rise to a new 4-coloured graph $\Gamma'(A)$ representing $TB(A)$, so that

$$\rho_{\bar{\epsilon}}(\Gamma'(A)) = \rho_{\bar{\epsilon}}(\Gamma(A)) - (\sum_{i,j\in\{0,1\}} |a_{ij}| - 2) = 3.$$

This completes the proof of statement (b).

Finally, let us consider the case of a matrix $A \in \tilde{GL}(2;\mathbb{Z})$ with $a_{11} = 0$. A direct check allows us to verify that, if τ is the cell of $\bar{H}$ corresponding to the face $\{0\} \times I \times I$ and f is the $\{2,3\}$-labelled edge of K_A having the baricenter of τ as an end-point, then the sequence of generalized dipoles transforming $\Gamma(A)$ into $\Gamma'(A)$ does not involve the vertices of the $\{0,1\}$-coloured cycle of $\Gamma(A)$, dual to f. Moreover, both in $\Gamma(A)$ and in $\Gamma'(A)$, f has exactly one common vertex with any $\{2,3\}$-coloured cycle, dual to one of the ($\{0,1\}$-labelled) edges of K_A triangulating $\{0\} \times I \times \{0\}$. Hence, the additional hypothesis $a_{11} = 0$ allows us to perform a further generalized dipole elimination, yielding a new graph $\Gamma''(A)$ representing $TB(A)$; statement (c) now directly follows:

$$a_{11} = 0 \quad \Longrightarrow \quad \rho_{\bar{\epsilon}}(\Gamma''(A)) = \rho_{\bar{\epsilon}}(\Gamma'(A)) - 1 = 2.$$

□

We are now able to easily prove the already quoted upper bound results on Heegaard genus of torus bundles.

Proof of Proposition (2.3). Statement (a) directly follows from Proposition (4.2)(b), via Lemma (2.1) and Remark (A).

On the other hand, statement (b) is a direct consequence of Proposition (4.2)(c), applied either to $A = \begin{pmatrix} m & \epsilon \\ 1 & 0 \end{pmatrix}$ or to the conjugate matrix $A' = \begin{pmatrix} m & -\epsilon \\ -1 & 0 \end{pmatrix}$ (see Lemma (2.1)). □

5. Recognition of torus bundles among elements of Lins's catalogue

In [LS] and [L] (resp. [C]) a complete catalogue $\mathcal{C}^{(28)}$ (resp. $\tilde{\mathcal{C}}^{(26)}$) of orientable (resp. nonorientable) 3-manifolds admitting coloured triangulations up to 28 (resp. 26) tetrahedra is obtained, and its elements are deeply analyzed; as a consequence, the following result is proved:

Proposition (5.1). a) [LS] - [L] *Exactly 69 non-homeomorphic prime orientable 3-manifolds exist, which admit a coloured triangulation consisting of at most 28 tetrahedra;*

b) [C] *Exactly 7 non-homeomorphic prime non-orientable 3-manifolds exist, which admit a coloured triangulation consisting of at most 26 tetrahedra.* □

In particular, all elements of $\tilde{\mathcal{C}}^{(26)}$ are topologically recognized [9], while some elements of $\mathcal{C}^{(28)}$ are simply proved to be non-homeomorphic manifolds by means of the computation of their fundamental group.

The first result about recognition of torus bundles among catalogues $\mathcal{C}^{(28)}$ and $\tilde{\mathcal{C}}^{(26)}$ concerns matrices $A \in GL(2,\mathbb{Z})$ with two null elements:

Proposition (5.2). *All orientable (resp. non-orientable) torus bundles having monodromy $A \in GL(2;\mathbb{Z})$ with $n_0(A) = 2$ belong to catalogue $\mathcal{C}^{(28)}$ (resp. $\tilde{\mathcal{C}}^{(26)}$). In particular, they are the following euclidean manifolds (whose corresponding crystallographic groups are indicated according to notations of international table of crystallography* [IT]*):*

$$TB(\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}) \cong \mathbb{S}^1 \times \mathbb{S}^1 \times \mathbb{S}^1 = \frac{\mathbb{E}^3}{G_1}, \quad \textit{with } G_1 = P1 \quad (\textit{corresponding to } r_1^{24} \in \mathcal{C}^{(28)})$$
$$TB(\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}) \cong \frac{\mathbb{E}^3}{G_2}, \quad \textit{with } G_2 = P2_1 \quad (\textit{corresponding to } r_4^{26} \in \mathcal{C}^{(28)})$$
$$TB(\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}) \cong \frac{\mathbb{E}^3}{G_4}, \quad \textit{with } G_4 = P4_1 \quad (\textit{corresponding to } r_{11}^{26} \in \mathcal{C}^{(28)})$$
$$TB(\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}) \cong \frac{\mathbb{E}^3}{B_1}, \quad \textit{with } B_1 = Pb \quad (\textit{corresponding to } \tilde{r}_{111}^{24} \in \tilde{\mathcal{C}}^{(26)})$$
$$TB(\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}) \cong \frac{\mathbb{E}^3}{B_2}, \quad \textit{with } B_2 = Bb \quad (\textit{corresponding to } \tilde{r}_{112}^{24} \in \tilde{\mathcal{C}}^{(26)})$$

Proof. A direct check allows us to prove that the only matrices $A \in GL(2;\mathbb{Z})$ with $n_0(A) = 2$ are (up to conjugation): $A_1 = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$, $A_2 = \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$, $A_3 = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$, $A_4 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$, and $A_5 = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$. Moreover, in case $n_0(A) = 2$, the first step of the algorithmic procedure of Theorem (4.1) yields two cell-complexes $\bar{K}_0$ and $\bar{K}_1$ consisting of just one 2-cell; so, for each $i \in \mathbb{N}_5$, the whole procedure may be very easily applied, giving rise to a 48 order 4-coloured graph $\Gamma(A_i)$ representing the associated torus bundle $TB(A_i)$. Now, a standard sequence of dipole moves may be performed (for example, by making use of

[9] The manifolds involved are proved to be (see [C]; Theorem I): the four euclidean non-orientable 3-manifolds, the non-trivial $\mathbb{S}^2$ bundle over $\mathbb{S}^1$, the topological product between the real projective plane $\mathbb{RP}^2$ and $\mathbb{S}^1$, and the torus bundle over $\mathbb{S}^1$, with monodromy induced by matrix $\begin{pmatrix}0 & 1\\ 1 & -1\end{pmatrix}$.

the corresponding function of DUKE III program), until we obtain a 4-coloured graph belonging to catalogues $\mathcal{C}^{(28)}$ (in case $i \in \{1,2,3\}$) or $\tilde{\mathcal{C}}^{(26)}$ (in case $i \in \{4,5\}$). Complete classification results about these catalogues allow us to directly prove the statement. $\square$

Finally, we will apply the above described algorithmic construction of edge coloured graphs representing torus bundles, in order to topologically recognize five further manifolds belonging to Lins's catalogue as torus bundles over $\mathbb{S}^1$. Note that they are the only topologically undetected elements of Lins's catalogue whose fundamental groups coincide with fundamental groups of torus bundles over $\mathbb{S}^1$.[10]

Proposition (5.3). (a) *The orientable 3-manifold corresponding to* $r^{26}_{14} \in \mathcal{C}^{(28)}$ *(whose fundamental group is* $\mathbb{Z} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$*) is the torus bundle* $TB(A)$, *with* $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$;

(b) *the orientable 3-manifold corresponding to* $r^{28}_{5} \in \mathcal{C}^{(28)}$ *(whose fundamental group is* $\mathbb{Z} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 3 \end{pmatrix}$*) is the torus bundle* $TB(A)$, *with* $A = \begin{pmatrix} 0 & 1 \\ -1 & 3 \end{pmatrix}$;

(c) *the orientable 3-manifold corresponding to* $r^{28}_{10} \in \mathcal{C}^{(28)}$ *(whose fundamental group is* $\mathbb{Z} \cdot \begin{pmatrix} 0 & 1 \\ -1 & -3 \end{pmatrix}$*) is the torus bundle* $TB(A)$, *with* $A = \begin{pmatrix} 0 & 1 \\ -1 & -3 \end{pmatrix}$;

(d) *the orientable 3-manifold corresponding to* $r^{28}_{42} \in \mathcal{C}^{(28)}$ *(whose fundamental group is* $\mathbb{Z} \cdot \begin{pmatrix} -1 & 0 \\ 2 & -1 \end{pmatrix}$*) is the torus bundle* $TB(A)$, *with* $A = \begin{pmatrix} -1 & 0 \\ 2 & -1 \end{pmatrix}$;

(e) *the orientable 3-manifold corresponding to* $r^{28}_{280} \in \mathcal{C}^{(28)}$ *(whose fundamental group is* $\mathbb{Z} \cdot \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$*) is the torus bundle* $TB(A)$, *with* $A = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$.

Proof. (a) Let $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$; even though $A \in \tilde{GL}(2;\mathbb{Z})$, let us apply the algorithmic procedure of the previous section to its conjugate matrix $A' = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \tilde{GL}(2;\mathbb{Z})$. The coloured pseudocomplex $K_{A'}$ turns out to be obtained by starring from an inner 3-labelled vertex the coloured complex depicted in Figure 3, where 2-simplices labelled x, x' have to be identified.

[10] Van Kampen's Theorem allows us to directly check that the fundamental group of the torus bundle $TB(A)$ is the semidirect product between $\mathbb{Z}$ and $\mathbb{Z} \times \mathbb{Z}$ induced by the matrix A, i.e. the group usually denoted by $\mathbb{Z} \cdot A$.

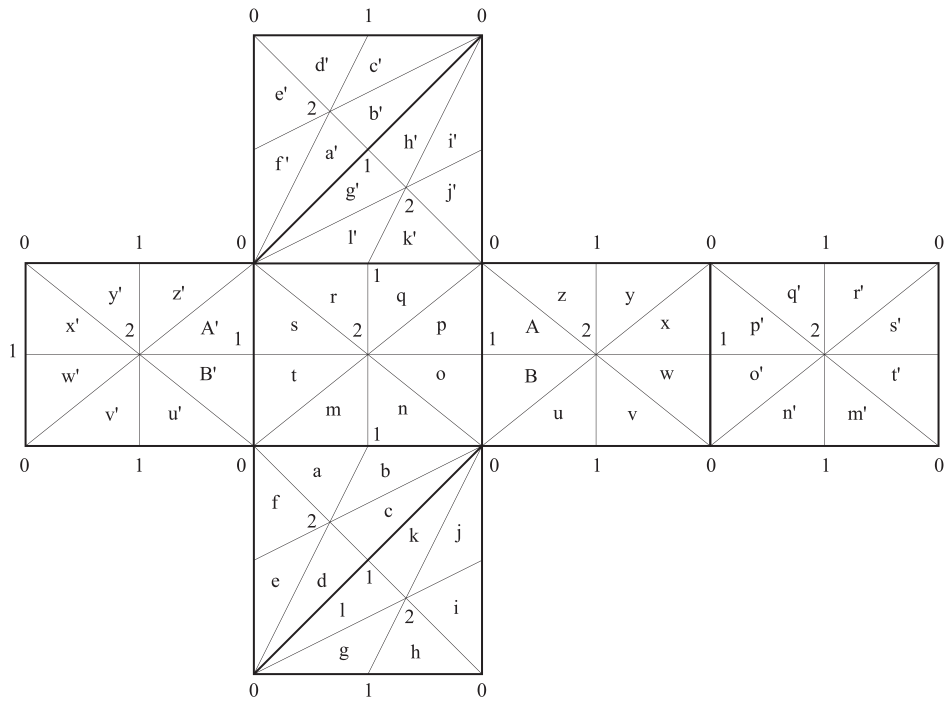


Figure 3

It is not difficult to check that the associated order 56 4-coloured graph $\Gamma(A')$ may be transformed by a finite sequence of dipole moves (for example, by making use of the corresponding function of DUKE III program) into an order 28 4-coloured graph $\Gamma'(A')$ having code

$$c(\Gamma'(A')) = \ dabchefgkijnlmknmedclihgajfbgfjnmkadlcbieh;$$

since this code identifies, up to permutation of vertices and colours, the element $r^{28}_{4478} \in \mathcal{C}^{(28)}$, and since Lins's classification ensures the represented 3-manifold to be the same as $r^{26}_{14} \in \mathcal{C}^{(28)}$, part (a) of the statement follows.

(b) Let $A = \begin{pmatrix} 0 & 1 \\ -1 & 3 \end{pmatrix}$; since $A \notin \tilde{GL}(2;\mathbb{Z})$, we apply the algorithmic procedure of the previous section to its conjugate matrix $A' = \begin{pmatrix} 3 & 1 \\ -1 & 0 \end{pmatrix} \in \tilde{GL}(2;\mathbb{Z})$ (see Lemma (2.1) in order to prove the conjugation). The coloured pseudocomplex $K_{A'}$ turns out to be obtained by starring from an inner 3-labelled vertex the coloured complex depicted in Figure 4, where 2-simplices labelled x, x' have to be identified.

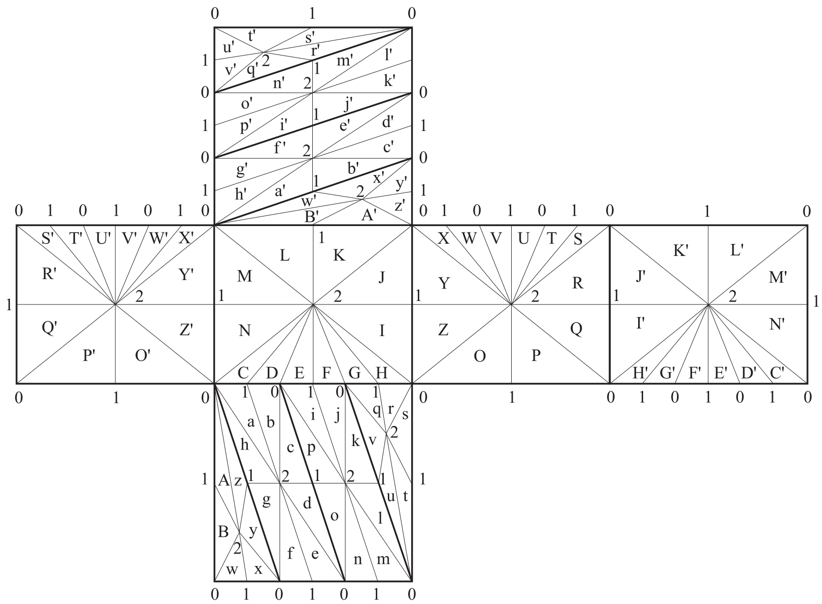


Figure 4

It is not difficult to check that the associated order 104 4-coloured graph $\Gamma(A')$ may be transformed by a finite sequence of dipole moves (for example, by making use of the corresponding function of DUKE III program) into an order 32 4-coloured graph $\Gamma'(A')$ having code

$$c(\Gamma'(A')) = fabcdejghimklpnomlonpgfebkjiadchojihkpalcbednmgf.$$

A direct computation allows us to prove $\Gamma'(A')$ admits a so called *cluster*, which is a combinatorial structure that may be easily eliminated by a finite sequence of elementary moves on edge-coloured graphs, yielding a new graph $\Gamma''(A')$ with two less vertices, representing the same 3-manifold (see Figure 5, or [L]; Proposition 24] for details). On the other hand, $\Gamma''(A')$ may be further simplified via a generalized dipole elimination and two dipole eliminations (for example, by making use of the corresponding functions of DUKE III program), so as to obtain the order 28 4-coloured graph $\Gamma'''(A')$ having code

$$c(\Gamma'''(A')) = \ dabchefgkijnlmknmedclihgajfbgfjlnkadmcbhie;$$

since this code identifies, up to permutation of vertices and colours, the element $r^{28}_{4538} \in \mathcal{C}^{(28)}$, and since Lins's classification ensures the represented 3-manifold to be the same as $r^{28}_{5} \in \mathcal{C}^{(28)}$, part (b) of the statement follows.

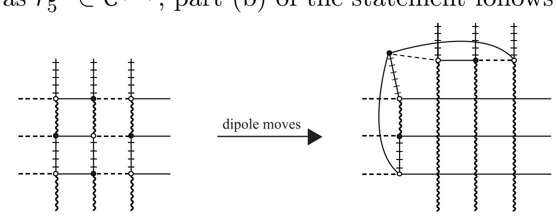


Figure 5

(c) Let $A = \begin{pmatrix} 0 & 1 \\ -1 & -3 \end{pmatrix}$; since $A \notin \tilde{GL}(2;\mathbb{Z})$, we apply the algorithmic procedure of the previous section to its conjugate matrix $A' = \begin{pmatrix} -3 & -1 \\ 1 & 0 \end{pmatrix} \in \tilde{GL}(2;\mathbb{Z})$ (see Lemma (2.1) in order to prove the conjugation). The coloured pseudocomplex $K_{A'}$ turns out to be obtained by starring from an inner 3-labelled vertex the coloured complex depicted in Figure 6, where 2-simplices labelled x, x' have to be identified.

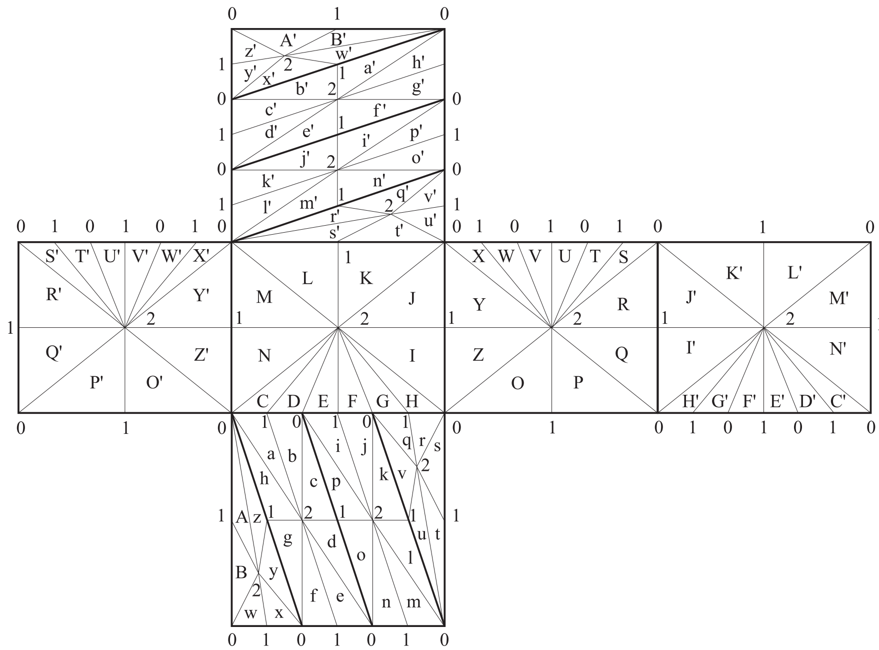


Figure 6

It is not difficult to check that the associated order 104 4-coloured graph $\Gamma(A')$ may be transformed by a finite sequence of dipole moves (for example, by making use of the corresponding function of DUKE III program) into an order 30 4-coloured graph $\Gamma'(A')$ having code

$$c(\Gamma'(A')) = eabcdifghljkomnlongfedcjimakhbnhkjomlbafcgdie.$$

A direct computation allows us to prove $\Gamma'(A')$ admits a *cluster*, whose elimination yields a new graph $\Gamma''(A')$ with two less vertices (i.e. 28 vertices), representing the same 3-manifold (see Figure 5, or [L]; Proposition 24 for details). Since its code

$$c(\Gamma''(A')) = \; dabcgefjhilknmjlnedchgmaibkfnfihkbalcejmdg$$

identifies, up to permutation of vertices and colours, the element $r^{28}_{206} \in \mathcal{C}^{(28)}$, and since Lins's classification ensures the represented 3-manifold to be the same as $r^{28}_{10} \in \mathcal{C}^{(28)}$, part (c) of the statement follows.

(d) Let $A = \begin{pmatrix} -1 & 0 \\ -2 & -1 \end{pmatrix}$; since $A \notin \tilde{GL}(2;\mathbb{Z})$, we apply the algorithmic procedure of the previous section to its conjugate matrix $A' = \begin{pmatrix} -1 & -2 \\ 0 & -1 \end{pmatrix} \in \tilde{GL}(2;\mathbb{Z})$ (see Lemma (2.1) in order to prove the conjugation). The coloured pseudocomplex $K_{A'}$ turns out to be obtained by starring from an inner 3-labelled vertex the coloured complex depicted in Figure 7, where 2-dimensional faces labelled x, x' have to be identified.

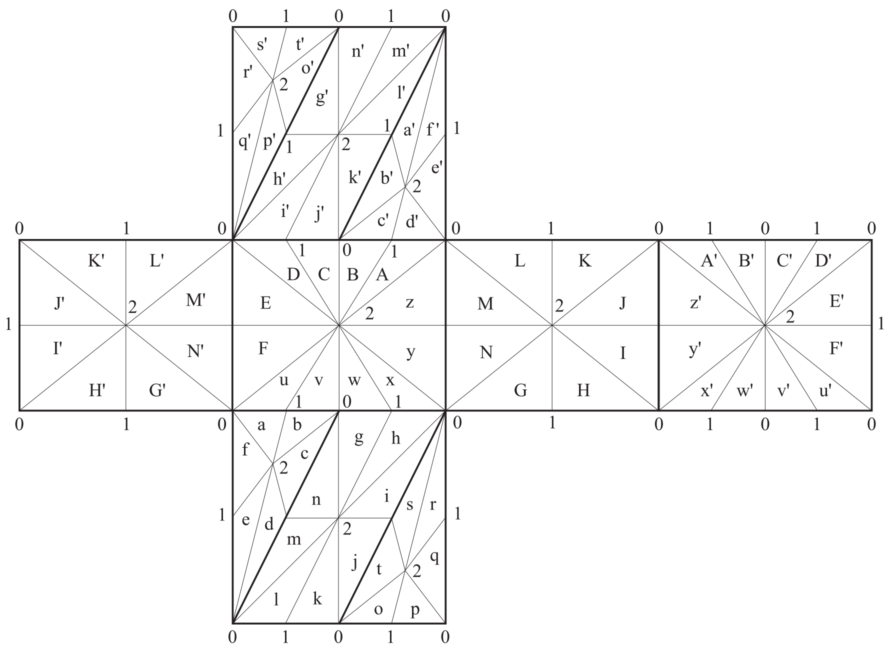


Figure 7

It is not difficult to check that the associated order 80 4-coloured graph $\Gamma(A')$ may be transformed by a finite sequence of dipole moves (for example, by making use of the corresponding function of DUKE III program) into an order 28 4-coloured graph $\Gamma'(A')$ having code

$$c(\Gamma'(A')) = \ dabchefgkijnlmknmedclihgajfbgfjnmkadlcbieh;$$

since this code identifies, up to permutation of vertices and colours, the element $r^{28}_{4540} \in \mathcal{C}^{(28)}$, and since Lins's classification ensures the represented 3-manifold to be the same as $r^{28}_{42} \in \mathcal{C}^{(28)}$, part (d) of the statement follows.

(e) Let $A = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$. As already pointed out in Example (I), the algorithmic procedure of the previous section, applied to its conjugate matrix $A' = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \in \tilde{GL}(2;\mathbb{Z})$, yields the coloured pseudocomplex $K_{A'}$ obtained by starring from an inner 3-labelled vertex the coloured complex depicted in Figure 1(c), where 2-simplices labelled x, x' have to be identified. Now, it is not difficult to check that the associated order 80 4-coloured graph $\Gamma(A')$ may be transformed by a finite sequence of dipole moves (for example, by making use of the corresponding function of DUKE III program) into an order 28 4-coloured graph $\Gamma'(A')$ having code

$$c(\Gamma'(A')) = \; dabchefgkijnlmknmedclihgajfbgfjmnkalecbhdi;$$

since this code identifies, up to permutation of vertices and colours, the element $r^{28}_{4539} \in C^{(28)}$, and since Lins's classification ensures the represented 3-manifold to be the same as $r^{28}_{280} \in \mathcal{C}^{(28)}$, part (e) of the statement follows, too. $\square$

Received May 01, 2003

Final version received April 05, 2004

Dipartimento di Matematica Pura ed Applicata
Università di Modena e Reggio Emilia
Via Campi 213 B
I-41100 Modena
Italy
casali.mariarita@unimore.it

References

[BCG] P. Bandieri - M. R. Casali - C. Gagliardi, *Representing manifolds by crystallization theory: foundations, improvements and related results*, Atti Sem. Mat. Fis. Univ. Modena **Suppl. 49** (2001), 283–337.

[BM] J. Bracho - L. Montejano, *The combinatorics of colored triangulations of manifolds*, Geom. Dedicata **22** (1987), 303–328.

[C] M. R. Casali, *Classification of non-orientable 3-manifolds admitting decompositions into 26 coloured tetrahedra*, Acta Appl. Math. **54** (1999), 75–97.

[CG] M. R. Casali - C. Gagliardi, *Classifying P.L. 5-manifolds up to regular genus seven*, Proc. Amer. Math. Soc. **120** (1) (1994), 275–283.

[CV] A. Costa - P. del Val-Melus, *Crystallizations and permutations representing 3-manifolds*, Boll. Un. Mat. Ital. **1-B** (1987), 477–490.

[FG_1] M. Ferri - C. Gagliardi, *The only genus zero n-manifold is* $\mathbb{S}^n$, Proc. Amer. Math. Soc. **85** (1982), 638–642.

[FG_2] M. Ferri - C. Gagliardi, *Crystallization moves*, Pacific J. Math. **100** (1982), 85–103.

[FGG] M. Ferri - C. Gagliardi - L. Grasselli, *A graph-theoretical representation of PL-manifolds. A survey on crystallizations*, Aequationes Math. **31** (1986), 121–141.

[G] C. Gagliardi, *Extending the concept of genus to dimension n*, Proc. Amer. Math. Soc. **81** (1981), 473–481.

[HW] P. J. Hilton - S. Wylie, An introduction to algebraic topology - Homology theory, Cambridge Univ. Press, 1960.

[IT] Theo Hahn (ed.), International Tables for X-Ray Crystallography, New edition: Vol A, Reidel, 1983.

[L] S. Lins, Gems, computers and attractors for 3-manifolds, Knots and Everything **5**, World Scientific, Singapore, 1995.

[LM] S. Lins - A. Mandel, *Graph-encoded 3-manifolds*, Discrete Math. **57** (1985), 261–284.

[LS] S. Lins - S. Sidki, *The fundamental groups of closed orientable 3-manifolds admitting colored triangualtions with less than 30 tetrahedra*, Int. J. Algebra and Computation **5** (1995), 205–250.

[M] S. Matveev, *Complexity theory of three-dimensional manifolds*, Acta Applicandae Math. **19** (1990), 101–130.

[MP] B. Martelli - C. Petronio, *Three-manifolds having complexity at most 9*, Experimental Mathematics **10 (2)** (2001), 207–236.

[RS] C. Rourke - B. Sanderson, Introduction to Piecewise-linear topology, Springer Verlag, 1972.

[S] S. Steenrod, *The topology of fiber bundles*, Princeton University Press, 1951.

[TO] M. Takahashi - M. Ochiai, *Heegaard diagrams of torus bundles over* $\mathbb{S}^1$, Comm. Math. Univ. Sancti Pauli **31 (1)** (1982), 63–69.

[V] A. Vince, *Combinatorial maps*, J. Combin. Theory Ser. B, **34** (1983), 1–21.

[ZVC] H. Zieschang - E. Vogt - H. D. Coldewey, Surfaces and planar discontinuous group, Lecture Notes in Mathematics **835**, 1980.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

STUDYING THE MULTIVARIABLE ALEXANDER POLYNOMIAL BY MEANS OF SEIFERT SURFACES

DAVID CIMASONI

Abstract. We show how Seifert surfaces, so useful for the understanding of the Alexander polynomial $\Delta_L(t)$, can be generalized in order to study the multivariable Alexander polynomial $\Delta_L(t_1,\dots,t_\mu)$. In particular, we give an elementary and geometric proof of the Torres formula.

1. Introduction

The technique of Seifert surfaces, discovered by Herbert Seifert [12] in 1935, enabled him to make great progress in the study of the Alexander polynomial of a knot. In particular, he succeeded in characterizing among all Laurent polynomials $\Delta(t)$ those that can be realized as the Alexander polynomial of a knot. The introduction by Ralph Fox of the multivariable Alexander polynomial $\Delta_L(t_1,\dots,t_\mu)$ of a μ-component oriented link L naturally gave rise to the corresponding question for this new invariant (see Problem 2 [6]). Guillermo Torres made use of the free differential calculus – developed at that time by Fox – to give several conditions for a polynomial Δ in $\mathbb{Z}[t_1^{\pm1},\dots,t_\mu^{\pm1}]$ to be the Alexander polynomial of a μ-component link [13, 5]. Since then, very little progress has been made on this question: it is known that the Torres conditions are not sufficient in general [7, 11], but a complete algebraic characterization remains out of reach.

In this paper, we present an original approach to this problem. We show how the technique of Seifert surfaces can be generalized to obtain a new geometric interpretation of $\Delta_L(t^{m_1},\dots,t^{m_\mu})$ for any integers $m_1,\dots,m_\mu$ (see Proposition (2.1) and Corollary (3.4)). If an equality holds for $\Delta_L(t^{m_1},\dots,t^{m_\mu})$ for any integers $m_1,\dots,m_\mu$, then it also holds for $\Delta_L(t_1,\dots,t_\mu)$ (Lemma (2.2)); therefore, it is possible to prove properties of Δ_L with this method. As an example, we give an elementary and geometric proof of the celebrated Torres formula, valid for any link in a homology 3-sphere. We also present several properties of Δ_L which turn out to be equivalent to the Torres conditions (Proposition (4.4)).

2. Preliminaries

Let us consider an oriented ordered link $L = L_1 \cup \dots \cup L_\mu$ in a homology 3-sphere Σ, and let X be the exterior of L. If $\widehat{X} \xrightarrow{\hat{p}} X$ denotes the universal abelian covering of X and $\widehat{X}^0$ the inverse image by $\hat{p}$ of a base point X^0 of X, the homology $\mathrm{H}_1(\widehat{X},\widehat{X}^0)$ is endowed with a natural structure of a module over

2000 *Mathematics Subject Classification*: 57M25.

Keywords and phrases: Alexander polynomial of a link, Seifert surface, Torres conditions.

the ring $\Lambda_\mu = \mathbb{Z}[t_1^{\pm 1}, \dots, t_\mu^{\pm 1}]$. Given an $m \times n$ presentation matrix of $\mathrm{H}_1(\widehat{X}, \widehat{X}^0)$ – that is, the matrix $\mathcal{P}$ corresponding to a presentation with n generators and m relations – the $(n-i)\times(n-i)$ minor determinants of $\mathcal{P}$ span an ideal of Λ_μ denoted by $\mathcal{E}_i\mathrm{H}_1(\widehat{X}, \widehat{X}^0)$. The greatest common divisor of these minor determinants is denoted by $\Delta_i\mathrm{H}_1(\widehat{X}, \widehat{X}^0)$; this invariant is well defined up to multiplication by units of Λ_μ, that is, by $\pm t_1^{\nu_1} \cdots t_\mu^{\nu_\mu}$ with $\nu_i \in \mathbb{Z}$. In the sequel, we will write $\Delta \doteq \Delta'$ if two elements Δ, Δ' of a ring R satisfy $\Delta = \varepsilon\Delta'$ for some unit ε of R. The Laurent polynomial $\Delta_1\mathrm{H}_1(\widehat{X}, \widehat{X}^0)$ is called the *Alexander polynomial* of the link L [1, 4]. It is denoted by $\Delta_L(t_1, \dots, t_\mu)$.

Our method will be to prove statements on this polynomial in an indirect way, by studying all the infinite cyclic coverings of X. Since these coverings are classified by $\mathrm{Hom}(\mathrm{H}_1(X), \mathbb{Z}) \simeq \mathrm{H}^1(X;\mathbb{Z}) \simeq \mathrm{H}_1(L) = \bigoplus_{i=1}^{\mu} \mathbb{Z}L_i$, this leads to the following definition [3]. A *multilink* is an oriented link $L = L_1 \cup \cdots \cup L_\mu$ in a homology sphere Σ together with an integer m_i associated with each component L_i, with the convention that a component L_i with multiplicity m_i is the same as $-L_i$ (L_i with reversed orientation) with multiplicity $-m_i$. Throughout this paper, we will write $\underline{m}$ for the ordered set of integers $m_1, \dots, m_\mu$, d for their greatest common divisor, and $L(\underline{m})$ for the multilink. Finally, we will also denote by $\underline{m}$ the morphism $\mathrm{H}_1(X) \to \mathbb{Z}$ given by $\underline{m}(\gamma) = \sum_{i=1}^{\mu} m_i \ell k(L_i, \gamma)$. Let $\widetilde{X} \xrightarrow{\tilde{p}} X$ be the regular $\mathbb{Z}$-covering determined by $\underline{m}$. If $\widetilde{X}^0 = \tilde{p}^{-1}(X^0)$, the homology $\mathrm{H}_1(\widetilde{X}, \widetilde{X}^0)$ can be thought of as a module over the ring $\mathbb{Z}[t^{\pm 1}]$. The Laurent polynomial $\Delta_{L(\underline{m})}(t) = \Delta_1\mathrm{H}_1(\widetilde{X}, \widetilde{X}^0)$ is called the *Alexander polynomial* of the multilink $L(\underline{m})$. Note that if $\underline{m} \neq \underline{0}$, the exact sequence of the pair $(\widetilde{X}, \widetilde{X}^0)$ implies at once that $\mathcal{E}_1\mathrm{H}_1(\widetilde{X}, \widetilde{X}^0) = \mathcal{E}_0\mathrm{H}_1(\widetilde{X})$. Therefore, $\Delta_{L(\underline{m})}(t)$ is also equal to $\Delta_0\mathrm{H}_1(\widetilde{X})$.

Here is the dictionary between the polynomials Δ_L and $\Delta_{L(\underline{m})}$:

PROPOSITION (2.1) (Eisenbud-Neumann [3]).

$$\Delta_{L(\underline{m})}(t) \doteq \begin{cases} \Delta_L(t^{m_1}) & \text{if } \mu = 1; \\ (t^d - 1)\,\Delta_L(t^{m_1}, \dots, t^{m_\mu}) & \text{if } \mu \geq 2. \end{cases}$$

Proof. To check this equality, we need the well-known fact that $\mathcal{E}_1\mathrm{H}_1(\widehat{X}, \widehat{X}^0) = (\Delta_*) \cdot I$, where I is the augmentation ideal $(t_1 - 1, \dots, t_\mu - 1)$ and Δ_* some polynomial in Λ_μ. This can be proved by purely homological algebraic methods using the fact that the group $\pi_1(X)$ has defect ≥ 1 (see Theorem 6.1 [3]). By considering a finite presentation of $\mathrm{H}_1(\widehat{X}, \widehat{X}^0)$ given by an equivariant cellular decomposition of $\widehat{X}$, it is easy to show that $\mathrm{H}_1(\widehat{X}, \widehat{X}^0) \otimes_{\Lambda_\mu} \mathbb{Z}[t^{\pm 1}] = \mathrm{H}_1(\widetilde{X}, \widetilde{X}^0)$, where $\mathbb{Z}[t^{\pm 1}]$ is endowed with the structure of Λ_μ-algebra given by $t_i \mapsto t^{m_i}$ for $i = 1, \dots, \mu$. Hence,

$$\begin{aligned} \mathcal{E}_1\mathrm{H}_1(\widetilde{X}, \widetilde{X}^0) &= \mathcal{E}_1(\mathrm{H}_1(\widehat{X}, \widehat{X}^0) \otimes_{\Lambda_\mu} \mathbb{Z}[t^{\pm 1}]) \\ &= (\Delta_*(t^{m_1}, \dots, t^{m_\mu})) \cdot (t^{m_1} - 1, \dots, t^{m_\mu} - 1) \\ &= (\Delta_*(t^{m_1}, \dots, t^{m_\mu})) \cdot (t^d - 1). \end{aligned}$$

Since $\Delta_L = (t_1 - 1)\,\Delta_*$ if $\mu = 1$ and $\Delta_L = \Delta_*$ if $\mu \geq 2$, the proposition is proved. □

In order to show that properties of $\Delta_{L(\underline{m})}$ translate directly into properties of Δ_L, we also need the following lemma.

LEMMA (2.2). *Consider two polynomials Δ and Δ' in Λ_μ such that*

$$\Delta(t^{m_1},\dots,t^{m_\mu}) \doteq \Delta'(t^{m_1},\dots,t^{m_\mu}) \quad \text{in} \quad \mathbb{Z}[t^{\pm 1}]$$

for all $(m_1,\dots,m_\mu)$ in $\mathbb{Z}^\mu$ except possibly a finite number of them. Then, $\Delta \doteq \Delta'$ in Λ_μ.

Proof. Without loss of generality, it may be assumed that $\Delta = \sum a_{i_1\cdots i_\mu} t_1^{i_1}\cdots t_\mu^{i_\mu}$ and $\Delta' = \sum b_{j_1\cdots j_\mu} t_1^{j_1}\cdots t_\mu^{j_\mu}$ with $a_{0\dots 0} > 0$, $b_{0\dots 0} > 0$, and only non-negative indices $i_k, j_k \geq 0$. By hypothesis, there are maps $\mathbb{Z}^\mu \xrightarrow{\varepsilon} \{\pm 1\}$ and $\mathbb{Z}^\mu \xrightarrow{\nu} \mathbb{Z}$ such that the equality

$$\sum a_{i_1\cdots i_\mu} t^{\sum_k m_k i_k} \;=\; \varepsilon(m_1,\dots,m_\mu)\, t^{\nu(m_1,\dots,m_\mu)} \sum b_{j_1\cdots j_n} t^{\sum_k m_k j_k}$$

holds for all but a finite number of $(m_1,\dots,m_\mu)$ in $\mathbb{Z}^\mu$. Let us choose an integer N greater than $\max_k \deg_{t_k}\Delta$ and $\max_k \deg_{t_k}\Delta'$, and set $m_1 = 1$, $m_2 = N$, $\dots, m_\mu = N^{\mu-1}$. By choosing N sufficiently large, it may be assumed that the equality above holds for this ordered set of integers. Since all these integers are positive as well as the coefficients $a_{0\dots 0}$ and $b_{0\dots 0}$, it follows that $\varepsilon(1,N,\dots,N^{\mu-1}) = +1$ and $\nu(1,N,\dots,N^{\mu-1}) = 0$. This gives

$$\sum a_{i_1\cdots i_\mu} t^{\,i_1+Ni_2+\cdots+N^{\mu-1}i_\mu} = \sum b_{j_1\cdots j_\mu} t^{\,j_1+Nj_2+\cdots+N^{\mu-1}j_\mu}.$$

But the equality $i_1 + Ni_2 + \cdots + N^{\mu-1}i_\mu = j_1 + Nj_2 + \cdots + N^{\mu-1}j_\mu$ with $0 \leq i_k, j_k < N$ for all k implies that $(i_1,\dots,i_\mu) = (j_1,\dots,j_\mu)$. Hence, $a_{i_1,\dots,i_\mu} = b_{i_1,\dots,i_\mu}$ for all multi-indices $(i_1,\dots,i_\mu)$, which proves the result. □

3. Generalized Seifert surfaces

One of the advantages of multilinks is that they can be studied via generalized Seifert surfaces [3]. A *Seifert surface* for a multilink $L(\underline{m})$ is an open embedded oriented surface $F \subset \Sigma\setminus L$ such that, if F_0 denotes $F \cap (\Sigma \setminus \operatorname{int}\mathcal{N}(L))$, the closure $c\ell(F)$ of F intersects a closed tubular neighborhood $\mathcal{N}(L_i)$ of L_i as follows for each i:

– If $m_i \neq 0$, $c\ell(F)\cap\mathcal{N}(L_i)$ consists of $|m_i|$ sheets meeting along L_i; F is oriented such that $\partial F_0 = m_i L_i$ in $\mathrm{H}_1(\mathcal{N}(L_i))$.

– If $m_i = 0$, $c\ell(F)\cap\mathcal{N}(L_i)$ consists of discs transverse to L_i; F is oriented such that the intersection number of L_i with each of these discs is the same (either always $+1$ or always -1).

This is illustrated in Figure (1). Note that $F \subset \Sigma\setminus L$ and $F_0 \subset \Sigma\setminus\operatorname{int}\mathcal{N}(L)$ determine each other up to isotopy; to simplify the notation, we will consider both of them as Seifert surfaces, and denote both by F. From now on, we will write $\overline{F}$ for the union of $F \subset \Sigma\setminus L$ and L.

LEMMA (3.1) (Eisenbud-Neumann [3]). *Let F be a Seifert surface for a multilink $L(\underline{m})$. Then, for $i = 1\dots,\mu$, the intersection $F\cap\partial\mathcal{N}(L_i)$ gives a d_i component link which is the (d_ip_i, d_iq_i)-cable about L_i, where p_i and q_i are coprime, $d_ip_i = m_i$ and $d_iq_i = -\sum_{j\neq i} m_j \ell k(L_i,L_j)$.*

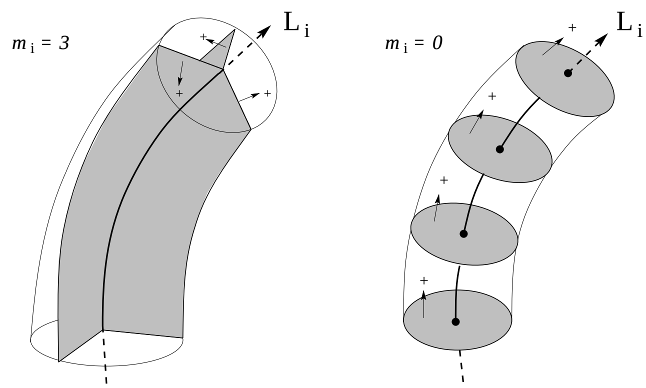


Figure 1. A Seifert surface near the multilink.

Proof. Let us denote by (P_i, M_i) a basis of $\mathrm{H}_1(\partial\mathcal{N}(L_i))$ given by a standard parallel and meridian. Since F is a Seifert surface for $L(\underline{m})$, $F \cap \partial\mathcal{N}(L_i) = m_i P_i + n_i M_i$ in $\mathrm{H}_1(\partial\mathcal{N}(L_i))$ for some integer n_i. Furthermore, $\partial F = \sum_{j\neq i} m_j L_j + m_i P_i + n_i M_i$ in $\mathrm{H}_1(\Sigma \setminus \operatorname{int}\mathcal{N}(L_i))$. By Alexander duality, this module is isomorphic to $\mathrm{H}^1(\mathcal{N}(L_i);\mathbb{Z}) = \mathbb{Z}$, and the isomorphism is given by the linking number with L_i. It follows that $0 = \ell k(L_i, \partial F) = \sum_{j\neq i} m_j \ell k(L_i, L_j) + n_i$, which gives the result. □

In the usual case of an oriented link, a Seifert surface needs to be connected in order to be useful. In the general case of a multilink, it has to be "as connected as possible". More precisely, a Seifert surface for $L(\underline{m})$ is a *good Seifert surface* if it has $d = \gcd(\underline{m})$ connected components.

Lemma (3.2). *Given a multilink $L(\underline{m})$, there exists a good Seifert surface for $L(\underline{m})$.*

Proof. One easily shows that there exists a Seifert surface for $L(\underline{m})$ (see Lemma 3.1 [3]). If $d > 1$, a good Seifert surface for $L(\underline{m})$ is given by d parallel copies of a connected Seifert surface for $L(\frac{\underline{m}}{d})$. Therefore, it may be assumed that $d = 1$. Let F be any Seifert surface for $L(\underline{m})$ without closed component, and let us denote by i_+ (resp. i_-) the epimorphism $\mathrm{H}_0(F) \to \mathrm{H}_0(\Sigma \setminus \overline{F})$ induced by pushing in the positive (resp. negative) normal direction off F. If i_+ and i_- are not isomorphisms, it is possible to reduce the number of connected components of F by handle attachment. So, let us assume that all the possible handle attachment(s) have been performed, yielding $F = F_1 \cup \cdots \cup F_n$ with isomorphisms $i_+, i_- \colon \mathrm{H}_0(F) \to \mathrm{H}_0(\Sigma \setminus \overline{F})$. The automorphism of $\mathrm{H}_0(F)$ given by $h = (i_-)^{-1} \circ i_+$ cyclically permutes the connected components of F. (Indeed, consider a component F_i of F; since $X = (\Sigma \setminus \overline{F}) \cup F$ is path connected and i_+, i_- are isomorphisms, there exists an integer m such that $F_i = h^m(F_1)$.) It easily follows that $\partial F_i = \partial F_j$ in $\mathrm{H}_1(\mathcal{N}(L))$ for $i, j = 1, \ldots, n$. Therefore, the equality $\sum_{i=1}^{\mu} m_i L_i = \partial F = \sum_{j=1}^{n} \partial F_j = n\,\partial F_1$ holds in $\mathrm{H}_1(\mathcal{N}(L)) = \bigoplus_{i=1}^{\mu} \mathbb{Z} L_i$. Hence, n divides m_i for $i = 1, \ldots, \mu$. Since $\gcd(m_1, \ldots, m_\mu) = 1$, F is connected. □

Let us now turn to the natural generalization to multilinks of the Seifert form. Given F a good Seifert surface for $L(\underline{m})$, the *Seifert forms* associated to F are the bilinear forms

$$\alpha_+, \alpha_- \colon \mathrm{H}_1(F) \times \mathrm{H}_1(\overline{F}) \longrightarrow \mathbb{Z}$$

given by $\alpha_+(x,y) = \ell k(i_+ x, y)$ and $\alpha_-(x,y) = \ell k(i_- x, y)$, where i_+ (resp. i_-) is the morphism $\mathrm{H}_1(F) \to \mathrm{H}_1(\Sigma \setminus \overline{F})$ induced by pushing in the positive (resp. negative) normal direction off F. (Note that we use the same notation for the morphisms $\mathrm{H}_0(i_\pm)$ and $\mathrm{H}_1(i_\pm)$; it will always be clear from the context which dimension is concerned.) Let us denote by A_+ and A_- the matrices associated with these forms, called *Seifert matrices*. Here is the generalization of Seifert's famous theorem.

Theorem (3.3). *Let F be a good Seifert surface for $L(\underline{m})$, and let A_+, A_- be associated Seifert matrices. Then, $A_+ - tA_-$ is a presentation matrix of the module $H_1(\widetilde{X})$.*

Proof. Given F a good Seifert surface for $L(\underline{m})$, let us denote $\Sigma \setminus \overline{F}$ by Y. By the proof of Lemma (3.2) it is possible to number the connected components $F = F_1 \cup \cdots \cup F_d$ and $Y = Y_1 \cup \cdots \cup Y_d$ so that $i_+ F_k = Y_k$ and $i_- F_k = Y_{k-1}$ (with the indices modulo d). Let us set $N = F \times (-1;1)$ an open bicollar of F, $N_+ = F \times (0;1)$, $N_- = F \times (-1;0)$ and $\{Y^i\}_{i\in\mathbb{Z}}$ (resp. $\{N^i\}_{i\in\mathbb{Z}}$) copies of Y (resp. N). Define

$$E = \bigsqcup_{i\in\mathbb{Z}} Y^i \sqcup \bigsqcup_{i\in\mathbb{Z}} N^i \Big/ \sim\,,$$

where $Y^i \supset N_+ \sim N_+ \subset N^i$ and $Y^i \supset N_- \sim N_- \subset N^{i+1}$. The obvious projection $E \xrightarrow{p} X$ is the infinite cyclic covering $\widetilde{X} \to X$ determined by $\underline{m}$. Indeed, a loop γ in X lifts to a loop in E if and only if the intersection number of γ with F is zero, that is, if $0 = \gamma \cdot F = \ell k(L(\underline{m}), \gamma) = \underline{m}(\gamma)$.

Consider the Mayer-Vietoris exact sequence of $\mathbb{Z}[t^{\pm1}]$-modules associated to the decomposition $\widetilde{X} = \left(\bigcup_i Y^i\right) \cup \left(\bigcup_i N^i\right)$; it gives

$$\begin{aligned}
(\mathrm{H}_1(F) \oplus \mathrm{H}_1(F)) \otimes \mathbb{Z}[t^{\pm1}] \quad &\xrightarrow{\phi_1} \quad (\mathrm{H}_1(Y) \oplus \mathrm{H}_1(F)) \otimes \mathbb{Z}[t^{\pm1}] \xrightarrow{\ \psi\ } \mathrm{H}_1(\widetilde{X}) \to \\
(\mathrm{H}_0(F) \oplus \mathrm{H}_0(F)) \otimes \mathbb{Z}[t^{\pm1}] \quad &\xrightarrow{\phi_0} \quad (\mathrm{H}_0(Y) \oplus \mathrm{H}_0(F)) \otimes \mathbb{Z}[t^{\pm1}],
\end{aligned}$$

where the homomorphism ϕ_0 is given by $(\alpha,\beta) \mapsto (i_+\alpha + t\, i_-\beta, \alpha + \beta)$. Since F is good, the homomorphisms $i_\pm \colon \mathrm{H}_0(F) \to \mathrm{H}_0(\Sigma \setminus \overline{F})$ are injective, and so is ϕ_0. Therefore, ψ is surjective and there is an exact sequence

$$(\mathrm{H}_1(F) \oplus \mathrm{H}_1(F)) \otimes \mathbb{Z}[t^{\pm1}] \xrightarrow{\phi_1} (\mathrm{H}_1(Y) \oplus \mathrm{H}_1(F)) \otimes \mathbb{Z}[t^{\pm1}] \to \mathrm{H}_1(\widetilde{X}) \to 0,$$

with $\phi_1(\alpha,\beta) = (i_+\alpha + t\, i_-\beta, \alpha+\beta)$. This can be transformed into

$$\mathrm{H}_1(F) \otimes \mathbb{Z}[t^{\pm1}] \xrightarrow{\tilde{\phi}} \mathrm{H}_1(Y) \otimes \mathbb{Z}[t^{\pm1}] \to \mathrm{H}_1(\widetilde{X}) \to 0,$$

where $\tilde{\phi}(\alpha) = i_+\alpha - t\, i_-\alpha$. Let us fix bases $\mathcal{B}$ for $\mathrm{H}_1(F)$, $\overline{\mathcal{B}}$ for $\mathrm{H}_1(\overline{F})$, and consider the basis $\overline{\mathcal{B}}^*$ for $\mathrm{H}_1(Y)$ which is dual to $\overline{\mathcal{B}}$ under Alexander duality. The matrix of i_+ (resp. i_-) with respect to $\mathcal{B}$ and $\overline{\mathcal{B}}^*$ is given by A_+^T (resp. A_-^T), where A_+ and A_- are the Seifert matrices with respect to the bases $\mathcal{B}$ and $\overline{\mathcal{B}}$. Therefore, a matrix of $\tilde{\phi}$ is given by $A_+^T - tA_-^T$. This concludes the proof. $\square$

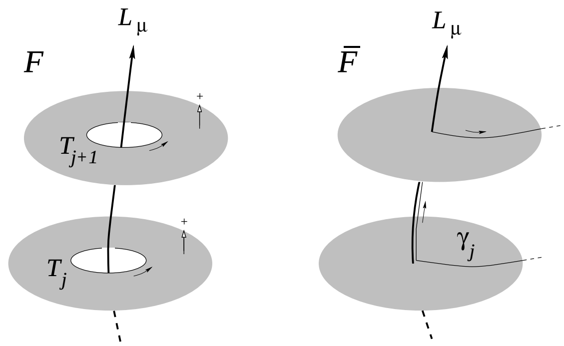


Figure 2. The proof of Lemma (4.1).

Corollary (3.4). *Let $L(\underline{m})$ be a multilink with $\underline{m} \neq \underline{0}$. If*

$$m_i = \sum_{j \neq i} m_j \ell k(L_i, L_j) = 0$$

for some index i, then $\Delta_{L(\underline{m})}(t) = 0$. If there is no such index then the matrices A_+ and A_- are square, and $\Delta_{L(\underline{m})}(t) \doteq \det(A_+ - tA_-)$.

Proof. By the proof of Lemma (3.2), a Seifert surface F is good if and only if $\operatorname{rk} \widetilde{\mathrm{H}}_0(F) = \operatorname{rk} \widetilde{\mathrm{H}}_0(\Sigma \setminus \overline{F})$, which is equal to $\operatorname{rk} \mathrm{H}_2(\overline{F})$ by Alexander duality. It is easy to show that $\operatorname{rk} \widetilde{\mathrm{H}}_0(\overline{F}) = r$, the number of indices i with $m_i = \sum_{j \neq i} m_j \ell k(L_i, L_j) = 0$. Since $\chi(F) = \chi(\overline{F})$, it follows that $\operatorname{rk} \mathrm{H}_1(\overline{F}) = \operatorname{rk} \mathrm{H}_1(F) + r$. So, if $r = 0$ then $A_+ - tA_-$ is a square presentation matrix of $\mathrm{H}_1(\widetilde{X})$, and if $r > 0$ then there are more generators than relations. It follows that $\Delta_0 \mathrm{H}_1(\widetilde{X}) \doteq \det(A_+ - tA_-)$ if $r = 0$, and $\Delta_0 \mathrm{H}_1(\widetilde{X}) = 0$ if $r > 0$. □

4. The Torres conditions

Let us now illustrate how Corollary (3.4), along with Proposition (2.1) and Lemma (2.2), can be used to study the multivariable Alexander polynomial. As an example, we present an elementary proof of the Torres formula [13], quite simpler than the original proof. (On the other hand, it should be mentioned that more perspicuous proofs have since been given, for example in [9]).

Throughout this section, we will denote by ℓ_{ij} the linking number $\ell k(L_i, L_j)$.

Lemma (4.1). *Let $L(\underline{m}) = L(m_1, \dots, m_{\mu-1}, 0)$ be a multilink, and let $L'(\underline{m}') = L'(m_1, \dots, m_{\mu-1})$ be the multilink obtained from $L(\underline{m})$ by removing the last component L_μ. Then,*

$$\Delta_{L(\underline{m})}(t) \doteq (t^{\sum_i m_i \ell_{i\mu}} - 1)\, \Delta_{L'(\underline{m}')}(t).$$

Proof. If $m_i = \sum_{j \neq i} m_j \ell_{ij} = 0$ for some index i then the lemma holds by Corollary (3.4). It may therefore be assumed that there is no such index. Let F be a good Seifert surface for $L(\underline{m})$; then, a good Seifert surface for $L'(\underline{m}')$ is given by $F' = F \cup (\overline{F} \cap \mathcal{N}(L_\mu))$. By Lemma (3.1), $\overline{F} \cap \mathcal{N}(L_\mu)$ consists of

$d_\mu = \sum_{i=1}^{\mu-1} m_i\ell_{i\mu}$ discs (recall Figure (1)). Furthermore, $\overline{F} = \overline{F}' \cup L_\mu$. Therefore, we have the natural isomorphisms

$$\mathrm{H}_1(F) = \mathrm{H}_1(F') \oplus \bigoplus_{j=1}^{d_\mu} \mathbb{Z}T_j \quad \text{and} \quad \mathrm{H}_1(\overline{F}) = \mathrm{H}_1(\overline{F}') \oplus \bigoplus_{j=1}^{d_\mu} \mathbb{Z}\gamma_j,$$

where the cycles T_j correspond to the boundaries of the discs, and the γ_j are the transverse cycles depicted in Figure (2). The associated Seifert matrices $A_\pm$ and $A'_\pm$ are related by

$$A_+ = \begin{pmatrix} A'_+ & & 0 & \\ & 1 & & \\ * & & \ddots & \\ & & & 1 \end{pmatrix} \quad \text{and} \quad A_- = \begin{pmatrix} A'_- & & & 0 & \\ & 0 & & & 1 \\ & 1 & 0 & & \\ * & & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix}.$$

Corollary (3.4) then gives

$$\begin{aligned} \Delta_{L(\underline{m})} &\doteq \det(A_+ - tA_-) = \begin{vmatrix} A'_+ - tA'_- & & & 0 & \\ & 1 & & & -t \\ & -t & 1 & & \\ * & & \ddots & \ddots & \\ & & & -t & 1 \end{vmatrix} \\ &= (t^{d_\mu} - 1)\det(A'_+ - tA'_-) \doteq (t^{d_\mu} - 1)\,\Delta_{L'(\underline{m}')}(t) \end{aligned}$$

and the lemma is proved. □

The demonstration of the Torres formula is now a mere translation of Lemma (4.1) via Proposition (2.1).

Torres formula (4.2). [13] *Let $L = L_1 \cup \cdots \cup L_\mu$ be an oriented link with $\mu \geq 2$ components, and let L' be the sublink $L_1 \cup \cdots \cup L_{\mu-1}$. Then,*

$$\Delta_L(t_1, \dots, t_{\mu-1}, 1) \doteq \begin{cases} \frac{t_1^{\ell_{12}} - 1}{t_1 - 1}\,\Delta_{L'}(t_1) & \text{if } \mu = 2; \\ (t_1^{\ell_{1\mu}} \cdots t_{\mu-1}^{\ell_{\mu-1,\mu}} - 1)\,\Delta_{L'}(t_1, \dots, t_{\mu-1}) & \text{if } \mu > 2. \end{cases}$$

Proof. Let us denote by Δ' the right-hand side of this formula, and let $m_1, \dots, m_{\mu-1}$ be arbitrary integers with $d = \gcd(m_1, \dots, m_{\mu-1}) > 0$. We have the equalities

$$\begin{aligned} \Delta'(t^{m_1}, \dots, t^{m_{\mu-1}}) &= \begin{cases} \frac{t^{m_1\ell_{12}} - 1}{t^{m_1} - 1}\,\Delta_{L'}(t^{m_1}) & \text{if } \mu = 2; \\ (t^{\sum_i m_i\ell_{i\mu}} - 1)\,\Delta_{L'}(t^{m_1}, \dots, t^{m_{\mu-1}}) & \text{if } \mu > 2, \end{cases} \\ \text{(Proposition (2.1))} &\doteq \frac{1}{t^d - 1}\,(t^{\sum_i m_i\ell_{i\mu}} - 1)\,\Delta_{L'(\underline{m}')}(t) \\ \text{(Lemma (4.1))} &\doteq \frac{1}{t^d - 1}\,\Delta_{L(\underline{m})}(t) \\ \text{(Proposition (2.1))} &\doteq \Delta_L(t^{m_1}, \dots, t^{m_{\mu-1}}, 1) \end{aligned}$$

and the proof is settled by Lemma (2.2). □

Using the same method, it is not hard to show the following result.

Fox-Torres relation (4.3). [13, 5] *Let $L = L_1 \cup \cdots \cup L_\mu$ be an oriented link with $\mu \geq 2$ components. Then,*

$$\Delta_L(t_1^{-1}, \ldots, t_\mu^{-1}) = (-1)^\mu \, t_1^{\nu_1 - 1} \cdots t_n^{\nu_\mu - 1} \, \Delta_L(t_1, \ldots, t_\mu)$$

with integers ν_i such that $\nu_i \equiv \sum_j \ell_{ij} (\text{mod } 2)$ if $\Delta_L \neq 0$.

These results provide necessary conditions for a polynomial Δ in Λ_μ to be the Alexander polynomial of a μ-component link with fixed $\ell k(L_i, L_j) = \ell_{ij}$. They are known as the *Torres conditions* (see [10] for a precise statement). Since these conditions are not sufficient [7], [11], the problem is now to find stronger conditions. By means of a close study of the homology $\mathrm{H}_1(F)$ and $\mathrm{H}_1(\overline{F})$, it is possible to find necessary conditions for a polynomial Δ in $\mathbb{Z}[t^{\pm 1}]$ to be the Alexander polynomial of a multilink. Via Proposition (2.1), this translates into the following result (see [2] for a proof).

Proposition (4.4). *Let L be an oriented link with $\mu \geq 2$ components. Then, its Alexander polynomial Δ_L satisfies the following conditions. For all integers $\underline{m} = (m_1, \ldots, m_\mu)$ with $d = gcd(m_1, \ldots, m_\mu)$ and $d_i = gcd(m_i, \sum_j m_j \ell_{ij})$, there exists some polynomial $\nabla_{L(\underline{m})}(t)$ in $\mathbb{Z}[t^{\pm d}]$ such that:*

– $\prod_{i=1}^{\mu} (t^{d_i} - 1) \, \nabla_{L(\underline{m})}(t) \doteq (t^d - 1)^2 \, \Delta_L(t^{m_1}, \ldots, t^{m_\mu})$*;*

– $\nabla_{L(\underline{m})}(t^{-1}) = \nabla_{L(\underline{m})}(t)$*;*

– $|\nabla_{L(\underline{m})}(1)| = \frac{d^2 \, D}{d_1 \cdots d_\mu m_1 \cdots m_\mu}$*, where D is any $(\mu - 1) \times (\mu - 1)$ minor determinant of the matrix*

$$\begin{pmatrix} -\sum_j m_1 m_j \ell_{1j} & m_1 m_2 \ell_{12} & \ldots & m_1 m_\mu \ell_{1\mu} \\ m_1 m_2 \ell_{12} & -\sum_j m_2 m_j \ell_{2j} & \ldots & m_2 m_\mu \ell_{2\mu} \\ \vdots & \vdots & \ddots & \vdots \\ m_1 m_\mu \ell_{1\mu} & m_2 m_\mu \ell_{2\mu} & \ldots & -\sum_j m_\mu m_j \ell_{\mu j} \end{pmatrix};$$

– *If $m_i = 0$ for some index i, then $\nabla_{L(\underline{m})} = \nabla_{L'(\underline{m}')}$, where L' denotes the sublink $L \setminus L_i$ and $\underline{m}' = (m_1, \ldots, \widehat{m_i}, \ldots, m_\mu)$.* □

This result easily implies the Torres conditions. It can also be thought of as a generalization of a theorem of Hosokawa [8], which corresponds to the case $m_1 = \cdots = m_\mu = 1$. At first sight, it might therefore seem more general than the Torres conditions. Unfortunately, this is not the case: it can be shown that every polynomial Δ which satisfies the Torres conditions also satisfies the conditions of Proposition (4.4) (see [2]).

By means of a somewhat closer study of the Seifert matrices $A_\pm$, it should be possible to find new properties of $\Delta_{L(\underline{m})}$. They would translate into properties of Δ_L, and provide new conditions, stronger than the ones of Torres.

Acknowledgments

I wish to express my thanks to Jerry Levine, to my advisor Claude Weber, and to Mathieu Baillif.

Received June 20, 2003

Final version received December 03, 2003

Section de Mathématiques
Université de Genève
2–4 rue du Lièvre
1211 Genève 24
Switzerland
David.Cimasoni@math.unige.ch

References

[1] J. W. Alexander, *Topological invariants of knots and links*, Trans. Amer. Math. Soc. **30** (1928), 275–306.
[2] D. Cimasoni, *Alexander invariants of multilinks*, PhD thesis, University of Geneva, 2002.
[3] D. Eisenbud and W. Neumann, Three-dimensional link theory and invariants of plane curve singularities, Ann. of Math. Stud. **110**, Princeton Univ. Press, Princeton, NJ, 1985.
[4] R. H. Fox, *Free differential calculus. II. The isomorphism problem of groups*, Ann. of Math. **59** (1954), 196–210.
[5] R. H. Fox and G. Torres, *Dual presentations of the group of a knot*, Ann. of Math. **59** (1954), 211–218.
[6] R. H. Fox, *Some problems in knot theory*, In: Topology of 3-manifolds and related topics, Proc. 1961 Top. Inst. Georgia, Prentice-Hall, Englewood Cliffs, 1962, pp. 168–176.
[7] J. Hillman, *The Torres conditions are insufficient*, Math. Proc. Cambridge Philos. Soc. **89** (1981), 19–22.
[8] F. Hosokawa, *On ∇-polynomial of links*, Osaka Math. J. **10** (1958), 273–282.
[9] A. Kawauchi, A survey of knot theory, Birkhäuser Verlag, Basel, 1996.
[10] J. Levine, *A method for generating link polynomials*, Amer. J. Math. **89** (1967), 69–84.
[11] M. L. Platt, *Insufficiency of Torres' conditions for two-component classical links*, Trans. Amer. Math. Soc. **296** (1986), 125–136.
[12] H. Seifert, *Über das Geschlecht von Knoten*, Mathematische Annalen **110** (1935), 571–592.
[13] G. Torres, *On the Alexander polynomial*, Ann. Math. **57** (1953), 57–89.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

A NOTE ON TORSION IN K_3 OF THE REAL NUMBERS

JOSÉ LUIS CISNEROS-MOLINA

Abstract. Following [9], we prove that every torsion element in $K_3(\mathbb{R})$ and in its indecomposable part $K_3^{\mathrm{ind}}(\mathbb{R})$ can be constructed using Brieskorn homology 3-spheres endowed with a representation of its fundamental group in $SL_4(\mathbb{R})$. Also, using the generators of the Bloch group $B(\mathbb{R})$ constructed via the dilogarithm identities in [7], we give an explicit map $K_3^{\mathrm{ind}}(\mathbb{R})_{\mathrm{tor}} \to B(\mathbb{R})_{\mathrm{tor}}$.

1. Introduction

The regulator map is a homomorphism $e\colon K_{2n+1}(\mathbb{C}) \to \mathbb{C}/\mathbb{Z}$ defined independently by Beilinson [3] and Karoubi [10] as a secondary Chern character. In [9] J. D. S. Jones and B. W. Westbury constructed elements in the algebraic K-group $K_n(R)$ using homology n-spheres endowed with a representation of their fundamental group in the general linear group over the ring R. They also computed the image of these elements in $K_3(\mathbb{C})$ under the regulator map. Using these computations, they proved that every torsion element in $K_3(\mathbb{C})$ can be constructed using Brieskorn homology 3-spheres. Finally, combining these computations with results of Merkurjev and Suslin [17] and Levine [15], they gave an explicit generator of the torsion subgroup of K_3 of the ring of algebraic integers in a non-trivial cyclotomic extension of the rationals of degree coprime to 6. The first aim of the present paper is to give analogous results for $K_3(\mathbb{R})$ and its indecomposable part $K_3^{\mathrm{ind}}(\mathbb{R})$, that is, every torsion element in $K_3(\mathbb{R})$ and $K_3^{\mathrm{ind}}(\mathbb{R})$ can be constructed using Brieskorn homology 3-spheres. Following [9], it would be good to find explicit generators in terms of Brieskorn homology 3-spheres of the torsion subgroup of K_3 of the real part of a non-trivial cyclotomic extension of the rationals. So far we have not been able to find such generators, but we mention a possible way to do it.

The Bloch group $B(F)$ of a field F is a group closely related with $K_3^{\mathrm{ind}}(F)$ by an exact sequence due to Suslin [21]. In [7] Frenkel and Szenes, using Roger's dilogarithm, constructed a map $\mathcal{L}\colon B(\mathbb{R}) \to \mathbb{R}/(\pi^2\mathbb{Z})$. Using dilogarithm identities they constructed generators of the torsion subgroup of the Bloch group of totally real fields. Using these generators, the representation of torsion elements in $K_3^{\mathrm{ind}}(\mathbb{R})$ by Brieskorn homology 3-spheres, and combining the computations of [9] and [7], we give an explicit homomorphism

$$K_3^{\mathrm{ind}}(\mathbb{R})_{\mathrm{tor}} \to B(\mathbb{R})_{\mathrm{tor}}$$

2000 *Mathematics Subject Classification*: 19D99, 11R70.

Keywords and phrases: K-theory of real numbers, homology spheres, bloch group, dilogarithm identities.

which, composed with the homomorphism $\mathcal{L}$, is related with the map $i_*\colon K_3(\mathbb{R}) \to K_3(\mathbb{C})$ induced by the inclusion composed with the regulator $e\colon K_3(\mathbb{C}) \to \mathbb{C}/\mathbb{Z}$.

2. Homology 3-spheres and torsion in $K_3(\mathbb{C})$

Let Σ be a homology n-sphere; since

$$0 = H_1(\Sigma, \mathbb{Z}) = \pi_1(\Sigma)/[\pi_1(\Sigma), \pi_1(\Sigma)],$$

$\pi_1(\Sigma)$ can have no abelian quotients and so it is perfect. Given a representation $\alpha\colon \pi_1(\Sigma) \to GL_N(R)$, let $f\colon \Sigma \to BGL_N(R)$ be the map which induces α on π_1. Composing this map with the inclusion $BGL_N(R) \to BGL(R)$ and applying Quillen's $+$-construction we get

$$S^n \simeq \Sigma^+ \to BGL(R)^+,$$

since the $+$-construction is functorial. Here $\simeq$ denotes homotopy equivalence. The homotopy class of this map gives us the element in K-theory

$$[\Sigma, \alpha] \in K_n(R) = \pi_n(BGL(R)^+).$$

Beilinson [3] and Karoubi [10] defined independently the **regulator** map

$$e\colon K_{2n+1}(\mathbb{C}) \to \mathbb{C}/\mathbb{Z} \tag{2.1}$$

as a secondary Chern character. Alternative constructions of this map can be found in [11] and [8]. The regulator map satisfies the following properties:

1. It is an isomorphism on $K_1(\mathbb{C}) \cong \mathbb{C}^* \to \mathbb{C}/\mathbb{Z}$.
2. The homomorphism e gives an isomorphism of the torsion subgroup of $K_{2n+1}(\mathbb{C})$ with $\mathbb{Q}/\mathbb{Z}$,
3. It vanishes on products.

For a proof see [10] or [9].

In [9] Jones and Westbury give a formula to compute the real part of $e[\Sigma, \alpha]$ when Σ is a Seifert homology sphere and α a representation in which the central element of $\pi_1(\Sigma)$ acts as a scalar multiple of the identity; for instance, this is the case when α is irreducible, and in general for any decomposable representation. This formula was obtained using the fact that

$$e([\Sigma, \alpha]) = \tilde{\xi}(\alpha, D) \in \mathbb{C}/\mathbb{Z} \tag{2.2}$$

given in [9, Thm. A], where $\tilde{\xi}(\alpha, D)$ is the reduced ξ-invariant of the Dirac operator D on Σ twisted by the representation α defined in [1, (3.2)].

They also study in detail the elements $[\Sigma(p, q, r), \alpha]$ where $\Sigma(p, q, r)$ is a Brieskorn homology 3-sphere and α some representation $\alpha\colon \pi_1(\Sigma(p, q, r)) \to SL_2(\mathbb{C})$. Brieskorn homology 3-spheres $\Sigma(p, q, r)$ with (p, q, r) pairwise coprime integers, are given explicitly by

$$\Sigma(p, q, r) = \{\, (z_1, z_2, z_3) \mid z_1^p + z_2^q + z_3^r = 0 \,\} \cap S^5 \subset \mathbb{C}^3.$$

The fundamental group $\pi_1(\Sigma(p, q, r))$ has a presentation of the form

$$\langle h, x_1, x_2, x_3 | \, [x_i, h] = 1, x_1^p = h^{-b_1}, x_2^q = h^{-b_2}, x_3^r = h^{-b_3}, x_1 x_2 x_3 = h^{-b_0} \rangle, \tag{2.3}$$

where

$$-pqrb_0 + qrb_1 + prb_2 + pqb_3 = 1.$$

It is always possible to choose b_1, b_2, b_3 to be odd; for example, if p is even then b_1 must be odd and if p is odd then replacing x_1 by hx_1 changes the parity of b_1. From now on we assume that b_1, b_2, b_3 are odd.

Consider a representation $\alpha\colon \pi_1(\Sigma(p,q,r)) \to SL_2(\mathbb{C})$; since $\pi_1(\Sigma(p,q,r))$ is perfect, any non-trivial representation of this kind is irreducible and, therefore, the central element h acts as an scalar multiple of the 2×2 identity matrix I. Thus $\alpha(h) = (-I)^f$ and, in view of the relations of the presentation (2.3) of $\pi_1(\Sigma(p,q,r))$, the representation α is given by the matrices $A = \alpha(x_1)$, $B = \alpha(x_2)$ and $C = \alpha(x_3)$ in $SL_2(\mathbb{C})$ satisfying the equations

$$A^p = (-I)^f, \quad B^q = (-I)^f, \quad C^r = (-I)^f, \quad ABC = (-I)^{fb_0}.$$

Let $\zeta_d = e^{2\pi i/d} \in \mathbb{C}$ be the standard primitive d-th root of unity. Then the respective eigenvalues of the matrices A, B and C are given by

$$\begin{aligned} &\zeta_{2p}^{k}, \zeta_{2p}^{-k} && 0<k<p,\\ &\zeta_{2q}^{l}, \zeta_{2q}^{-l} && 0<l<q,\\ &\zeta_{2r}^{m}, \zeta_{2r}^{-m} && 0<m<r, \end{aligned}$$

such that $k \equiv l \equiv m \equiv f \mod 2$. By [9, Lemma. 6.1] we have that if one of A, B, C is $\pm I$, then the representation α is trivial and by [9, Thm. 6.2] that the function given by $\alpha \mapsto (k,l,m)$ defines a one to one correspondence between conjugacy classes of non-trivial representations of $\pi_1(\Sigma(p,q,r))$ in $SL_2(\mathbb{C})$ and triples (k,l,m) with

$$0<k<p, \quad 0<l<q, \quad 0<m<r, \quad k\equiv l\equiv m \mod 2.$$

Using this characterisation of the non-trivial representations of $\pi_1(\Sigma(p,q,r))$ in $SL_2(\mathbb{C})$ and the aforementioned formula for the real part of $e[\Sigma(p,q,r),\alpha]$, Jones and Westbury proved in [9, Thm. D] that every element in $K_3(\mathbb{C})$ of finite order is of the form $[\Sigma(p,q,r),\alpha]$, for $\Sigma(p,q,r)$ a Brieskorn homology 3-sphere and some representation $\alpha\colon \pi_1(\Sigma(p,q,r)) \to SL_2(\mathbb{C})$.

The value of $e[\Sigma(p,q,r),\alpha]$ with α corresponding to the triple (k,l,m) is given by [9, Proof of Thm. D]

$$e([\Sigma(p,q,r),\alpha]) = -\frac{q^2r^2k^2+p^2r^2l^2+p^2q^2m^2}{4pqr}. \tag{2.4}$$

Let $\mathbb{Z}[\zeta_d]$ be the ring of integers in the cyclotomic field $\mathbb{Q}(\zeta_d)$. Then combining the results of Borel [4], Merkurjev and Suslin [17], and Levine [15], we have that

$$K_3(\mathbb{Z}[\zeta_d]) = K_3(\mathbb{Q}(\zeta_d)) = \mathbb{Z}/w_2(d) \oplus \mathbb{Z}^{r_2}, \tag{2.5}$$

where

$$w_2(d) = \operatorname{lcm}(24, 2d)$$

and r_2 is the number of complex places of $\mathbb{Q}(\zeta_d)$. In particular, note that if $(6,d)=1$ then the torsion subgroup of $K_3(\mathbb{Q}(\zeta_d))$ is exactly $\mathbb{Z}/24d$.

In [9, Thm. E] Jones and Westbury proved that if $(6,d)=1$ then there exists a representation $\alpha(d)\colon \pi_1(\Sigma(2,3,d)) \to SL_2(\mathbb{Z}[\zeta_d])$ such that the element $[\Sigma(2,3,d),\alpha(d)] \in SL_2(\mathbb{Z}[\zeta_d]) \subset SL_2(\mathbb{Q}(\zeta_d))$ is a generator of the torsion subgroup. They give explicitly the representation $\alpha(d)$ in [9, Proof of Thm. E].

3. Torsion in $K_3(\mathbb{R})$

By a result of Suslin [20, Thm. 4.9], we have that

$$K_3(\mathbb{R}) \cong \mathbb{Q}/\mathbb{Z} \oplus V,$$
$$K_3(\mathbb{C}) \cong \mathbb{Q}/\mathbb{Z} \oplus W,$$

where V and W are uniquely divisible groups, i.e., $\mathbb{Q}$-vector spaces. There are natural representations

$$u_n\colon GL_n(\mathbb{R}) \to GL_n(\mathbb{C}),$$
$$v_n\colon GL_n(\mathbb{C}) \to GL_{2n}(\mathbb{R}),$$

where u_n is the inclusion and, if $A \in GL_n(\mathbb{C})$, then

$$v_n(A) = \begin{pmatrix} \Re A & -\Im A \\ \Im A & \Re A \end{pmatrix}, \tag{3.1}$$

where $A = \Re A + i\Im A$. Note that $u_{2n}\circ v_n(A)$ is conjugate to $\left(\begin{smallmatrix} A & 0 \\ 0 & \bar{A} \end{smallmatrix}\right)$ by $\left(\begin{smallmatrix} \frac{1}{2}I & \frac{i}{2}I \\ iI & I \end{smallmatrix}\right)$ in $GL_n(\mathbb{C})$. These representations are compatible with stabilisation and therefore induce homomorphisms

$$i_*\colon K_3(\mathbb{R}) \to K_3(\mathbb{C}),$$
$$i^*\colon K_3(\mathbb{C}) \to K_3(\mathbb{R}),$$

where the former homomorphism corresponds to the one induced by the inclusion $i\colon \mathbb{R} \to \mathbb{C}$ and the later is called the **transfer homomorphism**. We have that the homomorphism i_* restricted to the torsion subgroup $i_*\colon \mathbb{Q}/\mathbb{Z} \xrightarrow{\times 2} \mathbb{Q}/\mathbb{Z}$ is given by multiplication by 2 and i^* restricted to the torsion subgroup $i^*\colon \mathbb{Q}/\mathbb{Z} \xrightarrow{\cong} \mathbb{Q}/\mathbb{Z}$ is an isomorphism [19, (1.18)].

Using the transfer homomorphism we can give the first result of this paper which is an analog of [9, Thm. D] for torsion elements of $K_3(\mathbb{R})$

Theorem (3.2). *Every element in $K_3(\mathbb{R})$ of finite order can be written as $[\Sigma(p,q,r),\beta]$ for some representation $\beta\colon \pi_1(\Sigma(p,q,r)) \to SL_4(\mathbb{R})$.*

Proof. By [9, Thm. D] any torsion element in $K_3(\mathbb{C})$ is of the form$[\Sigma(p,q,r),\alpha]$ with $\alpha\colon \pi_1(\Sigma(p,q,r)) \to SL_2(\mathbb{C})$. Since i^* is and isomorphism on torsion, any torsion element in $K_3(\mathbb{R})$ is of the form $i^*([\Sigma(p,q,r),\alpha])$. From the definitions of $[\Sigma(p,q,r),\alpha]$ and i^* we have that $i^*([\Sigma(p,q,r),\alpha])$=$[\Sigma(p,q,r),v_2(\alpha)]$. Thus, the theorem follows by taking $\beta = v_2(\alpha)$. We just need to check that β has image in $SL_4(\mathbb{R})$; but $\det\beta = \det u_{2n}(\beta) = \det\left(\begin{smallmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{smallmatrix}\right) = 1$ since $\alpha\colon \pi_1(\Sigma(p,q,r)) \to SL_2(\mathbb{C})$. ☐

Using Theorem (3.2) we can check that the homomorphism i_* on torsion is given by multiplication by 2. Let $a \in K_3(\mathbb{R})_{\mathrm{tor}}$; by Theorem (3.2) $a = [\Sigma(p,q,r),\beta]$ with $\beta = v_2(\alpha)$ for some representation $\alpha\colon \pi_1(\Sigma(p,q,r)) \to SL_2(\mathbb{C})$.

We have that

$$\begin{aligned} i_*(a) = i_*([\Sigma(p,q,r),\beta]) &= i_* \circ i^*([\Sigma(p,q,r),\alpha]) \\ &= [\Sigma(p,q,r), u_4(v_2(\alpha))] \\ &= [\Sigma(p,q,r), \alpha \oplus \overline{\alpha}] \\ &= [\Sigma(p,q,r),\alpha] + [\Sigma(p,q,r),\overline{\alpha}] \\ &= 2[\Sigma(p,q,r),\alpha], \end{aligned}$$

since α and $\overline{\alpha}$ are isomorphic representations because both of them determine the same triple (k,l,m).

Using formula (2.2) we can represent the unique element f of order 2 in $K_3(\mathbb{R})$ which is the generator of the kernel of $i_*\colon K_3(\mathbb{R}) \to K_3(\mathbb{C})$. The Brieskorn homology 3-sphere $\Sigma(2,3,5)$ is also known as the Poincaré homology 3-sphere. Its fundamental group is isomorphic to the binary icosahedral group $\tilde{I}$, which is the lifting of the group of isometries of a regular icosahedron $I \subset SO(3)$ to S^3 under the projection $S^3 \to SO(3)$. We have that $\Sigma(2,3,5) \cong S^3/\tilde{I}$, see [13]. The group $\tilde{I}$ has order 120 and it is the only finite group which is the fundamental group of a homology 3-sphere [12]. It has nine irreducible representations α_i, $i = 1,\ldots,9$ and the character table can be found in [6, Table IV]. In [6, p. 226] the author computed the invariants $\tilde{\xi}(\alpha_i, D)$ for the Dirac operator of $\Sigma(2,3,5)$ twisted by the representations α_i. In particular, we have that $\tilde{\xi}(\alpha_6, D) = \frac{5}{6}$ and $\tilde{\xi}(\alpha_8, D) = \frac{1}{6}$. Hence $\tilde{\xi}(\alpha_6 \oplus 4\alpha_8, D) = \frac{1}{2} \bmod \mathbb{Z}$ and by (2.2) we have that the element $[\Sigma(2,3,5), \alpha_6 \oplus 4\alpha_8] \in K_3(\mathbb{C})$ has order 2; therefore the generator of the kernel of i_* is given by

$$f = i^*([\Sigma(2,3,5), \alpha_6 \oplus 4\alpha_8]) \in K_3(\mathbb{R}).$$

4. Torsion in $K_3^{\text{ind}}(\mathbb{R})$

Let F be a field and let $K_3^{\text{dec}}(F)$ be the subgroup of $K_3(F)$ generated by products from $K_1(F)$. The indecomposable part $K_3^{\text{ind}}(F)$ of $K_3(F)$ is the quotient $K_3^{\text{ind}}(F) = K_3(F)/K_3^{\text{dec}}(F)$.

In [18] Milnor defined a graded ring $K_*^M(F)$, now known as the Milnor K-ring, to be the quotient of the tensor algebra of the multiplicative group $F^\times$ of F by the ideal generated by the homogeneous elements $x \otimes (1-x)$. The **Milnor K-group** $K_n^M(F)$ is defined to be the subgroup of elements of degree n. We shall write $\{x_1,\ldots,x_n\}$ for the image of $x_1 \otimes \cdots \otimes x_n$ in $K_n^M(F)$. There is a natural map

$$\phi_n\colon K_n^M(F) \to K_n(F),$$

which is an isomorphism for $0 \le n \le 2$. In the present paper we are interested in this map for $n = 3$ and we have that the image of $K_3^M(F)$ in $K_3(F)$ is precisely $K_3^{\text{dec}}(F)$. Therefore we have that

$$K_3^{\text{ind}}(F) = K_3(F)/K_3^M(F). \tag{4.1}$$

From [2] we have that $K_3^M(\mathbb{C})$ is a $\mathbb{Q}$-vector space, and from [18] that $K_3^M(\mathbb{R}) \cong \mathbb{Z}_2 \oplus H$, where H is a $\mathbb{Q}$-vector space. The summand $\mathbb{Z}_2$ is generated by the nontrivial symbol $\{-1,-1,-1\}$. From [5] it is known that $\phi_3\colon K_3^M(\mathbb{R}) \to K_3(\mathbb{R})$ is injective. The image of $\{-1,-1,-1\}$ under ϕ_3 is the unique element f of order 2 in $K_3(\mathbb{R})$ and generates the kernel of the homomorphism $i_*\colon K_3(\mathbb{R}) \to K_3(\mathbb{C})$.

Recall that the torsion subgroup of $K_3(\mathbb{R})$ is isomorphic to $\mathbb{Q}/\mathbb{Z}$, thus the subgroup of order 2 corresponding to the torsion subgroup of $K_3^M(\mathbb{R})$ is isomorphic to the subgroup $\frac{1}{2}\mathbb{Z}/\mathbb{Z}$ of $\mathbb{Q}/\mathbb{Z}$ and therefore $K_3^{\mathrm{ind}}(\mathbb{R})_{\mathrm{tor}} \cong \frac{\mathbb{Q}/\mathbb{Z}}{\frac{1}{2}\mathbb{Z}/\mathbb{Z}} \cong \mathbb{Q}/\frac{1}{2}\mathbb{Z}$. On the other hand, since $K_3^M(\mathbb{C})$ has no torsion, the projection $K_3(\mathbb{C}) \to K_3^{\mathrm{ind}}(\mathbb{C})$ is an isomorphism in the torsion subgroups. Hence we have the following commutative diagram

$$\begin{tikzcd} K_3(\mathbb{R})_{\mathrm{tor}} \cong \mathbb{Q}/\mathbb{Z} \arrow[r, "i_*"] \arrow[d, "p"'] & \mathbb{Q}/\mathbb{Z} \cong K_3(\mathbb{C})_{\mathrm{tor}} \arrow[d, "\cong"] \\ K_3^{\mathrm{ind}}(\mathbb{R})_{\mathrm{tor}} \cong \mathbb{Q}/\frac{1}{2}\mathbb{Z} \arrow[r, "\cong"'] & \mathbb{Q}/\mathbb{Z} \cong K_3^{\mathrm{ind}}(\mathbb{C})_{\mathrm{tor}}. \end{tikzcd} \tag{4.2}$$

Since i_* is given by multiplication by 2, we have that the lower isomorphism is given by

$$\begin{tikzcd} \mathbb{Q}/\frac{1}{2}\mathbb{Z} \arrow[r, shift left, "\times 2"] & \mathbb{Q}/\mathbb{Z}. \arrow[l, shift left, "\times \frac{1}{2}"] \end{tikzcd}$$

This agrees with the result proved independently and simultaneously by Levin [15, Cor. 4.6] and Mercurjev–Suslin [17, Prop. 11.3] which says that if E is an extension field of F then the natural homomorphism $K_3^{\mathrm{ind}}(F) \to K_3^{\mathrm{ind}}(E)$ is injective.

If $[\Sigma(p,q,r),\beta] \in K_3(\mathbb{R})_{\mathrm{tor}}$ we shall denote its image in $K_3^{\mathrm{ind}}(\mathbb{R})_{\mathrm{tor}}$ by $\langle\Sigma(p,q,r),\beta\rangle$, that is, $\langle\Sigma(p,q,r),\beta\rangle$ is the coset $[\Sigma(p,q,r),\beta] + \langle f\rangle$ with $\langle f\rangle$ the subgroup of order 2 generated by f. Thus we have a result analogous to Theorem (3.2) for $K_3^{\mathrm{ind}}(\mathbb{R})_{\mathrm{tor}}$.

Theorem (4.3). *Every element in $K_3^{\mathrm{ind}}(\mathbb{R})$ of finite order can be written as $\langle\Sigma(p,q,r),\beta\rangle$ for some representation $\beta\colon \pi_1(\Sigma(p,q,r)) \to SL_4(\mathbb{R})$.*

Let $\mathbb{Q}(\zeta_d)^+$ be the real part of the cyclotomic field $\mathbb{Q}(\zeta_d)$. Again by the results of Borel [4], Merkurjev and Suslin [17] and Levine [15] we have that

$$K_3(\mathbb{Q}(\zeta_d)^+) = \mathbb{Z}/2w_2(d) \oplus (\mathbb{Z}/2)^{r_1-1} \oplus \mathbb{Z}^{r_2} \tag{4.4}$$

where as before

$$w_2(d) = \mathrm{lcm}(24, 2d)$$

and r_1 and r_2 are respectively the number of real and complex places of $\mathbb{Q}(\zeta_d)$.

On the other hand, in [2] Bass and Tate proved that, for a number field F, $K_n^M(F) \cong (\mathbb{Z}/2)^{r_1}$ where r_1 is the number of real places of F. Since $\mathbb{Q}(\zeta_d)$ is totally imaginary ($r_1 = 0$), $K_3^M(\mathbb{Q}(\zeta_d))$ has no torsion. Then combining (4.1), (2.5) and (4.4) we have that

$$\begin{aligned} K_3^{\mathrm{ind}}(\mathbb{Q}(\zeta_d))_{\mathrm{tor}} &\cong K_3(\mathbb{Q}(\zeta_d))_{\mathrm{tor}} \cong \mathbb{Z}/w_2(d), \\ K_3^{\mathrm{ind}}(\mathbb{Q}(\zeta_d)^+)_{\mathrm{tor}} &\cong K_3(\mathbb{Q}(\zeta_d)^+)_{\mathrm{tor}}/(\mathbb{Z}/2) \cong \mathbb{Z}/w_2(d). \end{aligned}$$

Thus we have a commutative diagram analogous to (4.2):
(4.5)

$$\begin{tikzcd}
K_3(\mathbb{Q}(\zeta_d)^+)_{\mathrm{tor}} \cong \mathbb{Z}/2w_2(d) \oplus (\mathbb{Z}/2)^{r_1-1} \arrow[r, "i_*"] \arrow[d, "p"'] & \mathbb{Z}/w_2(d) \cong K_3(\mathbb{Q}(\zeta_d))_{\mathrm{tor}} \arrow[d, "\cong"] \\
K_3^{\mathrm{ind}}(\mathbb{Q}(\zeta_d)^+)_{\mathrm{tor}} \cong \mathbb{Z}/w_2(d) \arrow[r, "\cong"'] & \mathbb{Z}/w_2(d) \cong K_3^{\mathrm{ind}}(\mathbb{Q}(\zeta_d))_{\mathrm{tor}}.
\end{tikzcd}$$

Since i_* is given by multiplication by 2, we have that an element of order $2w_2(d)$ of $K_3(\mathbb{Q}(\zeta_d)^+)_{\mathrm{tor}}$ is sent by i_* to a generator of $K_3(\mathbb{Q}(\zeta_d))_{\mathrm{tor}}$.

Problem (4.5). It would be good to prove a theorem analogous to Theorem E of [9] for $K_3(\mathbb{Q}(\zeta_d)^+)$, that is, for some Brieskorn homology 3-sphere $\Sigma(p,q,r)$ find a representation $\beta(d)\colon \pi_1(\Sigma(p,q,r)) \to SL_n(\mathbb{Q}(\zeta_d)^+)$ such that the element $[\Sigma(p,q,r),\beta(d)] \in K_3(\mathbb{Q}(\zeta_d)^+)$ is an element of order $2w_2(d)$. Then using diagram (4.5) we would also get generators for the torsion subgroups of $K_3^{\mathrm{ind}}(\mathbb{Q}(\zeta_d)^+)$ and $K_3(\mathbb{Q}(\zeta_d))$.

One possible way to do this is to find an element $[\Sigma(p,q,r),\gamma]$ in $K_3(\mathbb{C})$ of order $2w_2(d)$, then the element $i^*([\Sigma(p,q,r),\gamma]) = [\Sigma(p,q,r),v_2(\gamma)] \in K_3(\mathbb{R})$ also has order $2w_2(d)$. It would then be enough to show that $v_2(\gamma)$ is conjugate in $SL_4(\mathbb{R})$ to a representation with image in $SL_4(\mathbb{Q}(\zeta_d)^+)$.

Remark (4.5). For the case $(6,d)=1$ we have that $K_3(\mathbb{Q}(\zeta_d)^+) \cong \mathbb{Z}/48d \oplus (\mathbb{Z}/2)^{r_1-1}$ and $K_3^{\mathrm{ind}}(\mathbb{Q}(\zeta_d)^+)_{\mathrm{tor}} \cong K_3^{\mathrm{ind}}(\mathbb{Q}(\zeta_d))_{\mathrm{tor}} \cong K_3(\mathbb{Q}(\zeta_d))_{\mathrm{tor}} \cong \mathbb{Z}/24d$. Using the representation $\alpha(d)$ of the generator (of order $24d$) $[\Sigma(2,3,d),\alpha(d)] \in K_3(\mathbb{Q}(\zeta_d))$ given in [9, Proof of Thm. E], one could try to find the representation $\beta(d)$ showing that the representation $\alpha(d)\colon \pi_1(\Sigma(2,3,d)) \to SL_2(\mathbb{Q}(\zeta_d))$ is conjugate in $SL_2(\mathbb{C})$ to a representation $\beta(d)$ with image in $SL_2(\mathbb{Q}(\zeta_d)^+)$. Such representation would give an element $[\Sigma(p,q,r),\beta(d)] \in K_3(\mathbb{Q}(\zeta_d)^+)$. Considering $\beta(d)$ as a complex representation, that is, taking $u_2(\beta)$, it is conjugate to $\alpha(d)$ and therefore $i_*([\Sigma(2,3,d),\beta(d)]) = [\Sigma(2,3,d),\alpha(d)]$. Since $[\Sigma(2,3,d),\alpha(d)]$ has order $24d$ and i_* is given by multiplication by 2, the element $[\Sigma(2,3,d),\beta(d)] \in K_3(\mathbb{Q}(\zeta_d)^+)$ would have order $48d$.

5. The Bloch group and dilogarithm identities

In this section we define the Bloch group $B(F)$ of a field F, which is a group closely related to $K_3^{\mathrm{ind}}(F)$. Next, we define the dilogarithm and state some of its functional identities, and show how Frenkel and Szenes used the dilogarithm identities in [7] to construct generators of the torsion subgroup of the Bloch group of totally real fields.

Let $F^\times$ the multiplicative group of F. Let $D(F)$ be the free abelian group generated by formal symbols $[x]$ with $x \in F \setminus \{0,1\}$. Let $C(F)$ be the kernel of the map

$$D(F) \xrightarrow{\lambda} F^\times \wedge F^\times, \qquad [x] \mapsto (x \wedge (1-x)). \tag{5.1}$$

One can check that the elements of the form

$$[x]-[y]+[y/x]-[(1-x^{-1})/(1-y^{-1})]+[(1-x)/(1-y)] \tag{5.2}$$

with $x \neq y \in F^\times \setminus \{1\}$ are contained in $C(F)$. The quotient $B(F)$ of $C(F)$ by the subgroup generated by the elements of this form is called the **Bloch group**.

The following exact sequence due to Suslin [21, Thm. 5.2] gives the precise relation between $K_3^{\mathrm{ind}}(F)$ and the Bloch group

$$0 \to \widetilde{\mathrm{Tor}(F^\times,F^\times)} \to K_3^{\mathrm{ind}}(F) \xrightarrow{\tau} B(F) \to 0, \tag{5.3}$$

where $\widetilde{\mathrm{Tor}(F^\times,F^\times)}$ is the unique non-trivial extension of $\mathbb{Z}/2$ by the group $\mathrm{Tor}(F^\times,F^\times)$.

From now on let F be a totally real field of algebraic numbers. Then

$$\widetilde{\mathrm{Tor}(F^\times,F^\times)} \cong \mathbb{Z}/4.$$

Therefore $B(F) \cong K_3^{\mathrm{ind}}(F)/\mathbb{Z}_4$. In particular, $B(\mathbb{Q}(\zeta_d)^+)$ is cyclic of order $\frac{6d}{(12,d)}$.

In [7] Frenkel and Szenes, using the Rogers' dilogarithm, defined a map $\mathcal{L}\colon B(\mathbb{R}) \to \mathbb{R}/(\pi^2\mathbb{Z})$ and, using dilogarithm identities, they constructed a set of generator for $B(F)$ with F a totally real field. We shall now sketch their construction:

The Rogers' dilogarithm is defined by

$$\mathrm{L}(x) = -\frac{1}{2}\int_0^x \Big(\frac{\log(1-y)}{y} + \frac{\log y}{1-y}\Big)dy$$

and it satisfies the following functional identities (see for instance [16, 14, 7]):

$$\mathrm{L}(x)+\mathrm{L}(1-x)=\mathrm{L}(1)=\frac{\pi^2}{6}, \tag{5.4}$$

$$\mathrm{L}(x)+\mathrm{L}(y)=\mathrm{L}(xy)+\mathrm{L}\big(\frac{x(1-y)}{1-xy}\big)+\mathrm{L}\big(\frac{y(1-x)}{1-xy}\big), \tag{5.5}$$

$$\sum_{j=1}^{k}\mathrm{L}\Big(\frac{\sin^2\frac{\pi}{k+2}}{\sin^2\frac{(j+1)\pi}{k+2}}\Big)=\frac{3k}{k+2}\cdot\frac{\pi^2}{6}. \tag{5.6}$$

Let $\bar{\mathcal{L}}'\colon D(\mathbb{R}) \to \mathbb{R}$ be the map which sends $[x]$ to $\mathrm{L}(x)$. We can restrict it to $C(\mathbb{R})$. Now consider the slightly modified map $\bar{\mathcal{L}} = \bar{\mathcal{L}}' - \frac{\pi^2}{6}\colon D(\mathbb{R}) \to \mathbb{R}$. Then for any element α of $C(\mathbb{R})$ of the form (5.2), by relations (5.4) and (5.5), one has $\bar{\mathcal{L}}(\alpha) = 0 \bmod \pi^2$. Hence this map gives rise to a well defined homomorphism

$$\mathcal{L}\colon B(\mathbb{R}) \to \mathbb{R}/(\pi^2\mathbb{Z}).$$

Let $\mathbb{Q}(\zeta_d)^+$ be the real part of the cyclotomic field $\mathbb{Q}(\zeta_d)$. We now describe the construction of torsion elements in $B(\mathbb{Q}(\zeta_d)^+)$. Put

$$\delta_j(d) = \frac{\sin^2\frac{\pi}{d}}{\sin^2\frac{\pi(j+1)}{d}}, \qquad j=1,\dots,d-3, \quad d>3.$$

Since $\zeta_d^k = \cos\frac{2\pi k}{d} + i\sin\frac{2\pi k}{d} \in \mathbb{Q}(\zeta_d)$, and $\sin^2\theta = \frac12 + \frac12\cos 2\theta$, we have that $\delta_j(d) \in \mathbb{Q}(\zeta_d)^+$. Define the symbols $\Delta_d \in D(\mathbb{Q}(\zeta_d)^+)$ by the formula

$$\Delta_d = 2\sum_{j=1}^{d-3}[\delta_j(d)].$$

By [7, Lemma 5.3], for each $d > 3$, $\Delta_d \in C(\mathbb{Q}(\zeta_d)^+)$. Thus the symbol Δ_d represents an element of $B(\mathbb{Q}(\zeta_d)^+)$.

Using the symbols Δ_d and the map $\mathcal{L}\colon B(\mathbb{R}) \to \mathbb{R}/(\pi^2\mathbb{Z})$, Frenkel and Szenes constructed generators of the Bloch group $B(F)$ for F a totally real field.

Proposition (5.7) (Frenkel–Szenes [7, Prop. 5.4]). *Let F be a totally real number field and m_p the maximal number $m \geq 0$ such that F contains $\mathbb{Q}(\zeta_{p^m})^+$. Then the symbols $\Delta_{p^{m_p}}$ and the symbol Δ_6, if $m_3 = 1$, generate the group $B(F) = K_3^{\mathrm{ind}}(F)/\mathbb{Z}_4$.*

Sketch of proof: The proof uses the following facts:

• By results of Merkurjev and Suslin [17] and Levine [15], for a totally real field F, the group $K_3^{\mathrm{ind}}(F)$ is isomorphic to the cyclic group of order $2\prod p^{m_p}$, where the product is taken over all primes. Hence the group $B(F)$ is cyclic of order $b(F) = \frac{1}{2}\prod p^{m_p}$.

• The symbol $\Delta_6 = 4[\frac{1}{3}] + 2[\frac{1}{4}] \in B(\mathbb{Q})$ belongs to $B(F)$.

• By the identity (5.6) we have

$$\begin{aligned} \mathcal{L}(\Delta_d) &= \Big(\frac{2-d}{3}\pi^2 - \frac{2}{d}\pi^2\Big) \bmod \pi^2, \\ \mathcal{L}(\Delta_6) &= \frac{\pi^2}{3} \bmod \pi^2, \end{aligned} \tag{5.8}$$

and therefore

$$\mathcal{L}(\Delta_{p^{m_p}}) - (2 - p^{m_p})\mathcal{L}(\Delta_6) = -\frac{2}{p^{m_p}}\pi^2 \bmod \pi^2.$$

The order of Δ_d is at least the order of its image, therefore these elements of $B(F)$ generate a cyclic group of order at least $b(F)$, hence they generate the whole group $B(F)$. Note that $m_3 \geq 1$ for any number field. If $m_3 = 1$, Δ_6 generates the 3-torsion subgroup of $B(F)$. If $m_3 > 1$, then $\Delta_6 = 3^{m_3-1}\Delta_{3^{m_3}}$. □

The third fact shows that the images of the elements Δ_d under the homomorphism $\mathcal{L}$ have the same order as the orders of these elements. This implies that for a totally real number field F the homomorphism $\mathcal{L}\colon B(F) \to \mathbb{R}/(\pi^2\mathbb{Z})$ is injective [7, Cor. 5.5].

6. Relations with the regulator map

It is known that the torsion subgroup of $B(\mathbb{R})$ is generated by the images of the groups $B(\mathbb{Q}(\zeta_d)^+)$ of real parts of cyclotomic fields, and is therefore isomorphic to $\mathbb{Q}/\mathbb{Z}$. Thus we see that it is generated by the symbols Δ_d, and that the map $\mathcal{L}$ is injective on the torsion subgroup of $B(\mathbb{R})$.

Now we discuss the connection between the map $\mathcal{L}$ and the regulator. Note that since the regulator (2.1) vanishes on products it descends to a homomorphism

$$K_3^{\mathrm{ind}}(\mathbb{C}) \to \mathbb{C}/\mathbb{Z}.$$

By results of Levin [15, Cor. 4.6] and Mercurjev–Suslin [17, Prop. 11.3] $K_3^{\mathrm{ind}}(\mathbb{R})$ embeds into $K_3^{\mathrm{ind}}(\mathbb{C})$. Under this embedding, the torsion part of $K_3^{\mathrm{ind}}(\mathbb{R})$ is mapped isomorphically onto the torsion subgroup of $K_3^{\mathrm{ind}}(\mathbb{C})$, which coincides with the subgroup $\operatorname{Tor}(\mathbb{C},\mathbb{C})^{\sim} \cong \mathbb{Q}/\mathbb{Z}$ of $K_3^{\mathrm{ind}}(\mathbb{C})$. Note that there is no torsion in the Bloch group $B(\mathbb{C})$, since $\operatorname{Tor}(\mathbb{C},\mathbb{C})^{\sim}$ is the torsion subgroup of $K_3^{\mathrm{ind}}(\mathbb{C})$ and by (5.3) $B(\mathbb{C}) \cong K_3^{\mathrm{ind}}(\mathbb{C})/\operatorname{Tor}(\mathbb{C},\mathbb{C})^{\sim}$. Because of that we cannot extend the map $\mathcal{L}$ from $B(\mathbb{R})$ to $B(\mathbb{C})$ (the five-term relation for the Rogers' dilogarithm does not hold for complex arguments).

In order to compare the map $\mathcal{L}$ with the regulator we need to renormalise the last one composing it with the homomorphism $\mathbb{C}/\mathbb{Z} \xrightarrow{\sim} \mathbb{C}/(2\pi i)^2\mathbb{Z}$ to get a map

$$r\colon K_3^{\mathrm{ind}}(\mathbb{C}) \to \mathbb{C}/(2\pi i)^2\mathbb{Z}.$$

Taking the composition

$$K_3^{\mathrm{ind}}(\mathbb{R}) \to K_3^{\mathrm{ind}}(\mathbb{C}) \xrightarrow{r} \mathbb{C}/(2\pi i)^2\mathbb{Z} \xrightarrow{\arg(\cdot)} \mathbb{R}/(2\pi i)^2\mathbb{Z}$$

and then taking the projection $\mathbb{R}/(2\pi i)^2\mathbb{Z} \to \mathbb{R}/(\pi^2\mathbb{Z})$ to kill the subgroup $\operatorname{Tor}(\mathbb{R}^\times,\mathbb{R}^\times)^{\sim}$ of $K_3^{\mathrm{ind}}(\mathbb{R})$, gives a map

$$\bar{r}\colon K_3^{\mathrm{ind}}(\mathbb{C}) \to \mathbb{R}/\pi^2\mathbb{Z} \tag{6.1}$$

which descends to a map

$$\tilde{r}\colon B(\mathbb{R}) \to \mathbb{R}/\pi^2\mathbb{Z}.$$

In [7] Frenkel and Szenes conjectured that this map coincides with the map $\mathcal{L}$. This is equivalent to saying that the diagram

$$\begin{tikzcd}
K_3^{\mathrm{ind}}(\mathbb{R}) \arrow[r] \arrow[d, "\tau"'] & K_3^{\mathrm{ind}}(\mathbb{C}) \arrow[r] & \mathbb{C}/(2\pi i)^2\mathbb{Z} \arrow[r, "\arg(\cdot)"] & \mathbb{R}/(2\pi i)^2\mathbb{Z} \arrow[d] \\
B(\mathbb{R}) \arrow[rrr, "\mathcal{L}"'] & & & \mathbb{R}/\pi^2\mathbb{Z}
\end{tikzcd} \tag{6.2}$$

commutes, where τ is the map in Suslin's exact sequence (5.3).

Instead of proving that $\bar{r} = \mathcal{L} \circ \tau$, one could try to prove that the image of $\bar{r}$ is contained in the image of $\mathcal{L}$ and using the fact that $\mathcal{L}$ is injective, one could take the composition $\psi = \mathcal{L}^{-1} \circ \bar{r}$ to get a homomorphism

$$\psi\colon K_3^{\mathrm{ind}}(\mathbb{R}) \to B(\mathbb{R}),$$

which by definition makes the diagram (6.2) commute after replacing τ by ψ. Hence, to prove the conjecture it would be enough to prove that ψ is precisely the homomorphism τ.

Combining Theorem (4.3) with Proposition (5.7) we can do this, but restricted to the torsion subgroups, and define an explicit homomorphism

$$\psi\colon K_3^{\mathrm{ind}}(\mathbb{R})_{\mathrm{tor}} \to B(\mathbb{R})_{\mathrm{tor}}.$$

THEOREM (6.3). *There is a homomorphism $\psi\colon K_3^{\mathrm{ind}}(\mathbb{R})_{\mathrm{tor}} \to B(\mathbb{R})_{\mathrm{tor}}$ given by*

$$\langle \Sigma(p,q,r), \beta\rangle \mapsto C\big((2-pqr)\Delta_6 - \Delta_{pqr}\big),$$

where $\beta = v_2(\alpha)$ with $\alpha\colon \pi_1(\Sigma(p,q,r)) \to SL_2(\mathbb{C})$, $C = q^2r^2k^2 + p^2r^2l^2 + p^2q^2m^2$, and (k,l,m) is the triple corresponding to the representation α.

Proof. Recall that

$$\begin{aligned} i_*(\langle \Sigma(p,q,r),\beta\rangle) &= [\Sigma(p,q,r), u_4 \circ v_2(\alpha)] \\ &= [\Sigma(p,q,r), \alpha \oplus \overline{\alpha}] \\ &= 2[\Sigma(p,q,r),\alpha]. \end{aligned}$$

Combining (2.4) with the definition of the map $\bar{r}\colon K_3^{\mathrm{ind}}(\mathbb{R}) \to \mathbb{R}/\pi^2\mathbb{Z}$ in page 126 we have that

$$\bar{r}([\Sigma(p,q,r),\beta]) = \frac{2C}{pqr}\pi^2,$$

but by (5.8) this is precisely the image of $C\big((2-pqr)\Delta_6 - \Delta_{pqr}\big)$ under the map $\mathcal{L}$. Finally, the injectivity of $\mathcal{L}$ makes the homomorphism well-defined. $\square$

Remark (6.3). As we mentioned above, it would be of interest to compare the map $\psi\colon K_3^{\mathrm{ind}}(\mathbb{R})_{\mathrm{tor}} \to B(\mathbb{R})_{\mathrm{tor}}$ of Theorem (6.3) with the map $\tau\colon K_3^{\mathrm{ind}}(\mathbb{R}) \to B(\mathbb{R})$ of Suslin's exact sequence (5.3) restricted to the torsion subgroup. If they turn out to be the same, this would prove the Frenkel–Szenes conjecture at the torsion level.

Acknowledgements

The author is grateful to the referees for their careful reading of the original manuscript and helpful suggestions to improve the presentation of the paper.

Received August 22, 2003

Final version received January 26, 2004

Instituto de Matemáticas, UNAM
Av. Universidad S/N
Col. Lomas de Chamilpa
62210 Cuernavaca, Morelos
México
jlcm@matcuer.unam.mx

References

[1] M. F. Atiyah, V. K. Patodi, and I. M. Singer, *Spectral asymmetry and Riemannian geometry II*, Math. Proc. Cam. Phil. Soc. **78** (1975), 405–432.

[2] H. Bass and J. Tate, *The Milnor ring of a global field*, Lecture Notes in Math. **342**, Springer-Verlag, Berlin, (1973), 349–446.

[3] A. A. Beilinson, *Higher regulators and values of L-functions of curves*, Funct. Anal. Appl. **14** (1980), 116–117.

[4] A. Borel, *Stable real cohomology of arithmetic groups*, Ann. Sci. Ecole Norm. Sup. **7** (1974), 235–272.

[5] W. Browder, *Algebraic K-theory with coefficients $\mathbb{Z}_p$*, Lecture Notes in Math. **657** Springer-Verlag, (1977), 40–84.

[6] J. L. Cisneros Molina, *The η-invariant of twisted Dirac operators of S^3/Γ*, Geometriae Dedicata **84** (2001), 207–228.

[7] E. Frenkel and A. Szenes, *Crystal bases, dilogarithm identities, and torsion in algebraic K-theory*, J. Am. Math. Soc. **8** (1995), 629–664.

[8] C. E. Hood and J. D. S. Jones, *Some algebraic properties of cyclic homology groups*, K-Theory **1** (1987), 361–384.

[9] J. D. S. Jones and B. W. Westbury, *Algebraic K-theory, homology spheres, and the η-invariant*, Topology **34** (1994), 929–957.

[10] M. Karoubi, *Homologie cyclique et K-théorie*, Astérisque. Société Mathématique de France, 1987.

[11] M. Karoubi, *Théorie générale des classes charactéristiques secondaires*, K-Theory **4** (1990), 55–87.

[12] M. A. Kervaire, *Smooth homology spheres and their fundamental groups*, Tran. Am. Math. Soc. **144** (1969), 67–72.

[13] R. C. Kirby and M. G. Scharlemann, *Eight faces of the Poincaré homology 3-sphere*, In James C. Cantrell, editor, Geometric Topology, 113–146. Academic Press, 1977. Proceedings of the 1977 Georgia Topology Conference.

[14] A. N. Kirillov and N. Reshetikhin, *Exact solution of the xxz Heisenberg model of spin S*. J. Soviet Math. **35** (1986), 2627–2643.

[15] M. Levine, *The indecomposable K_3 of fields*, Ann. Sci. Ecole Norm. Sup. **22** (1989), 255–344.

[16] L. Lewin, Dilogarithms and Associated Functions, Macdonald, London, 1958.

[17] A. S. Merkurjev and A. A. Suslin, *The group K_3 for a field*, Math. USSR Izvestiya, **36** (3), (1991), 541–565.

[18] J. Milnor, *Algebraic K-theory and quadratic forms*, Invent. Math. **9** (1970), 318–344.

[19] C. H. Sah, *Homology of classical lie groups made discrete III*, J. Pure Appl. Algebra **56** (1989), 269–312.

[20] A. A. Suslin, *On the K-theory of local fields*, J. Pure and Appl. Algebra **34** (1984), 301–318.

[21] A. A. Suslin, *K_3 of a field and the Bloch group*, Proc. Steklov Inst. Math. **4** (1991), 217–239.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

VOLUMES FOR TWIST LINK CONE-MANIFOLDS

D. DEREVNIN, A. MEDNYKH AND M. MULAZZANI

ABSTRACT. Recently, the explicit volume formulae for hyperbolic cone-manifolds, whose underlying space is the 3-sphere and the singular set is the knot 4_1 and the links 5_1^2 and 6_2^2, have been obtained by the second named author and his collaborators. In this paper we explicitly find the hyperbolic volume for cone-manifolds with the link 6_3^2 as singular set. Trigonometric identities (Tangent, Sine and Cosine Rules) between complex lengths of singular components and cone angles are obtained for an infinite family of two-bridge links containing 5_1^2 and 6_3^2.

1. Introduction

Starting from Alexander's works, polynomial invariants have become a very convenient instrument for knot investigation. Several kinds of knots polynomials have been discovered in the last twenty years. Among these, we recall the Jones-, Kaufmann-, HOMFLY-, A-polynomials and others ([12], [3], [8]). These polynomials relate knot theory to algebra and algebraic geometry. Algebraic techniques are used to find the most important geometrical characteristics of knots, such as volume, length of shortest geodesics and others.

The explicit volume formulae for hyperbolic cone-manifolds, whose underlying space is the 3-sphere and the singular set is the knot 4_1 and the links 5_1^2 and 6_2^2, have been obtained in [17], [19] and [15].

The aim of our paper is to explicitly find the hyperbolic volume for cone-manifolds with the link 6_3^2 as singular set. In order to do this, we will introduce a family of hyperbolic cone-manifolds $W_p(\alpha, \beta)$, with the two-bridge links W_p, with slope $(4p+4)/(2p+1)$ as singular set, and α, β as cone angles.

Trigonometric identities (Tangent, Sine and Cosine Rules) between complex lengths of singular components and cone angles for $W_p(\alpha, \beta)$ are obtained. Then the Schläfli formula applies in order to find explicit hyperbolic volumes for cone-manifolds $W_2(\alpha, \beta)$.

In the present paper links and knots are considered as singular subsets of the three-sphere endowed by a Riemannian metric of negative constant curvature.

2000 *Mathematics Subject Classification*: Primary 57M50; Secondary 57M25, 57M27.

Keywords and phrases: Hyperbolic orbifold, hyperbolic cone-manifold, complex length, Tangent Rule, Sine Rule, Cosine Rule, hyperbolic volume.

Partially supported by RFBR (grant 03-01-00104), INTAS (grant 03-51-3663) and Scientific Schools (grant SS-300.2003.1).

2. Trigonometric identities for knots and links

(2.1) Cone-manifolds, complex distances and lengths. We start with the definition of cone-manifold modelled in a hyperbolic, spherical, or Euclidian structure.

Definition (2.1.1). A 3-dimensional *hyperbolic cone-manifold* is a Riemannian 3-dimensional manifold of constant negative sectional curvature with cone-type singularities along simple closed geodesics.

To each component of the singular set is associated a real number $n \geq 1$ such that the cone-angle around the component is $\alpha = 2\pi/n$. The concept of hyperbolic cone-manifold generalizes that of hyperbolic manifold, which appears in the partial case when all cone-angles are 2π. Hyperbolic cone-manifolds are also a generalization of hyperbolic 3-orbifolds, which arises when all associated numbers n are integers. Euclidean and spherical cone-manifolds are defined similarly.

In the present paper hyperbolic, spherical or Euclidean cone-manifolds C are considered whose underlying space is the three-dimensional sphere and the singular set $\Sigma = \Sigma^1 \cup \Sigma^2 \cup \ldots \cup \Sigma^k$ is a link consisting of the components $\Sigma^j = \Sigma^j(\alpha_j)$, $j = 1, 2, \ldots, k$ with cone-angles $\alpha_1, \ldots, \alpha_k$ respectively.

We recall a few well-known facts from hyperbolic geometry.

Let $\mathbb{H}^3 = \{(z, \xi) \in \mathbb{C} \times \mathbb{R} : \xi > 0\}$ be the upper half space model of the 3 -dimensional hyperbolic space endowed by the Riemannian metric

$$ds^2 = \frac{dzd\overline{z} + d\xi^2}{\xi^2}.$$

We identify the group of orientation preserving isometries of $\mathbb{H}^3$ with the group $PSL(2, \mathbb{C})$, consisting of linear fractional transformations

$$A' \colon z \in \mathbb{C} \to \frac{az + b}{cz + d}.$$

By a canonical procedure, A' can be uniquely extended to an isometry of $\mathbb{H}^3$. We prefer to deal with the matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{C})$ rather than the element $A' \in PSL(2, \mathbb{C})$. The matrix A is uniquely determined by the element A', up to a sign. In the following we will use the same letter A for both A and A', as long as this does not create confusion.

Let C be a hyperbolic cone-manifold with the singular set Σ. Then C defines a nonsingular but incomplete hyperbolic manifold $\mathcal{M} = C - \Sigma$. Denote by Φ the fundamental group of the manifold $\mathcal{M}$.

The hyperbolic structure of $\mathcal{M}$ defines, up to conjugation in $PSL(2, \mathbb{C})$, a holonomy homomorphism

$$\hat{h} \colon \Phi \to PSL(2, \mathbb{C}).$$

It is shown in [23] that the holonomy homomorphism of an orientable cone-manifold can be lifted to $SL(2, \mathbb{C})$ if all cone-angles are at most π. Denote by $h : \Phi \to SL(2, \mathbb{C})$ this lifting homomorphism. Choose an orientation on the link $\Sigma = \Sigma^1 \cup \Sigma^2 \cup \ldots \cup \Sigma^k$ and fix a meridian-longitude pair $\{m_j, l_j\}$ for each component $\Sigma^j = \Sigma^j(\alpha_j)$. Then the matrices $M_j = h(m_j)$ and $L_j = h(l_j)$ satisfy the following properties:

$$\mathrm{tr}\,(M_j) = 2\cos(\alpha_j/2), \quad M_jL_j = L_jM_j\,, \;\; j = 1, 2, \ldots, k\,.$$

Now we point out some definitions and results from the book [4]. A matrix $A \in SL(2,\mathbb{C})$ satisfying $\mathrm{tr}\,(A) = 0$ is called a (normalized) *line matrix.* We have by definition that $A^2 = -I$, where I is the identity matrix. Hence any line matrix determines a half-turn about a line in $\mathbb{H}^3$, and this line determines the matrix up to sign. According to [4, p. 63], there exists a natural one-to-one correspondence between line matrices and oriented lines in $\mathbb{H}^3$. Hereby, if a line matrix A determines an oriented line $\lambda_A = [e, e']$ with end points e and e', then the line matrix $-A$ determines the line $[e', e]$. Moreover, if a matrix $F \in SL(2,\mathbb{C})$ is considered as a motion of $\mathbb{H}^3$, then the matrix FAF^{-1} determines the line $[F(e), F(e')]$.

Definition (2.1.2). Let λ_A and λ_B be oriented lines determined by the line matrices A and B. A complex number μ is called a *complex distance* from λ_A to λ_B if its real part $\Re\,\mu$ is the distance from λ_A to λ_B, and its imaginary part $\Im\,\mu$ is the angle from λ_A to λ_B chosen in $[0, 2\pi)$.

We have [4, p. 68]

$$\cosh\mu = -\frac{1}{2}\,\mathrm{tr}\,(AB). \tag{2.1.3}$$

From now on, all lines in this paper will be assumed to be oriented.

Any isometry A of $\mathbb{H}^3$ which is neither parabolic nor the identity has two fixed points u and v in $\widehat{\mathbb{C}}$. It acts as a translation of distance r_A along the axis $\lambda_A = [u, v]$ and rotation of φ_A about λ_A.

Definition (2.1.4). We call *displacement* of A the complex number $\delta(A) = r_A + i\varphi_A$.

The isometry A, without an orientation of its axis, determines $\delta(A)$ up to sign. By [4, p. 46], for the isometry given by a matrix $A \in SL(2,\mathbb{C})$ we have

$$2\cosh\delta(A) = \mathrm{tr}\,(A^2) = \mathrm{tr}^{\,2}(A) - 2.$$

We remark that if $\delta(A) \neq 0$ then A has two different fixed points, so it admits an axis determined by these points. The line matrix $\widetilde{A}$ of this axis is defined by

$$\widetilde{A} = \frac{A - A^{-1}}{2i\sinh\frac{\delta(A)}{2}} \tag{2.1.5}$$

Since $\delta(A^{-1}) = -\delta(A)$, the matrices A and A^{-1} define the same line matrix $\widetilde{A} = \widetilde{A^{-1}}$ (see [4]).

Definition (2.1.6). The *complex length* γ_j of a singular component Σ^j of the cone-manifold C is the displacement $\delta(L_j)$ of the isometry L_j, where $L_j = h(l_j)$ is represented by the longitude l_j of Σ^j.

Immediately from the definition we get [4, p. 46]

$$2\cosh\gamma_j = \mathrm{tr}\,(L_j^2). \tag{2.1.7}$$

We note [2, p. 38] that the meridian-longitude pair $\{m_j, l_j\}$ of the oriented link is uniquely determined up to a common conjugating element of the group Φ. Hence, the complex length $\gamma_j = r_j + i\,\varphi_j$ is uniquely determined (mod $2\pi i$), up to a sign, by the above definition.

We need two conventions to correctly choose real and imaginary parts of γ_j. The first convention is the following. By the assumptions on the singular set we have $r_j \neq 0$. Hence, we can choose γ_j in such a way that $r_j > 0$. The second convention concerns the imaginary part φ_j. We want to choose φ_j so that the following identity holds

$$\cosh\frac{\gamma_j}{2} = -\frac{1}{2}\,\mathrm{tr}\,(L_j) \tag{2.1.8}$$

By virtue of the identity $\mathrm{tr}^2(L_j) - 2 = \mathrm{tr}\,(L_j^2)$, the equality (2.1.7) is a consequence of (2.1.8), but the converse, in general, is true only up to a sign. Under the second convention (2.1.7) and (2.1.8) are equivalent. The two above conventions lead to convenient analytic formulas in order to calculate γ_j and r_j. More precisely, there are simple relations between these numbers and the eigenvalues of the matrix L_j. Recall that $\det(L_j) = 1$. Since L_j is loxodromic, it has two eigenvalues f_j and $1/f_j$. We choose f_j so that $|f_j| > 1$. The case $|f_j| = 1$ is impossible because in this case the matrix L_j is elliptic and therefore $r_j = 0$. Hence

$$f_j = -e^{\frac{\gamma_j}{2}}, \;\; |f_j| = e^{\frac{r_j}{2}}\,.$$

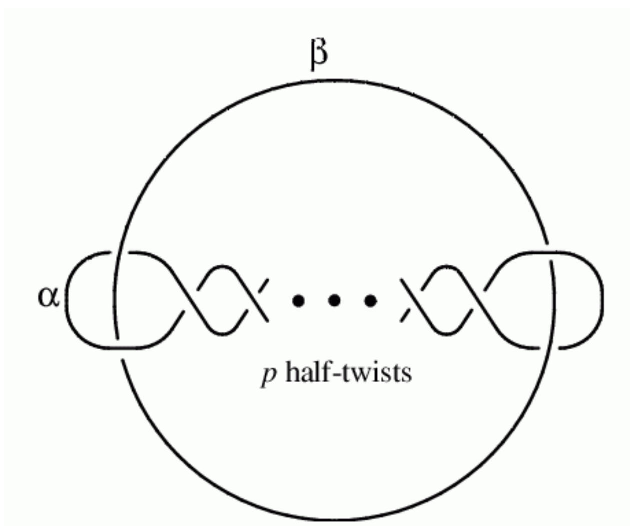


Figure 1. The cone-manifold $W_p(\alpha, \beta)$.

In this paper we consider a family of cone-manifolds whose singular sets are links which are generalizations of the Whitehead link. The link W_p, $p \geq 0$, is the two-component link depicted in Figure 1, where p is the number of half twists of one component. For this reason we will call them *twist links.* It is easy to see that W_0 is the torus link of type $(2,4)$ and W_1 is the Whitehead link. All twist links are two-bridge links, in particular W_p is the two-bridge link with slope $(4p+4)/(2p+1)$, for all $p \geq 0$. They are all hyperbolic, except for W_0.

Denote by $W_p(\alpha,\beta)$ the cone-manifold whose underlying space is the 3-sphere and whose singular set consists of the twist link W_p with cone angles $\alpha = 2\pi/m$ and $\beta = 2\pi/n$ (see Figure 1). It follows from Thurston's theorem that $W_p(\alpha,\beta)$, with $p \neq 0$, admits a hyperbolic structure for all sufficiently small α and β.

By Kojima's rigidity theorem [13] the hyperbolic structure is unique, up to isometry, if $0 \leq \alpha, \beta \leq \pi$.

In our paper we deal only with this range of angles.

Let us investigate the hyperbolic structure of the cone-manifold $W_p(\alpha,\beta)$. Its singular set $\Sigma = \Sigma^1 \cup \Sigma^2$ of consists of two components $\Sigma^1 = \Sigma^1(\alpha)$ and $\Sigma^2 = \Sigma^2(\beta)$ with cone-angles α and β respectively. $W_p(\alpha,\beta)$ defines a nonsingular but incomplete hyperbolic manifold $\mathcal{M} = W_p(\alpha,\beta) - \Sigma$. The fundamental group of the manifold $\mathcal{M}$ has the following presentation

$$\Phi_p = \langle s, t \mid \; sl_s = l_s s \rangle = \langle s, t \mid \; tl_t = l_t t \rangle \,,$$

where s and t (resp. l_s and l_t) are meridians (resp. longitudes) of the components Σ^1 and Σ^2 respectively.

We use the following expression of l_s in terms of s and t:

$$l_s = [s,t]^{\frac{p+1}{2}} [s,t^{-1}]^{\frac{p+1}{2}}, \quad \text{if } \; p \; \text{ is odd,} \tag{2.1.9}$$

$$l_s = s^{-1} [t,s]^{\frac{p}{2}} tst [s^{-1},t^{-1}]^{\frac{p}{2}}, \quad \text{if } \; p \; \text{ is even}\,, \tag{2.1.10}$$

where $[s,t] = sts^{-1}t^{-1}$.

The expressions for l_t can be easily obtained by exchanging s and t in the previous formulae.

Let

$$\hat{h} = \hat{h}_{\alpha,\beta} : \Phi_p \to PSL(2,\mathbb{C})$$

and

$$h = h_{\alpha,\beta} : \Phi_p \to SL(2,\mathbb{C})$$

be holonomy homomorphisms and $\Gamma_{\alpha,\beta} = h_{\alpha,\beta}(\Phi_p)$. The images $\hat{h}_{\alpha,\beta}(s)$ and $\hat{h}_{\alpha,\beta}(t)$ of s and t are rotations in $\mathbb{H}^3$ of angles α and β, respectively. The group $\Gamma_{\alpha,\beta}$ is generated by the two matrices $S = h_{\alpha,\beta}(s)$ and $T = h_{\alpha,\beta}(t)$ with the following properties:

$$\operatorname{tr}(S) = 2\cos\frac{\alpha}{2}, \quad \operatorname{tr}(T) = 2\cos\frac{\beta}{2}, \quad SL_S = L_S S,$$

where $L_S = h_{\alpha,\beta}(l_s)$.

(2.2) Complex distance equation for two-bridge links. The fundamental group of (the exterior of) a link K is generated by two meridians if and only if K is a two-bridge link [1]. Moreover, a two-bridge link is hyperbolic if and only if its slope is different from $p/1$ and $p/(p-1)$ (see [21]).

Proposition (2.2.1). *Let $\Phi = \langle s,t \rangle$ be the fundamental group of a hyperbolic two-bridge link K generated by the two meridians s and t. Let $\Gamma_{\alpha,\beta} = h_{\alpha,\beta}(\Phi)$*

be the image of Φ *under the holonomy homomorphism of the hyperbolic cone-manifold* $K(\alpha,\beta)$. *Then, up to conjugation in* $SL(2,\mathbb{C})$, *the generators* $S = h_{\alpha,\beta}(s)$ *and* $T = h_{\alpha,\beta}(t)$ *of* $\Gamma_{\alpha,\beta}$ *can be chosen in such a way that*

$$(2.2.2)\quad S = \begin{pmatrix} \cos\frac{\alpha}{2} & i\,e^{\frac{\rho}{2}}\sin\frac{\alpha}{2} \\ i\,e^{-\frac{\rho}{2}}\sin\frac{\alpha}{2} & \cos\frac{\alpha}{2} \end{pmatrix}, \quad T = \begin{pmatrix} \cos\frac{\beta}{2} & i\,e^{-\frac{\rho}{2}}\sin\frac{\beta}{2} \\ i\,e^{\frac{\rho}{2}}\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix},$$

where ρ *is the complex distance between the axis of* S *and* T.

Proof. After a suitable conjugation in the group $SL(2,\mathbb{C})$, one can assume that the oriented axes of the elliptic elements S and T are $\lambda_S = [-e^{\frac{\rho}{2}}, e^{\frac{\rho}{2}}]$ and $\lambda_T = [-e^{-\frac{\rho}{2}}, e^{-\frac{\rho}{2}}]$. Since $\mathrm{tr}\,(S) = 2\cos\frac{\alpha}{2}$ and $\mathrm{tr}\,(T) = 2\cos\frac{\beta}{2}$, the matrices S and T are given by (2.2.2). Check that ρ coincides with the complex distance $\rho(S,T)$ between λ_S and λ_T. The line matrices $\widetilde{S}$ and $\widetilde{T}$, corresponding to these axes, can be obtained by (2.1.5). Since $\delta(S) = i\,\alpha$ and $\delta(T) = i\,\beta$, we have $\widetilde{S} = \begin{pmatrix} 0 & -ie^{\frac{\rho}{2}} \\ -ie^{-\frac{\rho}{2}} & 0 \end{pmatrix}$ and $\widetilde{T} = \begin{pmatrix} 0 & -ie^{-\frac{\rho}{2}} \\ -ie^{\frac{\rho}{2}} & 0 \end{pmatrix}$ respectively. By [4, p. 68] we get $\cosh\rho(S,T) = -\frac{1}{2}\,\mathrm{tr}\,(\widetilde{S}\widetilde{T}) = \cosh\rho$. □

The following two propositions can be obtained by direct calculation from the above statement.

PROPOSITION (2.2.3). *Let*

$$\Phi_2 = \langle s,t : sl = ls,\ l = s^{-1}tst^{-1}s^{-1}tsts^{-1}t^{-1}st\rangle$$

be the fundamental group of the two-bridge link W_2 *with slope* $12/5$ *and* $\Gamma_{\alpha,\beta} = h_{\alpha,\beta}(\Phi_2) = \langle S,T\rangle$ *be the image of* Φ_2 *under the holonomy homomorphism of the hyperbolic cone-manifold* $W_2(\alpha,\beta)$. *Denote by* $\rho = \rho(S,T)$ *the complex distance between the axes of* $S = h_{\alpha,\beta}(s)$ *and* $T = h_{\alpha,\beta}(t)$. *Then* $u = \cosh\rho$ *is a non-real root of the complex distance equation*

$$(2.2.4)\quad 4z^3 - 4abz^2 + (3a^2b^2 + 3a^2 + 3b^2 - 1)z - ab(a^2b^2 + a^2 + b^2 - 3) = 0,$$

where $a = \cot\frac{\alpha}{2}$ *and* $b = \cot\frac{\beta}{2}$.

Proof. Denote by $L = S^{-1}TST^{-1}S^{-1}TSTS^{-1}T^{-1}ST$ the image of the longitude l under the holonomy homomorphism $h = h_{\alpha,\beta} : \Phi_2 \to SL(2,\mathbb{C})$. Then we have $SL = LS$.

Let N be a line matrix corresponding to the common normal to the axes of S and T. If S and T are represented in the form (2.2.2) then one can take $N = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$. It is not difficult to verify that $NSN^{-1} = S^{-1}$ and $NTN^{-1} = T^{-1}$.

To complete the proof, we need the following lemma, which gives simple criteria for matrices S and L to be permutable.

LEMMA (2.2.5). *The following conditions are equivalent: (i)* $SL = LS$; *(ii)* $NLN^{-1} = L^{-1}$; *(iii)* $\mathrm{tr}\,(NL) = 0$.

Proof. First we show that (i) and (ii) are equivalent. Indeed, since $L = S^{-1}TST^{-1}S^{-1}TSTS^{-1}T^{-1}ST$ we have

$$NLN^{-1} = ST^{-1}S^{-1}TST^{-1}S^{-1}T^{-1}STS^{-1}T^{-1} = SL^{-1}S^{-1}.$$

Hence (ii) holds if and only if S and L^{-1} commute. The last property is equivalent to (i). Because of $N^2 = -I$ the condition (ii) can be rewritten in the form $NLNL = -I$; this is equivalent to (iii). □

By this lemma and direct calculation we have

$$\operatorname{tr}(NL) = \frac{-4i\sinh\rho}{(1+a^2)^3(1+b^2)^3} \cdot (4u^2 + a^2b^2 + a^2 + b^2 - 3) \cdot$$
$$\cdot (4u^3 - 4abu^2 + (3a^2b^2 + 3a^2 + 3b^2 - 1)u - ab(a^2b^2 + a^2 + b^2 - 3)) = 0,$$

where $u = \cosh\rho$.

Now we have to show that u is a non-real root of (2.2.4). Since $\Gamma_{\alpha,\beta}$ is the holonomy group of a hyperbolic cone-manifold, it is non-elementary[1] and is not conjugate to a subgroup of $SL(2,\mathbb{R})$ [8].

If $\sinh\rho = 0$ then the axes S and T coincide, and the group $\Gamma_{\alpha,\beta}$ is elementary.

If u is a root of equation

$$4u^2 + a^2b^2 + a^2 + b^2 - 3 = 0$$

then by the equality

$$\operatorname{tr} L - 2 = -\frac{4(a^2+u^2)(4u^2 + a^2b^2 + a^2 + b^2 - 3)^2}{(a^2+1)^3(b^2+1)^3}$$

we have $\operatorname{tr} L = 2$. From (2.1.8) we obtain

$$\cosh\frac{\gamma_S}{2} = -\frac{1}{2}\operatorname{tr}(L) = -1.$$

Hence $\gamma_S = r_S + i\varphi_S = 2\pi i$ and the real length r_S of the link component Σ_1 is equal to zero, which is a contradiction.

Suppose that $u = \cosh\rho$ is a real root. Let

$$R(z_1, z_2, z_3, z_4) = \frac{(z_3 - z_1)(z_4 - z_2)}{(z_3 - z_2)(z_4 - z_1)}$$

be the cross ratio of the four points $z_1, z_2, z_3, z_4 \in \widehat{\mathbb{C}}$. Then

$$R(-e^{\frac{\rho}{2}}, e^{\frac{\rho}{2}}, -e^{-\frac{\rho}{2}}, e^{-\frac{\rho}{2}}) = (\cosh\rho - 1)/(\cosh\rho + 1) \in \mathbb{R}\cup\{\infty\}\,.$$

We have that the axes $[-e^{\frac{\rho}{2}}, e^{\frac{\rho}{2}}]$ and $[-e^{-\frac{\rho}{2}}, e^{-\frac{\rho}{2}}]$ of S and T lie in a common plane. If the axes intersect then the group $\Gamma_{\alpha,\beta} = \langle S, T\rangle$ has a fixed point and is elementary. If they do not intersect, $\Gamma_{\alpha,\beta}$ is conjugate to a subgroup of $SL(2,\mathbb{R})$.

Therefore, we have shown that u is a non-real root of (2.2.4) and the proof of Proposition (2.2.3) is complete. □

The next proposition can be proved by similar arguments.

PROPOSITION (2.2.6). *Let*

$$\Phi_3 = \langle s, t:\ sl = ls,\ l = sts^{-1}t^{-1}sts^{-1}t^{-1}st^{-1}s^{-1}tst^{-1}s^{-1}t\rangle$$

be the fundamental group of the two-bridge link W_3 with the slope $16/7$ and $\Gamma_{\alpha,\beta} = h_{\alpha,\beta}(\Phi_3) = \langle S, T\rangle$ the image of Φ_3 under the holonomy homomorphism of a hyperbolic cone-manifold $W_3(\alpha,\beta)$ generated by $S = h_{\alpha,\beta}(s)$ and $T = h_{\alpha,\beta}(t)$.

[1] A subgroup G of $SL(2,\mathbb{C})$ is called *elementary* if it has a finite orbit in $\mathbb{H}^3\cup\widehat{\mathbb{C}}$.

Denote by $\rho = \rho(S, T)$ *the complex distance between the axes of* S *and* T*. Then* $u = \cosh\rho$ *is a non-real root of the complex distance equation*

$$0 = 8u^5 + 8abu^4 + 8(a^2b^2 + a^2 + b^2 - 1)u^3 + 4ab(a^2b^2 + a^2 + b^2 - 3)u^2 +$$

$$(a^4b^4 + 2a^4b^2 + 2a^2b^4 - 4a^2b^2 + a^4 + b^4 - 6a^2 - 6b^2 + 1)u - 4ab(a^2b^2 + a^2 + b^2 - 1),$$

where $a = \cot\frac{\alpha}{2}$ *and* $b = \cot\frac{\beta}{2}$*.*

(2.3) Tangent, Sine and Cosine rules. If we set $z = \mathrm{tr}\,(S^{-1}T)$ then, from the presentation in Proposition (2.2.1), we have

$$z = 2(\cos\frac{\alpha}{2}\cos\frac{\beta}{2} + u\sin\frac{\alpha}{2}\sin\frac{\beta}{2}),$$

where $u = \cosh\rho$.

The algebraic equation for z and its behaviour was considered in a number of papers (see [3], [5], [8] and others) devoted to $PSL(2,\mathbb{C})$-representation of two-generator groups.

In general, the equation for u (as well as for z) is very complicated, even for twist links. In spite of this, since $u = \cosh\rho$ has a very clear geometric sense, we are able to produce some general results for twist links without calculating u.

PROPOSITION (2.3.1). *Let* $W_p(\alpha, \beta)$ *be a hyperbolic twist link cone-manifold. Denote by* $S = h_{\alpha,\beta}(s)$ *and* $T = h_{\alpha,\beta}(t)$ *the images of the generators of the group* $\Phi_p = \langle s, t \mid sl_s = l_s s\rangle$ *under the holonomy homomorphism* $h_{\alpha,\beta} : \Phi_p \to SL(2,\mathbb{C})$*. Set* $u = \cosh\rho$*, where* ρ *is the complex distance between the axes of* S *and* T*, such that* $\Im u > 0$*. Moreover, denote by* γ_α *and* γ_β *the complex lengths of the singular components of* $W_p(\alpha, \beta)$ *with cone-angles* α *and* β *respectively. Then*

$$u = i\,\cot\frac{\alpha}{2}\coth\frac{\gamma_\beta}{4} = i\,\cot\frac{\beta}{2}\coth\frac{\gamma_\alpha}{4}.$$

Proof. To prove the statement we need to calculate the complex distance between axes of elliptic elements S and T in two ways. By definition, $L_S = h_{\alpha,\beta}(l_s)$ and $L_T = h_{\alpha,\beta}(l_t)$, where l_s and l_t are the longitudes of the singular components of $W_p(\alpha, \beta)$ with cone angles α and β, respectively.

First of all, we fix an orientation on the axes of S and T by the following line matrices

$$\widetilde{S} = \frac{S - S^{-1}}{2\,i\,\sinh\frac{i\,\alpha}{2}}, \qquad \widetilde{T} = \frac{T - T^{-1}}{2\,i\,\sinh\frac{i\,\beta}{2}}.$$

Then the complex distance $\rho(S, T)$ between the oriented axes of S and T is defined by (2.1.3):

$$\cosh\rho(S, T) = -\frac{1}{2}\,\mathrm{tr}\,(\widetilde{S}\widetilde{T}).$$

Using (2.1.5) we define the line matrices for L_S and L_T as

$$\widetilde{L_S} = \frac{L_S - L_S^{-1}}{2i\sinh\frac{\gamma_\alpha}{2}}, \qquad \widetilde{L_T} = \frac{L_T - L_T^{-1}}{2i\sinh\frac{\gamma_\beta}{2}}.$$

To continue the proof, we need two lemmas:

LEMMA (2.3.2). *For every* S, T *we have* $\widetilde{S} = -\widetilde{L_S}$ *and* $\widetilde{T} = -\widetilde{L_T}$*.*

Proof. Up to conjugation in $SL(2,\mathbb{C})$, we can assume that S is given by

$$S = \begin{pmatrix} e^{\frac{i\alpha}{2}} & 0 \\ 0 & e^{-\frac{i\alpha}{2}} \end{pmatrix}.$$

Note that L_S is a loxodromic element, with displacement γ_α, and commutes with S. Since $\widetilde{L_S^{-1}} = \widetilde{L_S}$, we can assume that

$$L_S = \begin{pmatrix} \pm e^{\frac{\gamma_\alpha}{2}} & 0 \\ 0 & \pm e^{-\frac{\gamma_\alpha}{2}} \end{pmatrix}$$

By convention (see formula (2.1.8)) we have

$$\operatorname{tr}(L_S) = -2\cosh\frac{\gamma_\alpha}{2}.$$

Hence

$$L_S = \begin{pmatrix} -e^{\frac{\gamma_\alpha}{2}} & 0 \\ 0 & -e^{-\frac{\gamma_\alpha}{2}} \end{pmatrix}$$

and we obtain

$$\widetilde{L}_S = \frac{L_S - L_S^{-1}}{2i\sinh\frac{\gamma_\alpha}{2}} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

and

$$\widetilde{S} = \frac{S - S^{-1}}{2\,i\sinh\frac{i\,\alpha}{2}} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}.$$

□

Lemma (2.3.3). *For every S,T we have $\operatorname{tr}(S) = \operatorname{tr}(S^{-1}L_T)$ and $\operatorname{tr}(T) = \operatorname{tr}(T^{-1}L_S)$.*

Proof. To prove $\operatorname{tr}(T) = \operatorname{tr}(T^{-1}L_S)$ it is enough to show that $T^{-1}L_S$ is conjugate to either T or T^{-1} in the group $\Gamma_{\alpha,\beta}$. If p is odd, we have from (2.1.9):

$$T^{-1}L_S = T^{-1}[S,T]^{\frac{p+1}{2}}\,[S,T^{-1}]^{\frac{p+1}{2}} = [T^{-1},S]^{\frac{p+1}{2}}\,T^{-1}[T^{-1},S]^{-\frac{p+1}{2}}.$$

If p is even, we have from (2.1.10):

$$T^{-1}L_S = T^{-1}S^{-1}[T,S]^{\frac{p}{2}}\,TST\,[S^{-1},T^{-1}]^{\frac{p}{2}} = T^{-1}S^{-1}[T,S]^{\frac{p}{2}}\,T\,[T,S]^{-\frac{p}{2}}\,ST.$$

The equality $\operatorname{tr}(S) = \operatorname{tr}(S^{-1}L_T)$ can be obtained in a similar way. □

To complete the proof of Proposition (2.3.1), we note that $\operatorname{tr}(XY) = \operatorname{tr}(X)\operatorname{tr}(Y) - \operatorname{tr}(X^{-1}Y)$, $\operatorname{tr}(X^{-1}) = \operatorname{tr}(X)$ and $\operatorname{tr}(XY) = \operatorname{tr}(X^{-1}Y^{-1})$ holds for all $X,Y \in SL(2,\mathbb{C})$. By Lemma (2.3.2), Lemma (2.3.3) and formulae $\operatorname{tr}(S) = 2\cos\frac{\alpha}{2}$, $\operatorname{tr}(L_S) = -2\cosh\frac{\gamma_\alpha}{2}$, we have

$$\cosh\rho(S,T) = -\frac{1}{2}\operatorname{tr}(\widetilde{S}\widetilde{T}) = \frac{1}{2}\operatorname{tr}(\widetilde{S}\widetilde{L_T}) =$$

$$= \frac{1}{2}\operatorname{tr}\left(\frac{S - S^{-1}}{2\sin\frac{\alpha}{2}}\,\frac{L_T - L_T^{-1}}{2i\sinh\frac{\gamma_\beta}{2}}\right) = \frac{\operatorname{tr}(SL_T - S^{-1}L_T - SL_T^{-1} + S^{-1}L_T^{-1})}{8i\sin\frac{\alpha}{2}\sinh\frac{\gamma_\beta}{2}} =$$

$$= \frac{2(\operatorname{tr}(SL_T) - \operatorname{tr}(S^{-1}L_T))}{8i\sin\frac{\alpha}{2}\sinh\frac{\gamma_\beta}{2}} = \frac{\operatorname{tr}(S)\operatorname{tr}(L_T) - 2\operatorname{tr}(S^{-1}L_T)}{4i\sin\frac{\alpha}{2}\sinh\frac{\gamma_\beta}{2}} =$$

$$= \frac{\mathrm{tr}\,(S)\mathrm{tr}\,(L_T) - 2\mathrm{tr}\,(S)}{4i\sin\frac{\alpha}{2}\sinh\frac{\gamma_\beta}{2}} = \frac{\mathrm{tr}\,(S)(2 - \mathrm{tr}\,(L_T))}{-4i\sin\frac{\alpha}{2}\sinh\frac{\gamma_\beta}{2}}$$

$$= \frac{2\cos\frac{\alpha}{2}(2 + 2\cosh\frac{\gamma_\beta}{2})}{-4i\sin\frac{\alpha}{2}\sinh\frac{\gamma_\beta}{2}} = i\cot\frac{\alpha}{2}\coth\frac{\gamma_\beta}{4}.$$

Since $\cosh\rho(S,T) = \cosh\rho(T,S) = u$ the statement follows. □

As an immediate consequence of the previous proposition, we have the following result.

THEOREM (2.3.4). (The Tangent Rule). *Suppose that $W_p(\alpha,\beta)$ is a hyperbolic cone-manifold. Denote by γ_α and γ_β complex lengths of the singular geodesics of $W_p(\alpha,\beta)$ with cone angles α and β respectively. Then*

$$\frac{\tanh\frac{\gamma_\alpha}{4}}{\tanh\frac{\gamma_\beta}{4}} = \frac{\tan\frac{\alpha}{2}}{\tan\frac{\beta}{2}}.$$

The following two theorems are consequences of the Tangent Rule.

THEOREM (2.3.5). (The Sine Rule). *Let $\gamma_\alpha = r_\alpha + i\,\varphi_\alpha$ and $\gamma_\beta = r_\beta + i\,\varphi_\beta$ be the complex lengths of the singular geodesics of a hyperbolic cone-manifold $W_p(\alpha,\beta)$ with cone angle α and β respectively. Then*

$$\frac{\sin\frac{\varphi_\alpha}{2}}{\sinh\frac{r_\alpha}{2}} = \frac{\sin\frac{\varphi_\beta}{2}}{\sinh\frac{r_\beta}{2}}.$$

Proof. By the Tangent Rule we have

$$\frac{\tanh\frac{\gamma_\alpha}{4}}{a} = \frac{\tanh\frac{\gamma_\beta}{4}}{b},$$

where $a = \tan\frac{\alpha}{2}$ and $B = \tan\frac{\beta}{2}$ are real numbers. Hence

$$\frac{\Re(\tanh\frac{\gamma_\alpha}{4})}{a} = \frac{\Re(\tanh\frac{\gamma_\beta}{4})}{b},$$

and

$$\frac{\Im(\tanh\frac{\gamma_\alpha}{4})}{a} = \frac{\Im(\tanh\frac{\gamma_\beta}{4})}{b}.$$

Dividing one equation by the other we obtain

$$\frac{\Re(\tanh\frac{\gamma_\alpha}{4})}{\Im(\tanh\frac{\gamma_\alpha}{4})} = \frac{\Re(\tanh\frac{\gamma_\beta}{4})}{\Im(\tanh\frac{\gamma_\beta}{4})}.$$

By direct calculations we have

$$\Re(\tanh\frac{\gamma_\alpha}{4}) = \frac{1}{2}(\tanh\frac{\gamma_\alpha}{4} + \tanh\frac{\bar{\gamma}_\alpha}{4}) = \frac{\sinh\frac{r_\alpha}{2}}{\cosh\frac{r_\alpha}{2} + \cos\frac{\varphi_\alpha}{2}}$$

and

$$\Im(\tanh\frac{\gamma_\alpha}{4}) = \frac{1}{2i}(\tanh\frac{\gamma_\alpha}{4} - \tanh\frac{\bar{\gamma}_\alpha}{4}) = \frac{\sin\frac{\varphi_\alpha}{2}}{\cosh\frac{r_\alpha}{2} + \cos\frac{\varphi_\alpha}{2}}.$$

Since $r_\alpha > 0$, we have $\cosh\dfrac{r_\alpha}{2} > 1$. Therefore $\cosh\dfrac{r_\alpha}{2} + \cos\dfrac{\varphi_\alpha}{2} > 0$ and the result follows. □

Theorem (2.3.6). (The Cosine Rule). *Let $\gamma_\alpha = r_\alpha + i\,\varphi_\alpha$ and $\gamma_\beta = r_\beta + i\,\varphi_\beta$ be the complex lengths of the singular geodesics of a hyperbolic cone-manifold $W_p(\alpha,\beta)$ with cone angle α and β respectively. Then*

$$\frac{\cos\frac{\varphi_\alpha}{2}\cosh\frac{r_\beta}{2} - \cos\frac{\varphi_\beta}{2}\cosh\frac{r_\alpha}{2}}{\cosh\frac{r_\alpha}{2}\cosh\frac{r_\beta}{2} - \cos\frac{\varphi_\alpha}{2}\cos\frac{\varphi_\beta}{2}} = \frac{\cos\alpha - \cos\beta}{1-\cos\alpha\cos\beta}.$$

Proof. By the Tangent Rule

$$\frac{\tanh\frac{\gamma_\alpha}{4}\tanh\frac{\bar\gamma_\alpha}{4}}{a^2} = \frac{\tanh\frac{\gamma_\beta}{4}\tanh\frac{\bar\gamma_\beta}{4}}{b^2},$$

where $a = \tan\frac{\alpha}{2}$ and $b = \tan\frac{\beta}{2}$. Hence

$$\frac{1+\cos\alpha}{1-\cos\alpha}\,\frac{\cosh\frac{r_\alpha}{2} - \cos\frac{\varphi_\alpha}{2}}{\cosh\frac{r_\alpha}{2} + \cos\frac{\varphi_\alpha}{2}} = \frac{1+\cos\beta}{1-\cos\beta}\,\frac{\cosh\frac{r_\beta}{2} - \cos\frac{\varphi_\beta}{2}}{\cosh\frac{r_\beta}{2} + \cos\frac{\varphi_\beta}{2}}.$$

Set

$$p = \cos\alpha, \quad q = \cos\beta, \quad p' = \frac{\cos\frac{\varphi_\alpha}{2}}{\cosh\frac{r_\alpha}{2}}, \quad q' = \frac{\cos\frac{\varphi_\beta}{2}}{\cosh\frac{r_\beta}{2}}$$

and rewrite the above equation in the form

$$\frac{1+p}{1-p}\,\frac{1-p'}{1+p'} = \frac{1+q}{1-q}\,\frac{1-q'}{1+q'},$$

or, equivalently, as

$$\log\frac{1+p}{1-p} + \log\frac{1-p'}{1+p'} = \log\frac{1+q}{1-q} + \log\frac{1-q'}{1+q'}.$$

Since $\operatorname{arctanh} p = \dfrac{1}{2}\log\dfrac{1+p}{1-p}$ we have

$$\operatorname{arctanh} p - \operatorname{arctanh} p' = \operatorname{arctanh} q - \operatorname{arctanh} q'.$$

and

$$\operatorname{arctanh} p - \operatorname{arctanh} q = \operatorname{arctanh} p' - \operatorname{arctanh} q'.$$

Hence

$$\frac{p-q}{1-pq} = \frac{p'-q'}{1-p'q'}$$

and, after substituting the expressions for p, q, p', q' in the last formula, we obtain the desired identity. □

We remark that, in the case of Whitehead link cone-manifolds, Tangent and Sine rules were obtained in [14].

3. Explicit volume calculation for twist link cone-manifolds

(3.1) The Schläfli formula. In this section we will obtain explicit formulae for the volume of some special cone-manifolds in the hyperbolic and spherical geometries. In the case of complete hyperbolic structure on the simplest knot and link complements such formulas, in terms of the Lobachevsky function, are well-known and widely represented in [21]. In general, a hyperbolic cone-manifold can be obtained by completion of a non-complete hyperbolic structure on a suitable knot or link complement. If the cone-manifold is compact, explicit formulas are only known in a few cases [9], [10], [11], [15], [16], [17], [18], [19]. In all these cases the starting point for the volume calculation is the Schläfli formula (see, for example [11]).

THEOREM (3.1.1). (The Schläfli volume formula). *Suppose that C_t is a smooth 1-parameter family of (curvature K) cone-manifold structures on an n-manifold, with singular locus Σ of a fixed topological type. Then the derivative of volume of C_t satisfies*

$$(n-1)K dV(C_t) = \sum_\sigma V_{n-2}(\sigma)\, d\theta(\sigma)$$

where the sum is over all the components σ of the singular locus Σ, and $\theta(\sigma)$ is the cone angle along σ.

In the present paper we will deal mostly with three-dimensional cone-manifold structures of negative constant curvature $K = -1$. The Schläfli formula in this case reduces to

$$dV = -\frac{1}{2}\sum_i r_i d\theta_i,$$

where the sum is taken over all the components of the singular set Σ with lengths r_i and cone angles θ_i.

Our aim is to obtain the volume formulas for twist link hyperbolic cone-manifolds $W_2(\alpha,\beta)$. We note that the volume formula for $W_1(\alpha,\beta)$ were obtained earlier in [16] and [19].

PROPOSITION (3.1.2). *Let $W_2(\alpha,\beta)$ be a hyperbolic cone-manifold and r_α, r_β the lengths of its singular components, with cone angles α and β respectively. If $a = \cot\frac{\alpha}{2}$ and $b = \cot\frac{\beta}{2}$, then*

$$r_\alpha = 2i\ \arctan\frac{a}{\zeta} - 2i\ \arctan\frac{a}{\overline{\zeta}}, \tag{3.1.3}$$

$$r_\beta = 2i\ \arctan\frac{b}{\zeta} - 2i\ \arctan\frac{b}{\overline{\zeta}}, \tag{3.1.4}$$

where ζ is a root of the equation

$$4(z^2+a^2)(z^2+b^2) - (1+a^2)(1+b^2)(z-z^2)^2 = 0, \tag{3.1.5}$$

with $\Im(\zeta) > 0$.

Proof. By Proposition (2.3.1) we have

$$i\,b\coth\frac{\gamma_\alpha}{4} = i\,a\coth\frac{\gamma_\beta}{4} = u, \tag{3.1.6}$$

where $u = \cosh\rho$ and ρ is a complex distance between the axes of S and T, chosen so that $\Im u > 0$. By Proposition (2.2.3), u is a root of the cubic equation

$$4z^3 - 4abz^2 + (3a^2b^2 + 3a^2 + 3b^2 - 1)z - ab(a^2b^2 + a^2 + b^2 - 3) = 0\,.$$

From (3.1.6), for a suitable choice of analytical branches,

$$r_\alpha = \frac{\gamma_\alpha}{2} + \frac{\overline{\gamma}_\alpha}{2} = 2i\ \arctan\frac{\overline{u}}{b} - 2i\ \arctan\frac{u}{b} = 2i\ \arctan\frac{a}{\zeta} - 2i\ \arctan\frac{a}{\overline{\zeta}}\,,$$

where $\zeta = ab/\overline{u}$, $\Im(\zeta) > 0$, satisfies the equation

$$Q(z) = (a^2b^2 + a^2 + b^2 - 3)z^3 - (3a^2b^2 + 3a^2 + 3b^2 - 1)z^2 + 4a^2b^2 z - 4a^2b^2 = 0\,.$$

To finish the proof we note that

$$(z+1)Q(z) = -4(z^2+a^2)(z^2+b^2) + (1+a^2)(1+b^2)(z-z^2)^2.$$

□

In the next section we will apply this result to calculate the volume of $W_2(\alpha,\beta)$ via the Schläfli formula.

We remark that formulae (3.1.3) and (3.1.4), as a consequence of the Tangent Rule, also hold for all twist links W_p, with $\zeta = ab/\bar{u}$, where $u = \cosh\rho$.

For example, an analog for the algebraic equation (3.1.5), in the case of twist link W_3, can easily be obtained from Proposition (2.2.6). But in this case the equation becomes too complicated, and we are not able to explicitly find the integrand in the Schläfli formula.

(3.2) Volume of twist link cone-manifolds. The case of the Whitehead link cone-manifolds $W_1(\alpha,\beta)$ has already been solved (see [16] and [19]).

THEOREM (3.2.1). [16, 19] *Let $W_1(\alpha,\beta)$ be a hyperbolic Whitehead link cone-manifold with cone angles α and β. Then the volume of $W_1(\alpha,\beta)$ is given by the formula*

$$\operatorname{Vol} W_1(\alpha,\beta) = i\int_{\bar{\zeta}}^{\zeta} \log\left[\frac{2(z^2+a^2)(z^2+b^2)}{(1+a^2)(1+b^2)(z^2-z^3)}\right]\frac{dz}{z^2-1}.$$

where $a = \cot\frac{\alpha}{2}$, $b = \cot\frac{\beta}{2}$ and ζ is a non-real root, with $\Im(\zeta) > 0$, of the equation

$$2(z^2+a^2)(z^2+b^2) - (1+a^2)(1+b^2)(z^2-z^3) = 0.$$

The main result of this section is the following.

THEOREM (3.2.2). *Let $W_2(\alpha,\beta)$ be a hyperbolic twist link cone-manifold with cone angles α and β. Then the volume of $W_2(\alpha,\beta)$ is given by the formula*

$$\operatorname{Vol} W_2(\alpha,\beta) = i\int_{\bar{\zeta}}^{\zeta} \log\left[\frac{4(z^2+a^2)(z^2+b^2)}{(1+a^2)(1+b^2)(z-z^2)^2}\right]\frac{dz}{z^2-1}. \tag{3.2.3}$$

where $a = \cot\frac{\alpha}{2}$, $b = \cot\frac{\beta}{2}$ and ζ is a non-real root, with $\Im(\zeta) > 0$, of the equation

$$4(z^2+a^2)(z^2+b^2) - (1+a^2)(1+b^2)(z-z^2)^2 = 0. \tag{3.2.4}$$

Proof. Denote by $V = \operatorname{Vol} W_2(\alpha,\beta)$ the hyperbolic volume of $W_2(\alpha,\beta)$. Then by virtue of the Schläfli formula we have

$$\frac{\partial V}{\partial \alpha} = -\frac{r_\alpha}{2}, \quad \frac{\partial V}{\partial \beta} = -\frac{r_\beta}{2}, \tag{3.2.5}$$

where r_α and r_α are the lengths of the singular geodesics having cone angles α and β respectively.

We note that for $\alpha = \beta$ and $\Im(\zeta) \to 0$ the geometrical limit of the cone-manifold $W_2(\alpha,\,\alpha)$ is a Euclidean cone-manifold $W_2(\alpha_0,\,\alpha_0)$, where $\alpha_0 = 2.7243...$ $< \pi$. (See Example 1 in Section 3.3 below). Hence, by Theorem 7.1.2 of [13], we have

$$V \to 0 \;\text{ as }\; \alpha = \beta \;\text{ and }\; \Im(\zeta) \to 0. \tag{3.2.6}$$

We set

$$W = \int_{\overline{\zeta}}^{\zeta} F(z,a,b)\,dz,$$

where

$$F(z,a,b) = \frac{i}{z^2-1}\,\log\frac{4(z^2+a^2)(z^2+b^2)}{(1+a^2)(1+b^2)(z-z^2)^2}.$$

Now we show that W satisfies conditions (3.2.5) and (3.2.6). So $W = V$ and the theorem follows.

By the Leibniz formula we have

$$\frac{\partial W}{\partial \alpha} = F(\zeta,a,b)\frac{\partial \zeta}{\partial \alpha} - F(\overline{\zeta},a,b)\frac{\partial \overline{\zeta}}{\partial \alpha} + \int_{\overline{\zeta}}^{\zeta} \frac{\partial F(z,a,b)}{\partial a}\,\frac{\partial a}{\partial \alpha}\,dz \tag{3.2.7}$$

We note that $F(\overline{\zeta},a,b) = F(\zeta,a,b) = 0$ if $\overline{\zeta},\zeta,a$ and b are the same as in the statement of the theorem. Moreover, since $\alpha = 2\,\mathrm{arccot}\,a$ we have $\dfrac{\partial a}{\partial \alpha} = -\dfrac{1+a^2}{2}$ and

$$\frac{\partial F(z,a,b)}{\partial a}\,\frac{\partial a}{\partial \alpha} = -\frac{ia}{z^2+a^2}.$$

Hence, by Proposition (3.1.2), we obtain from (3.2.7)

$$\frac{\partial W}{\partial \alpha} = -ia\int_{\overline{\zeta}}^{\zeta}\frac{dz}{z^2+a^2} = -i\,\arctan\frac{a}{\zeta} + i\,\arctan\frac{a}{\overline{\zeta}} = -\frac{r_\alpha}{2}.$$

The equation $\dfrac{\partial W}{\partial \beta} = -\dfrac{r_\beta}{2}$ can be obtained in the same way. The boundary condition (3.2.6) for the function W follows from the integral formula. □

(3.3) Particular cases and examples.

1. Case $\alpha = \beta$. In this case, Equation (3.2.4) splits into two quadratic equations:

$$(1+a^2)(z-z^2) + 2(z^2+a^2) = 0$$

and

$$(1+a^2)(z-z^2) - 2(z^2+a^2) = 0.$$

The first has two real roots $z = -1$ and $z = 2a^2/(a^2-1)$. The second has two non-real roots

$$z_{1,2} = \frac{1+a^2 \pm \sqrt{1-22a^2-7a^4}}{2(3+a^2)}.$$

By [10], $\Delta = 1-22a^2-7a^4$ is <0 in the hyperbolic case, $=0$ in the Euclidean case and >0 in the spherical case. In the Euclidean case we obtain $a^2 = \cot^2(\alpha_0/2) = (\sqrt{128}-11)/7 = 0.0448...$ and $a = a_0 = \cot(\alpha_0/2) = 0.2116...$. So the cone-manifold is hyperbolic for $0 \le \alpha < \alpha_0 = 2.7243...$ and is Euclidean for $\alpha = \alpha_0$.

From (3.2.3) we have

$$\mathrm{Vol}\,W_2(\alpha,\alpha) = i\int_{z_1}^{z_2}\log\left[\frac{2(z^2+a^2)}{(z-z^2)(1+a^2)}\right]^2\frac{dz}{z^2-1}.$$

By differentiation with respect to a and then by integration with respect to z we obtain

$$\operatorname{Vol} W_2(\alpha,\alpha) = 4\int_{a_0}^{a} \operatorname{arctanh} \frac{\sqrt{7t^4+22t^2-1}}{t(5+t^2)}\frac{dt}{t^2+1}.$$

Since the integrand is purely imaginary for $0 \le t < a_0$ we are able to compute the volume in the more convenient way

$$\operatorname{Vol} W_2(\alpha,\alpha) = 4\Re\int_{0}^{a} \operatorname{arctanh} \frac{\sqrt{7t^4+22t^2-1}}{t(5+t^2)}\frac{dt}{t^2+1},$$

where $a = \cot\frac{\alpha}{2}$.

2. Case $\alpha=\beta=\pi/2$. In this case equation (3.2.4) becomes

$$(z+1)(z^2-z+2)=0.$$

Hence, the non-real roots are

$$z_{1,2} = \frac{1\pm i\sqrt{7}}{2}$$

and

$$\operatorname{Vol} W_2(\pi/2,\pi/2) = 2i\int_{\frac{1-i\sqrt{7}}{4}}^{\frac{1+i\sqrt{7}}{4}} \log\frac{z^2+1}{z-z^2}\frac{dz}{z^2-1} = 2.6667...$$

3. Case $\alpha=\beta=0$. Recall that $W_2(0,0)$ is the complete hyperbolic manifold $S^3 \setminus W_2$. By arguments similar to those of the previous case, we obtain

$$\operatorname{Vol} W_2(0,0) = 2i\int_{\frac{1-i\sqrt{7}}{2}}^{\frac{1+i\sqrt{7}}{2}} \log\frac{2}{z-z^2}\frac{dz}{z^2-1} = 5.3334...$$

Note that $\operatorname{Vol} W_2(0,0) = 2\operatorname{Vol} W_2(\pi/2,\pi/2)$.

4. Case $\alpha=0$, $\beta=\pi/3$. In this case equation (3.2.4) reduces to

$$(1+z)(3-3z+3z^2-z^3)=0.$$

Hence, the non-real roots are

$$z_{1,2} = 1-\frac{1\pm i\sqrt{3}}{\sqrt[3]{4}}$$

and

$$\operatorname{Vol} W_2(0,\pi/3) = i\int_{1-\frac{1+i\sqrt{3}}{\sqrt[3]{4}}}^{1-\frac{1-i\sqrt{3}}{\sqrt[3]{4}}} \log\frac{z^2+3}{(z-z^2)^2}\frac{dz}{z^2-1} = 4.6165...$$

The results of the above numerical calculation coincide with the corresponding results obtained by Weeks's SnapPea program [22].

Acknowledgement

Work performed under the auspices of the G.N.S.A.G.A. of I.N.d.A.M. (Italy) and the University of Bologna, funds for selected research topics.
The authors are thankful to referees for helpful remarks and suggestions.

Received June 30, 2003

Final version received February 04, 2004

Dmitriy Derevnin
Novosibirsk State University
Novosibirsk 630090
Russia
derevnin@mail.ru

Alexander D. Mednykh
Sobolev Institute of Mathematics
Novosibirsk 630090
Russia
mednykh@math.nsc.ru

Michele Mulazzani
Department of Mathematics and C.I.R.A.M.
University of Bologna
Italy
mulazza@dm.unibo.it

References

[1] M. Boileau, B. Zimmermann, *The π-orbifold group of a link,* Math. Z. **200** (1989), 187–208.

[2] G. Burde, H. Zieschang, Knots, De Gruyter Studies in Mathematics vol. 5, Walter de Gruyter, Berlin-New York, 1985.

[3] D. Cooper, M. Culler, H. Gillet, D. D. Long and P. B. Shalen, *Plane curves associated to character varieties of 3-manifolds*, Invent. Math. **118** (1994), 47–84.

[4] W. Fenchel, Elementary Geometry in Hyperbolic Space, De Gruyter Studies in Mathematics Vol. 11, Walter de Gruyter, Berlin-New York, 1989.

[5] F. González-Acuña, J. M. Montesinos-Amilibia, *On the character varieties of group representations in $SL(2,\mathbb{C})$ and $PSL(2,\mathbb{C})$,* Math. Z. **214** (1993), 627–652.

[6] H. Helling, A. C. Kim, J. L. Mennicke, *Some honey-combs in hyperbolic 3-space,* Comm. Algebra **23** (1995), 5169–5206.

[7] H. M. Hilden, M. T. Lozano, J. M. Montesinos-Amilibia, *On the Borromean orbifolds: geometry and arithmetic,* in: Topology'90, eds. B. A. Apanasov, W. Newmann, A. Reid, L. Siebenmann, de Gruyter, Berlin, 1992, 133–167.

[8] H. M. Hilden, M. T. Lozano, J. M. Montesinos-Amilibia, *On the arithmetic 2-bridge knots and link orbifolds and a new knot invariant,* J. Knot Theory Ramifications **4** (1995), 81–114.

[9] H. M. Hilden, M. T. Lozano, J. M. Montesinos-Amilibia, *On a remarkable polyhedron geometrizing the figure eight cone manifolds,* J. Math. Sci. Univ. Tokyo **2** (1995), 501–561.

[10] H. M. Hilden, M. T. Lozano, J. M. Montesinos-Amilibia, *Volumes and Chern-Simons invariants of cyclic coverings over rational knots,* in: Proceedings of the 37th Taniguchi Symposium on Topology and Teichmüller Spaces held in Finland, July 1995, ed. by Sadayoshi Kojima et al., 1996, 31–35.

[11] C. D. Hodgson, *Schläfli revisited: Variation of volume in constant curvature spaces,* preprint.

[12] L. Kauffman, *New invariants in the theory of knots,* Amer. Math. Monthly **95** (1988), 195–242.

[13] S. Kojima, *Deformation of hyperbolic 3-cone-manifolds,* J. Differential Geom. **49** (1998), 469–516.

[14] A. Mednykh, *On the the remarkable properties of the hyperbolic Whitehead link cone-manifolds,* in: Knots in Hellas '98 (C. McA. Gordon, V. F. R. Jones, L. H. Kauffman, S. Lambropoulou, J. H. Przytycki Eds.), World Scientific, 2000, 290–305.

[15] A. Mednykh, *Trigonometric identities and geometrical inequalities for links and knots,* in: Proceeding of the Third Asian Mathematical Conference (23 - 27 October 2000, Manila, Philippines), World Scientific, 2002, 352–368.

[16] A. Mednykh, *On hyperbolic and spherical volumes for knot and link come-manifolds,* in: Y. Komori, V. Markovic, C. Series (Eds.) Kleinian Groups and Hyperbolic 3-Manifolds, Lond. Math. Soc. Lec. Notes **299**, Cambridge Univ. Press, 2003, 145–165.

[17] A. Mednykh, A. Rasskazov, *Volumes and degeneration of cone-structures on the figure-eight knot,* preprint, 2002, available in http://cis.paisley.ac.uk/research/reports/index.html.

[18] A. Mednykh, A. Vesnin, *Hyperbolic volumes of Fibonacci manifolds,* Siberian Math. J. **36** (1995), 235–245.

[19] A. Mednykh, A. Vesnin, *On the volume of hyperbolic Whitehead link cone-manifolds*, Sci. Ser. A Math. Sci. (N.S.) **8** (2002), 1–11.

[20] D. Rolfsen, Knots and links, Publish or Perish Inc., Berkeley CA., 1976.

[21] W. P. Thurston, The geometry and topology of three-manifolds, Princeton Univ. Press, 1979.

[22] J. Weeks, SnapPea, Computer program for work with hyperbolic structures, available in http://www.geometrygames.org/SnapPea/index.html.

[23] Q. Zhou, *The moduli space of hyperbolic cone structures,* J. Differential Geom. **51** (1999), 517–550.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

ACYLINDRICAL SURFACES IN 3-MANIFOLDS AND KNOT COMPLEMENTS

MARIO EUDAVE-MUÑOZ AND MAX NEUMANN-COTO

Dedicado a Fico en su 60 aniversario

Abstract. We consider closed acylindrical surfaces in 3-manifolds and in knot and link complements, and show that the genus of these surfaces is bounded linearly by the number of tetrahedra in a triangulation of the manifold and by the number of rational (or alternating) tangles in a projection of a link (or knot). For each g we find knots with tunnel number 2 and manifolds of Heegaard genus 3 containing acylindrical surfaces of genus g. Finally, we construct 3-bridge knots containing quasi-Fuchsian surfaces of unbounded genus, and use them to find manifolds of Heegaard genus 2 and homology spheres of Heegaard genus 3 containing infinitely many incompressible surfaces.

1. Introduction

A closed incompressible surface F embedded in a 3-manifold M is called *acylindrical* if the manifold $M_F = M - intN(F)$, obtained by cutting M along F contains no essential annuli (a properly embedded annulus in a 3-manifold is essential if it is incompressible and not boundary parallel). Acylindrical surfaces are interesting in connection with geometry, as every totally geodesic surface in a hyperbolic 3-manifold is acylindrical, and every acylindrical surface in a hyperbolic link complement is quasi-Fuchsian. Moreover, if F is an acylindrical surface in a closed, irreducible and atoroidal 3-manifold M then M_F admits a hyperbolic metric with totally geodesic boundary [21].

In [12] Hass proved that for the finite volume hyperbolic 3-manifolds there is a constant C, independent of the manifold, so that each acylindrical surface in a manifold M has genus at most $C \cdot vol(M)$. He used this result to show that in any compact 3-manifold there is only a finite number of acylindrical surfaces. It seems natural to ask if there are similar bounds which hold for all 3-manifolds and depend not on volume, but on some topological measures of complexity. Some candidates could be the number of tetrahedra in a triangulation or the Heegaard genus of the manifold, and in the case of knots and links, the crossing number, the bridge number or the tunnel number. Such bounds must exist in the case of the number of tetrahedra in a triangulation or the crossing number of a link, as there are only finitely many manifolds and links for each number n. We find explicit bounds in these cases, and furthermore show that there is a

2000 *Mathematics Subject Classification*: 57N10, 57M25.

Keywords and phrases: acylindrical surface, quasi-Fuchsian surface, incompressible surface, triangulations, Heegaard genus, tangles, tunnel number.

The authors were supported in part by a CONACyT grant.

linear bound in terms of the number of rational tangles in a link projection or the number of alternating tangles in a prime knot projection.

The fact that 3-manifolds with Heegaard genus 2 and the complements of knots with tunnel number 1 contain no separating acylindrical surfaces ([19], [4]) could suggest that -at least for small Heegaard genus or tunnel number- there could be bounds for the genus of such surfaces. We show here that for each g, there are tunnel number 2 knots which contain a closed acylindrical surface of genus g. By performing suitable Dehn surgeries, we get closed manifolds of Heegaard genus 3 which contain closed acylindrical surfaces of genus g. These examples show that Heegaard genus 3 manifolds and tunnel number 2 knots are already quite complicated.

We also consider what happens when the acylindrical assumption is weakened to require that there are no essential annuli running from the surface to a boundary torus (in the case of hyperbolic knots and links this means that the surface is quasi-Fuchsian). We show that a knot that can be decomposed into two alternating tangles cannot contain any quasi-Fuchsian surfaces in its complement. On the other hand, we find hyperbolic 3-bridge knots whose complements contain infinitely many quasi-Fuchsian surfaces. These knots have an essential branched surface which carries quasi-Fuchsian surfaces of arbitrarily high genus. These examples show that there are no bounds for the genus of quasi-Fuchsian surfaces based on volume, crossing number or the number of tetrahedra. Finally, by means of suitable Dehn fillings and double covers, we produce manifolds of Heegaard genus 2 and homology spheres of Heegaard genus 3 which contain infinitely many incompressible surfaces. These examples are interesting, for it seems that all known examples of hyperbolic manifolds with infinitely many surfaces have noncyclic homology, and in the case of knots with infinitely many surfaces, it seems that the only known explicit examples are some satellite knots (see for example [17]). The examples are also interesting for the study of surfaces in the complement of 3-bridge knots, as they supplement results of Finkelstein and Moriah [6], who showed that many 3-bridge knots contain an incompressible but meridionally compressible surface, and of Ichihara and Ozawa [15], who proved that any closed surface in the complement of a 3-bridge knot is meridionally compressible or annular.

2. Bounds for the genus of acylindrical surfaces

PROPOSITION (2.1). *If a closed 3-manifold M admits a (pseudo)triangulation with n tetrahedra then the genus of a 2-sided closed acylindrical surface in M is at most $\frac{n+1}{2}$.*

Proof. Let T be a (pseudo)triangulation of M with n tetrahedra, and denote by T_i the i-skeleton of T.

Let F be an incompressible surface in M in normal position with respect to the triangulation, so F intersects the faces of the tetrahedra along arcs and the interior of the tetrahedra along discs which are triangles or squares. Assume further that F has been isotoped to minimize the number of intersections with T_1. Let $\overline{F}$ be the boundary of a regular neighborhood N of F. As F is two-sided, $\overline{F}$ consists of two copies of F. By definition F is acylindrical iff $M - intN$ contains no essential annuli.

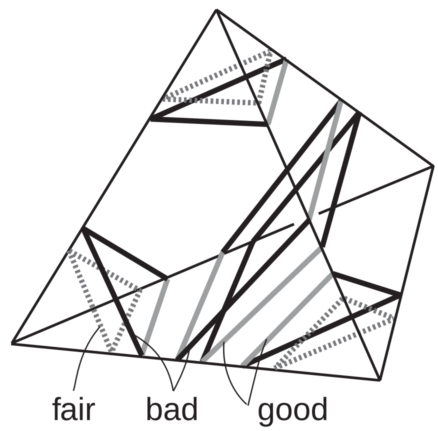


Figure 1.

The edges of $\overline{F}$ in each face of a tetrahedron split the face into triangles, quadrangles, pentagons and/or hexagons, and each edge is adjacent to a quadrangle (which lies in N). Call an edge *good* if the other adjacent region (which lies in $M - intN$) is also a quadrangle. Notice that if an embedded curve c in $\overline{F}$ is made of good edges, then the union of these adjacent quadrangles in $M - intN$ forms an annulus A that joins c with another curve c' in $\overline{F}$.

We claim that if c is essential in $\overline{F}$ then the annulus A is essential. Otherwise A would be isotopic to an annulus A' bounded by c and c' in $\overline{F}$ (in particular, c must be 2-sided in $\overline{F}$). As F is 2-sided in M, then A' is parallel to an annulus A'' in F and the isotopy from A' to A can be used to isotope A'' (pushing it even further across A) to reduce the number of intersections of F with T_1.

So if $\overline{F}$ is acylindrical, then the good edges of $\overline{F}$ carry no embedded essential curves, and so they carry no essential curves at all. But as the edges of $\overline{F}$ split $\overline{F}$ into discs, they must carry all of $H_1(\overline{F})$.

So there must be at least as many non-good edges in $\overline{F}$ as the rank of $H_1(\overline{F})$. As the number of non-good edges in a face of a tetrahedron is at most 6, the total number of non-good edges in $\overline{F}$ is at most $12n$, so $12n \geq \text{rank } H_1(\overline{F}) = 2 \cdot \text{genus } \overline{F}$, and so the genus of F is at most $3n$.

In order to get the better estimate one needs to look more carefully at the graph Q formed by the edges of $\overline{F}$. Divide the non-good edges in each tetrahedron in two classes: those lying in triangles of $\overline{F}$ that cut off outermost corners of the tetrahedron will be called *fair edges* and the others (which may lie in squares or triangles) will be called *bad edges* (see Figure 1).

Let Q_g and Q_f denote the subgraphs of Q made of good edges and fair edges respectively.

Observe that Q_g and Q_f are disjoint, that is, have no vertices in common. As the components of Q_f lie in the links of the vertices of a triangulation of the manifold M, all curves contained in Q_f are contractible in M, and so as $\overline{F}$ is incompressible then Q_f contains only trivial curves of $\overline{F}$. All curves contained in Q_g are also trivial because $\overline{F}$ is acylindrical. So as $Q_g \cup Q_f$ contains only trivial curves of $\overline{F}$, by attaching to $Q_g \cup Q_f$ some of the complementary pieces of $\overline{F}$ we obtain a (possibly empty or disconnected) simply connected subcomplex $\overline{F}_S$ of $\overline{F}$.

Now the Euler characteristic of $\overline{F}$ is $\chi(\overline{F}) = \chi(\overline{F}_S) + v - e + f$ where v, e, and f count the vertices, edges and faces of $\overline{F}$ that do not lie in $\overline{F}_S$. So e counts some bad edges -some others may lie in $\overline{F}_S$- and f counts the triangles and squares adjacent to them. It can be shown directly that in each tetrahedron Δ, every subcollection f_Δ of the set of squares and triangles of $\overline{F} \cap \Delta$ with bad edges satisfies the inequality $\frac{1}{2}e_\Delta - f_\Delta \leq 2$. In particular, we may take f_Δ to be the set of squares and triangles with bad edges not contained in $\overline{F}_S$. As $\overline{F}$ has two components and each of them contains a component of $\overline{F}_S$ or a vertex, it follows that

$\chi(\overline{F}) \geq 2 - e + f = 2 + \sum_{\Delta \in T_3} -\frac{1}{2}e_\Delta + f_\Delta \geq 2 - 2n$

and so $\operatorname{genus}(F) = \frac{1}{4} \operatorname{rank} H_1(\overline{F}) = \frac{1}{4}(4 - \chi(\overline{F})) \leq \frac{1}{4}(2 + 2n)$. □

The genus of an acylindrical surface in a manifold is not bounded in terms of its Heegaard genus, as we show in Section 3. However, there is a bound depending on the complexity of a Heegaard splitting. Let $M = H \cup H'$ be a Heegaard splitting of M of genus g, and let $D_1, D_2, ..., D_g$ and $D'_1, D'_2, ..., D'_g$ be discs splitting H and H' into 3-balls B and B'. The complexity of the Heegaard splitting with respect to these discs is just the minimal intersection number between the boundaries of the discs. The complexity of a Heegaard splitting is the minimum complexity among all such systems of discs.

Proposition (2.2). *If a closed 3-manifold M admits an irreducible Heegaard splitting of genus g and complexity n then the genus of a closed acylindrical surface in M is at most $(n - \frac{3}{2}g)$.*

Proof. Let $M = H \cup H'$ be a Heegaard splitting of M of genus g as above, with $\cup D_i$ meeting $\cup D'_j$ in n points. Let F be an acylindrical surface in M. As F is incompressible, we may assume that F meets H' along g stacks of parallel discs in $N(D'_j)$ (some stacks may be empty). We may also assume that F meets B along discs and that it meets each D_i along stacks of parallel arcs connecting different components of $\partial D_i \cap N(D'_j)$.

As before, consider the graph of intersection Q of $\overline{F}$ with $\partial H \cup_i D_i$. Call an edge of Q on D_i *good* if it is an interior arc of a stack, otherwise call it *bad.* Call an edge of Q in ∂H *good* if it is part of the boundary of an interior disc of a stack. Otherwise (i.e., if it is part of the boundary of an outermost disc of a stack) call it *fair.* See Figure 2.

Observe that the subgraphs Q_g and Q_f made of good edges and fair edges do not meet. As the components of Q_f are contained in the boundaries of discs in H' then Q_f carries no essential curves of $\overline{F}$, and as F is acylindrical we may assume as in the proof of 2.1 that Q_g carries no essential curves either.

So as $Q_g \cup Q_f$ carries no essential curves and Q splits $\overline{F}$ into discs, the rank of $H_1(\overline{F})$ is bounded above by the number of bad edges. If D_i meets the D'_j in n_i points then $n_i > 1$ (because the Heegaard splitting is irreducible) and D_i contains at most $4n_i - 6$ bad edges, so the rank of $H_1(\overline{F})$ is at most $\sum_{D_i} (4n_i - 6) = 4n - 6g$ and so the genus of F is at most $n - \frac{3}{2}g$. □

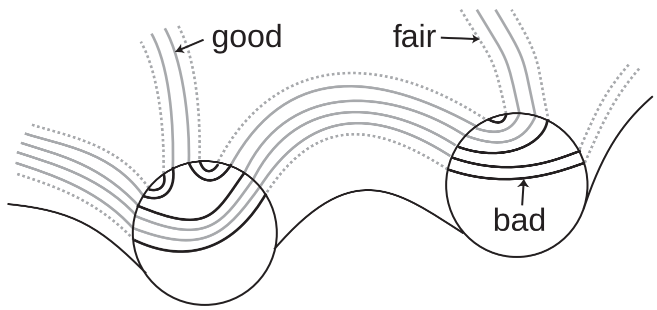


Figure 2.

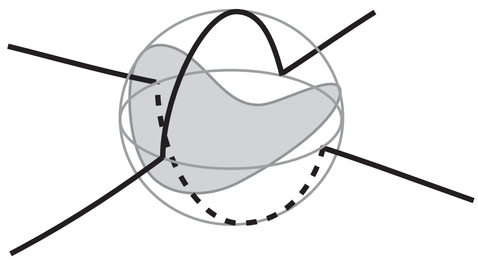


Figure 3.

There are other ways of measuring the complexity of a Heegaard splitting, for example, by means of the curve complex, as defined in [14]. Note however that no such bound for the genus of acylindrical surfaces exists for this complexity, for in fact, all the examples constructed in Section 3 have a Heegaard splitting of genus 3 which comes from a certain bridge presentation of a knot, and then by a similar proof to Theorem 1.4 of [14], the distance in the curve complex is ≤ 2.

We now consider bounds for the genus of acylindrical surfaces in the exterior of knots and links in the 3-sphere.

PROPOSITION (2.3). *If k is a knot or link with n crossings then the genus of a closed acylindrical surface in the exterior of k is at most $\frac{3}{2}n - 3$.*

Proof. Draw k on a projection sphere S, except for the crossings which lie on the surface of n small spheres S_1, S_2,...,S_n. Let S_0 be the part of the projection sphere outside the S_i's. Then $S_0 \cup_i S_i$ cuts S^3 into $n+2$ polyhedral balls B^-, B^+ and B_1, B_2, ..., B_n, with faces determined by the equators of the bubbles and the arcs of k. If F is an incompressible surface in the exterior of k then F can be isotoped to meet B^+ and B^- along discs, meet each B_i along parallel saddle-shaped discs, and meet their faces along arcs. See Figure 3.

Let $\overline{F}$ be the boundary of a regular neighborhood of F, and let Q be the graph of intersection of $\overline{F}$ with $S_0 \cup_i S_i$. So Q splits $\overline{F}$ into discs. As before, consider the edges of Q on each face of $S_0 \cup_i S_i$, call those that have parallel edges on both sides *good*, those which are closest to arcs of k and are parallel to

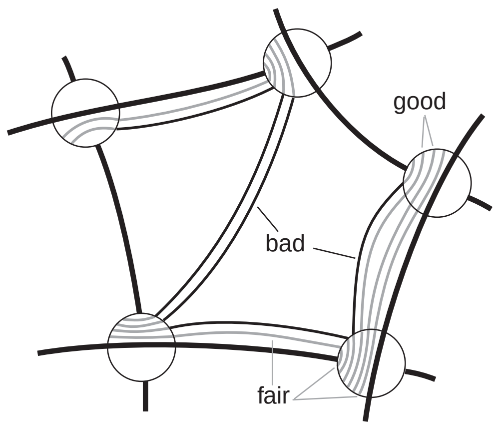


Figure 4.

them *fair*, those lying on some S_i and parallel to an arc of ∂S_0 which contain a point of k are also fair, and all the others are called *bad* edges (so the faces of the B_i's contain no bad edges). See Figure 4.

Again the subgraphs Q_g and Q_f of Q are disjoint, and if $\overline{F}$ is acylindrical then Q_g carries no essential curves of $\overline{F}$. On the other hand, Q_f can be regarded as lying on the boundary tori of a regular neighborhood of the link k. But as $\overline{F}$ is acylindrical, there can be no essential annuli running from $\overline{F}$ to k, so Q_f contains no essential curves of $\overline{F}$. So, as the graph Q carries all of $H_1(\overline{F})$, there must be at least as many bad edges as the rank of $H_1(\overline{F})$. As there are at most $3i-6$ bad edges on each i-gon determined by the projection of k into S, the number of bad edges in $\overline{F}$ is at most
$\sum_{i-\text{gons in } P}(3i-6) = 3(2(\text{ arcs of } k \text{ in } S)) - 6\,(\text{regions determined by } k \text{ in } S)$
$= 12n - 6(2+n) = 6n - 12$

So the rank of $H_1(\overline{F})$ is at most $6n-12$ and the genus of F is at most $\frac{6}{4}n - \frac{12}{4}$. $\square$

After we proved Proposition (2.3), we learned that Agol and D. Thurston, following Lackenby [16], showed that the volume of a hyperbolic knot of link is bounded above by $10v_3(t(D)-1)$ where v_3 is the volume of a hyperbolic ideal tetrahedra and t is the twist number of k (the minimum number of twists in a diagram of k, where a twist is a string of 2-gons or a crossing in the diagram). Agol has also shown [1] that if a hyperbolic manifold M has an acylindrical surface of genus g, then $Vol(M) \geq 4v_3(g-1)$. It follows that the genus of an acylindrical surface in the exterior of a hyperbolic link k is at most $\frac{5}{2}t$. These results suggested the following.

Recall that a tangle is a 3-ball B together with two properly embedded arcs. The tangle is rational if the arcs are isotopic (rel ∂) to arcs in ∂B. We will say that a knot or link k in S^3 is decomposed into tangles if there is a sphere S and 3-balls B_1, B_2, ..., B_n each intersecting S in a disc, so that $k \cap B_i$ is a tangle, and the part of k outside these balls is a collection of arcs lying on $S_0 = S - int(\cap B_i)$.

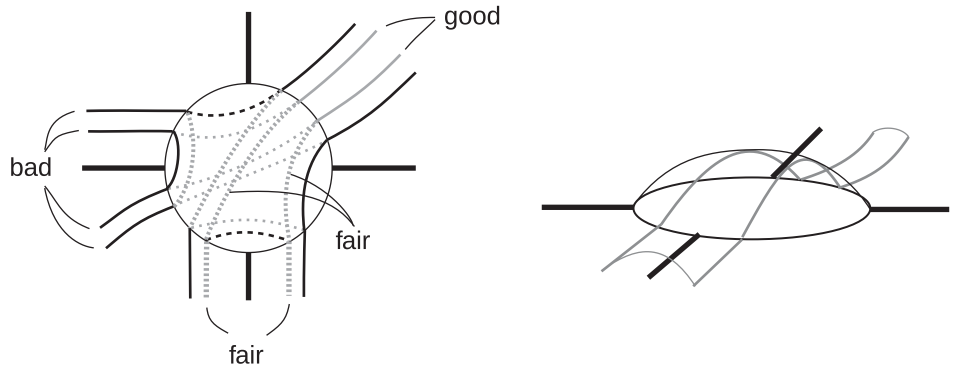


Figure 5.

THEOREM (2.4). *If a link is decomposed into n rational tangles, then the genus of a closed acylindrical surface in its complement is at most* $2n-4$.

Proof. Draw the projection of the link k as the union of n rational tangles in the interior of n disjoint spheres S_1, S_2,...,S_n joined by $2n$ disjoint arcs in the projection sphere. Let S_0 be the part of the projection sphere outside these spheres. Then $S_0 \cup_i S_i$ cuts S^3 into polyhedral balls B_+, B_- and B_1,B_2,...,B_n with faces determined by the equators of the spheres and the arcs of k in S_0.

If F is an incompressible surface in the exterior of k we may isotope F so that intersects B^- and B^+ along discs, and intersects S and the hemispheres of each S_i along arcs. Moreover, as $k \cap B_i$ is a rational tangle, we may isotope F to intersect $B_i - k$ along parallel discs that separate the strings of the tangle, and we may assume that their boundaries meet each hemisphere of S_i along 2 or 3 families of parallel arcs -2 if the tangle is a crossing and 3 otherwise (a single family of parallel arcs implies that the discs are vertical and the tangle has no crossings of k).

Let $\overline{F}$ be the boundary of a regular neighborhood N of F. The intersection of $\overline{F}$ with $S_0 \cup_i S_i$ gives a cell decomposition of $\overline{F}$ and cuts the faces of $S_0 \cup_i S_i$ into quadrangles that lie in N and other polygons that lie in $S^3 - N$; as before, let Q be the graph of intersection. Call an edge of Q in a face of $S_0 \cup_i S_i$ *good* if the adjacent polygon in $S^3 - N$ is a quadrangle with another edge on $\overline{F}$ (so the two edges are parallel in that face). Otherwise, call an edge in Q *fair* if it is adjacent to a quadrangle in S_0 with a side in $k \cap S_0$ that is adjacent to another quadrangle in S_0 with a side in Q (so both edges of Q are parallel to this arc of k) or if it is adjacent to a polygon in a hemisphere of S_i with exactly 2 sides in Q (so the other sides lie in the equator and are separated by points of $k \cap S_i$). Call the other edges of Q *bad*. Note that edges lying on some S_i and parallel to an arc of ∂S_0 which contain a point of k are bad. See Figure 5.

One can use the fair edges as well as the good edges to construct annuli for $\overline{F}$, by taking the quadrangles that lie between two fair edges in $S_0 \cup_i S_i$ (but that may intersect k) and pushing them outside the corresponding S_i, or if they lie in S_0, to the side of S_0 that doesn't contain an edge of Q connecting the two fair edges (there can't be connecting edges on both sides because the union of the four edges would be a meridian of k, and so F would be meridionally compressible). This creates quadrangles in $S^3 - k$ connecting pairs of fair edges, and one can see that the quadrangles corresponding to consecutive fair or good edges match well.

As before, if a simple essential curve in $\overline{F}$ is made of good and fair edges then the annulus formed by the union of the adjacent quadrangles is essential or else F could be isotoped to reduce its intersection with $S_0 \cup_i S_i$. So, if F is acylindrical, the subgraph of Q consisting of the good and fair edges cannot contain any essential curve of $\overline{F}$, so it is contained in a simply connected subcomplex $\overline{F}_S$ of $\overline{F}$. Again, as $\overline{F}$ has two components and Q divides them into discs,

$\chi(\overline{F}) = \chi(\overline{F}_S) + v - e + f \geq 2 - e + f$ where v, e, and f count the vertices, edges and discs in $\overline{F} - \overline{F}_S$, and so $\operatorname{rank} H_1(\overline{F}) = (4 - \chi(\overline{F}) \leq 2 + e - f$.

As there are at most 4 bad edges and 8 fair edges on each S_i, all contained in the 2 outermost discs of $\overline{F} \cap B_i$, the number of bad edges minus the number of discs that contain them in $\cup S_i$ is at most $2n$.

There are at most $i-3$ families of parallel edges on each face of S_0 determined by $i > 1$ arcs of k, not including the families of edges parallel to the arcs of k, and they produce at most $2i - 6$ bad edges on each face. If an arc of k has parallel families edges of Q on both sides, then there are two bad edges in these families, for the edges closest to k are fair. If an arc of k has only edges of Q on one side, then there are two bad edges in this family, since in this case the edge closest to k is not fair.

So, if no face of S_0 is a monogon the number of bad edges in S_0 is at most $\sum_{edgesofk} 2 + \sum_{i-\text{gons in } P} (2i - 6) = 4n + 8n - 6(2 + n) = 6n - 12$.

When $i = 1$, the previous formula undercounts the number of bad edges in the monogon as -3 instead of 0 -there are no edges in the monogon as they could be isotoped into B_i to eliminate two intersection curves of $\overline{F}$ with S_i-. In this case there cannot be bad edges around the endpoints of the monogon in S_i and so the discs of intersection of $\overline{F}$ with S_i are vertical and the tangle is trivial -unless $\overline{F}$ does not meet S_i at all, so there is an overcount on the number of bad edges in $\cup S_i$ by at least 2 and also on the number of bad edges in the face of S_0 adjacent to the monogon. So the previous bound also holds when some faces of S_0 are monogons.

Finally observe that since $\overline{F}$ has 2 components and each of them must meet B_+ and B_-, there must be at least 2 discs of $\overline{F} - \overline{F}_S$ inside each of these balls.

So, $\operatorname{genus}(F) = \frac{1}{4} \operatorname{rank} H_1(\overline{F}) \leq \frac{1}{4}(2 + e - f) \leq \frac{1}{4}(2 + 2n + (6n - 12) - 4) = 2n - \frac{7}{2}$ □

Consider a tangle as above, i.e., it is determined by the intersection of a 3-ball B with a link k, so that $B \cap S$ is a disc, where S is a projection sphere, and $k \cap \partial B$ consists of 4 points lying on S. We say that the tangle is alternating if its arcs can be isotoped, keeping ∂B fixed, to have an alternating projection on the sphere S. Note that each rational tangle is alternating.

The next result extends Theorem (2.4) to allow alternating tangles.

THEOREM (2.5). *If a prime knot is decomposed into alternating tangles, n of them rational, then the genus of a closed acylindrical surface in its complement is at most $2n-4$.*

The proof is based on the following:

Claim (2.6). Let k be a nonseparable link or a knot and S a sphere that meets k in 4 points. Then each acylindrical surface F in S^3-k is isotopic to one that either i) is disjoint from S or ii) intersects S in one curve or iii) meets one of the components of S^3-k-S along parallel discs.

Proof. The sphere S separates k into two tangles. Isotope F to minimize its intersection with the 4-punctured sphere $S-k$. The intersection then contains no trivial curves, and as F is meridionally incompressible then it does not contain curves surrounding only one puncture, so all the curves c_1,c_2,...c_n in which F intersects S must be parallel in $S-k$. As F is acylindrical, if there is more than one c_i then the annuli connecting two of them in S cannot be essential, so either one annulus is isotopic (rel ∂) to an annulus in F (and the isotopy can be used to remove two c_i's) or all the c_i's bound discs of F. So at least one of them, say c_1, bounds a disc D_1 in F that lies completely on one side of S. But then, as all c_i's are parallel to ∂D_1, one can draw parallel discs D_i in S^3-k on that side of S that meet F at c_i (and nowhere else). The union of the discs bounded by the c_i's in F and the D_i's form spheres in S^3-k, and if k is a knot or a nonseparable link these spheres bound balls in S^3-k, so the D_i's must be isotopic to the discs in F, and the isotopy reduces the number of curves unless the discs in F were already on one side of S. □

Claim (2.7). If k is a prime knot and $k \cap B_i$ is an alternating tangle, then every acylindrical surface in the complement of k can be isotoped to meet B_i-k along parallel discs or be disjoint from it.

Proof of theorem. Assume for the moment that Claim (2.7) is true, and isotope the surface F to meet only the B_i's corresponding to separable tangles. To estimate the genus of F we would like to count the number of bad edges and discs of $\overline{F}$ that contain them by replacing each nonseparable tangle in the diagram of k by a trivial tangle to get a knot k' and counting the bad edges of $\overline{F}$ in its diagram.

Now some bad edges in the diagram of k may become fair in the diagram of k' as in Figure 6, but in this case we may regard them as originally being "almost fair" -there is a quadrangle joining them that lies above or below the nonseparable tangles that were between them in the diagram of k. The quadrangles corresponding to almost fair edges match well with the other quadrangles corresponding to good and fair pairs of edges, so they can be used as well to construct annuli for $\overline{F}$. So the same bound for the number of bad edges and discs -and therefore the same bound for the genus of F- holds. □

The proof of Claim (2.7) is based on the following extension of the Meridional Lemma of Menasco [18].

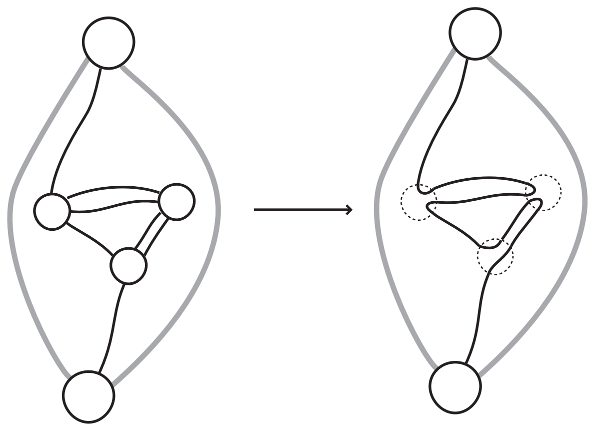


Figure 6.

LEMMA (2.8). *If a link k intersects a ball B in an alternating tangle, then every meridionally incompressible surface in the complement of k can be isotoped to intersect B along copies of a surface that separates the strings of the tangle.*

Proof. Draw B as a round ball with $k \cap B$ lying in an equatorial disc except at the crossings, that lie on the surface of small "bubbles" B_1, B_2,....as in Figure 3. Let ∂B_{i+} and ∂B_{i-} be the hemispheres of ∂B_i, and let D_0 denote the part of the equatorial disc outside the bubbles. Let $D_+ = D_0 \cup_i \partial B_{i+}$ and $D_- = D_0 \cup_i \partial B_{i-}$, and let B_+ and B_- be the parts of B above and below D_+ and D_-.

If F is a meridionally incompressible surface in the complement of k then by isotoping F to minimize its intersection with $\partial B \cup D_0 \cup_i \partial B_i$ we can assume that F meets ∂B along parallel curves that separate 2 points of $\partial B \cap k$ from the other 2, that F meets D_0 and each hemisphere of ∂B and ∂B_i along arcs and that meets B_+ and B_- along discs and each B_i along parallel saddle-shaped discs. So F intersects D_+ and D_- along curves and arcs with endpoints in ∂B.

Following Menasco, one can show that the curves and arcs of intersection of F with D_+ (and similarly with D_-) have the following properties:

1. As F is incompressible, each curve (and each arc) crosses at least one bubble.

2. As F is meridionally incompressible, each curve (or arc) crosses each bubble at most once.

3. As the diagram of $k \cap B$ is alternating, if a curve (or arc) crosses two bubbles B_i and B_j in succession, then the 2 arcs $k \cap \partial B_{i+}$ and $k \cap \partial B_{j+}$ lie on opposite sides of the curve. See Figure 7.

So there can be no closed curves in D_+, because by properties 1 and 3 an innermost such curve would have to leave an arc of $k \cap \partial B_{i+}$ inside (so there would be another curve inside) unless the curve crossed the same bubble twice, contradicting property 2.

Let k^1 and k^2, k_1 and k_2 be the 4 segments of $k \cap D_0$ that start on ∂D_0, and end in overcrossings or undercrossings of k respectively. Note that ∂D_0

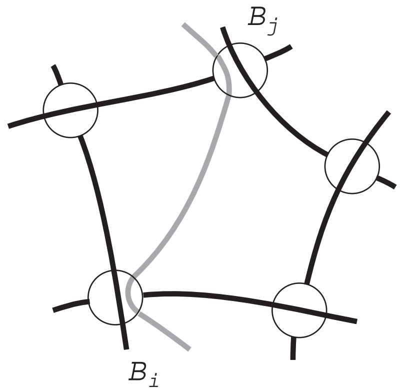


Figure 7.

encounters them in the order k^1, k_1, k^2, k_2, for otherwise there is an arc on D_0 separating the strings of the tangle, but then the knot will be composite. Properties 1, 2 and 3 for arcs imply that each outermost arc in D_+ goes around k_1 or k_2 and so every arc in D_+ must separate k_1 from k_2. See Figure 8a.

Now let F_0 be a surface consisting of one or more components of $F \cap B$. If F_0 does not separate the strings of the tangle then each path in B joining the strings must meet F_0 in an even number of points, so F_0 intersects each bubble in an even number of discs, and so the number of curves and arcs cross ∂B_{i+} on each side of $k \cap \partial B_{i+}$ is even. We claim that in these conditions $F \cap D_+$ consists of pairs of parallel arcs.

To show this, order the arcs according to its distance from k_1, and assume that the first $2n$ are paired and let a be the next one. Let B_i and B_j be two consecutive bubbles crossed by a, so the segments of $k \cap \partial B_{i+}$ and $k \cap \partial B_{j+}$ are on opposite sides of a as in Figure 8a. Since all the curves on one side of a are paired and each side of the bubbles is crossed by an even number of arcs, there must be other arcs a' and a'' crossing B_{i+} and B_{j+} next to a. If a' and a'' are different, then one of them cannot separate k_1 from k_2 (see Figure 8b). If $a' = a''$ then either a and a' run parallel from B_i to B_j or else a' crosses other bubbles between B_i and B_j. If so, let B_l be the bubble crossed by a' immediately after B_j. See Figure 8c. Then $k \cap \partial B_{l+}$ lies between a and a', and so there must be another arc between a and a', and this arc would have to cross B_i or B_j between a and a', and this is impossible. Therefore a' must run parallel to a from the first bubble to the last bubble crossed by a. It remains to show that a' runs parallel to a from the first bubble to the boundary of D_+ and from the last bubble to the boundary of D_+, i.e., that a' does not meet other bubbles in its way to the boundary and that the region between a and a' does not contain other bubbles. As k_1 and k_2 lie outside the region between a and a', this region does not contain any other arc a''. So k^1 and k^2 also lie outside this region, because if k^1 were between a and a' the number of arcs between k_1 and k^1 would be odd, so F_0 would separate these strings of k. Now if there were any segments of $k \cap D_+$ in that region, k would have to enter and leave the region at 2 bubbles crossed by a'

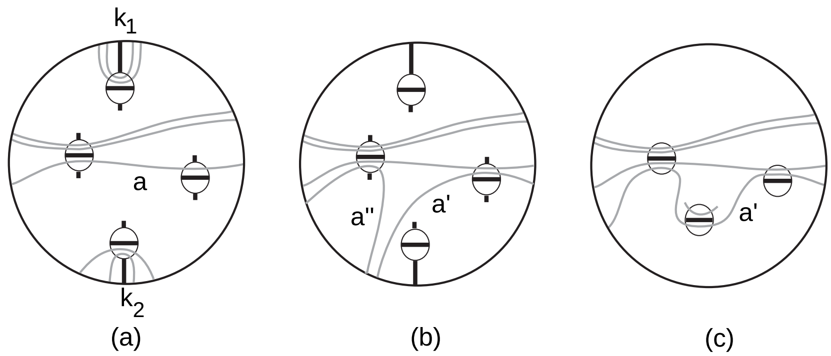


Figure 8.

on its way to the boundary. But we know that for any two consecutive bubbles crossed by a' the segments of k in their upper hemispheres lie on opposite sides of a', so one of them is in the region between a and a' and so there must be an arc in that region, a contradiction.

Now observe that each pair of parallel arcs of F_0 in D_+ must be adjacent to a pair of parallel arcs of F_0 in ∂B_+: an arc of F_0 in ∂B_+ cannot go around the endpoints of k^1 or k^2 because F would be meridionally compressible and something analogous holds for the arcs of F_0 in D_-. So the intersection of F_0 with ∂B_+, ∂B_- and with each ∂B_i consists of pairs of parallel curves, and as F_0 is assembled by attaching discs to these parallel curves, F_0 must consist of pairs of parallel surfaces.

Finally, as F is meridionally incompressible, the intersection of F with the 4-punctured sphere $\partial B - k$ consists of curves surrounding 2 punctures, and if there is more than one curve these are parallel. So if $F \cap B$ has several components, and F_0 consists of any two of them, then any path in B joining the strings of the tangle must intersect F_0 in an even number of points, and this is all that we needed before to show that F_0 consists of parallel surfaces. □

Proof of Claim (2.7). Isotope F to minimize its intersection with ∂B_i. By the previous lemma if $F \cap B_i$ is not empty then it consists of parallel copies of a surface F_0 that separates the strings of the tangle. As k is a knot F cannot separate the strings of $k \cap B$, so there must be an even number of copies of F_0. Now by the previous claim either $F \cap B_i$ or $F \cap S^3 - B_i$ consists of discs, and in the second case F would be the union of the components of $F \cap B_i$ with discs, and since there are at least two such components F would not be connected. □

In [2] Adams *et al.* extended the Meridional Lemma of Menasco to almost alternating knots, i.e. knots that can be obtained by changing one crossing of an alternating knot. The following corollary extends it to knots that can be obtained from an alternating one by mirroring any (2-string) tangle.

Corollary (2.9). *If a knot k can be decomposed into 2 alternating tangles, then k admits no meridionally incompressible surfaces in its complement.*

Proof. By Lemma (2.8), a meridionally incompressible surface F in the complement of k can be isotoped to meet each of the balls B_1 and B_2 that determine the tangles along an even number of parallel copies of a surface F_i that separates the strings of the tangle.

So $F \cap B_i$ is the boundary of a regular neighborhood N_i of one or more copies of F_i, and N_i is determined by painting the components of $B_i - F$ in a chessboard fashion and choosing those whose color is different from that of the regions that contain the strings of the tangle. So N_1 and N_2 match on $\partial B_1 = \partial B_2$ to form the regular neighborhood of a single surface in S^3, and F is its boundary, so F cannot be connected. □

Corollary (2.10). *The total genus of a disjoint family of closed, embedded, totally geodesic surfaces in a hyperbolic 3-manifold or link complement is bounded above by:*

- $\frac{3}{2}t$ *where t is the number of tetrahedra in a triangulation.*
- $n - \frac{3}{2}g$ *for manifolds of Heegaard genus g and complexity n.*
- $\frac{3}{2}c - 3$ *for a link with c crossings.*
- $\frac{5}{2}r - 3$ *for a link that admits a projection made of r rational tangles.*
- $\frac{5}{2}r - 3$ *for a prime knot decomposed into alternating tangles, r of them rational.*

Proof. If M is a hyperbolic 3-manifold and F_1,F_2,...,F_k are disjoint totally geodesic surfaces in M, then each F_i is acylindrical and there are no essential annuli in M connecting two F_i's. For, the preimages of the F_i's in the universal covering of M are disjoint totally geodesic planes in $\mathbb{H}^3$, and each preimage of an essential annulus is an infinite strip of bounded height connecting two lines in different planes. These lines lie at a bounded distance from geodesic lines representing the preimages of the boundaries of the annulus, so they determine 2 different points at infinity were the two planes meet, but two disjoint totally geodesic planes in H^3 can only meet at 1 point.

So we may consider the family F_1,F_2,...,F_k as a single disconnected acylindrical surface. The arguments above show the existence of essential annuli for a surface F if the rank of $H_1(F)$ is higher than the number of bad edges, independently of the number of components of F. The bounds arise from a count of the number of bad edges in each case. □

3. Acylindrical surfaces in tunnel number two complements

Let S be a closed surface of genus g standarly embedded in S^3, that is, it bounds a handlebody on each of its sides. A knot K has a (b, g)-presentation if can be isotoped to intersect S transversely in $2b$ points that divide K into $2b$ arcs, so that the b arcs in each side can be isotoped, keeping the endpoints fixed, to disjoint arcs on S. We say that a knot K is a (b, g)-knot if it has a (b, g)-presentation. Consider a product neighborhood $S \times I$ of S. To say that a knot K has a (b, g)-presentation is equivalent to say that K can be isotoped to lie in $S \times I$, so that $K \cap (S \times \{0\})$ and $K \cap (S \times \{1\})$ consist each of b arcs (or b tangent points), and the rest of the knot consist of $2b$ straight arcs in $S \times I$, that is, arcs which intersect each leaf $S \times \{t\}$ in the product exactly in one point. It is not difficult to see that if K is a (b, g)-knot, then the tunnel number of K,

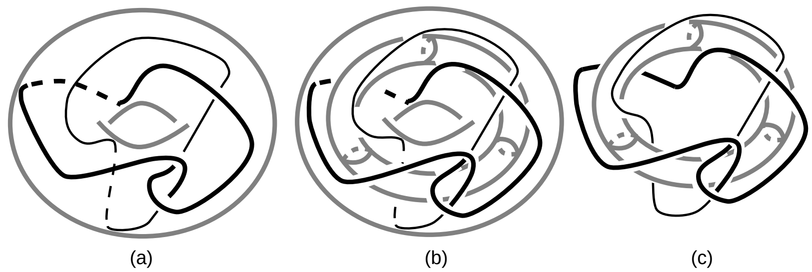


Figure 9.

denoted $tn(K)$, satisfies $tn(K) \leq b + g - 1$. In this section we construct $(2,1)$-knots, which are in fact tunnel number 2 knots, which contain an acylindrical surface of genus g.

Let T be a standard torus in S^3, and let $I = [0,1]$. Consider $T \times I \subset S^3$. $T \times \{0\}$ bounds a solid torus R_0, and $T \times \{1\}$ bounds a solid torus R_1, such that $S^3 = R_0 \cup (T \times I) \cup R_1$. Choose $n+1$ distinct points on I, $e_0 = 0,\ e_1, \ldots, e_n = 1$, so that $e_i < e_{i+1}$, for all $0 \leq i \leq n-1$. Consider the tori $T \times \{e_i\}$. By a vertical arc in a product $T \times [a,b]$ we mean an embedded arc which intersects every torus $T \times \{x\}$ in the product in at most one point.

Let γ_i be a simple closed essential curve embedded in the product $T \times [e_{i-1}, e_i]$, for $i = 1, \ldots, n$, so that it has only one local maximum and one local minimum with respect to the projection to $[e_{i-1}, e_i]$. Let α_i, for $i = 1, \ldots, n-1$, be a vertical arc in $T \times [0,1]$, joining the maximum point of γ_i with the minimum of γ_{i+1}. Let Γ be the 1-complex consisting of the union of all the curves γ_i and the arcs α_j. So Γ is a trivalent graph embedded in S^3. Let $R'_0 = R_0 \cup (T \times [e_0, e_1])$ and $R'_1 = R_1 \cup (T \times [e_{n-1}, e_n])$.

Suppose each curve γ_i satisfies the following:

1. γ_i is not in a 3-ball contained in $T \times [e_{i-1}, e_i]$, or in R'_0 or R'_1, that is, it is not a trivial knot in that region.
2. γ_i is not isotopic in $T \times [e_{i-1}, e_i]$, or in R'_0 or R'_1, to a knot lying on the torus $T \times \{e_i\}$.
3. γ_i is not a cable of a knot lying in $T \times [e_{i-1}, e_i]$ or in R'_0 or R'_1 (it can be proved that this is equivalent to say that γ_i is not isotopic to a cable of a knot lying on the torus $T \times \{e_i\}$.)
4. There is no annulus B in $T \times \{e_0\}$ so that $B \times [0,1]$ contains Γ. If that happens then each curve γ_i would be contained in a product $B \times [e_{i-1}, e_i]$.
5. There is no Möbius band in R'_0 (R'_1) disjoint from γ_1 (γ_n).

It is not difficult to see that there exist plenty of knots satisfying the conditions required for the curves γ_i, say by taking each γ_i to be a $(1,1)$-knot which is not a torus knot nor a satellite knot. For example, each γ_i could be a copy of the figure eight knot, as shown in Figure 9(a) in the case of γ_1, Figure 9(b) for $\gamma_2, \ldots, \gamma_{n-1}$, and Figure 9(c) for γ_n. In the figures the knot is divided in two arcs; the thin arc contains the minimum point of the knot, and the bold arc contains the maximum. When assembled we get the graph Γ, shown for $n = 2$ in Figure 10.

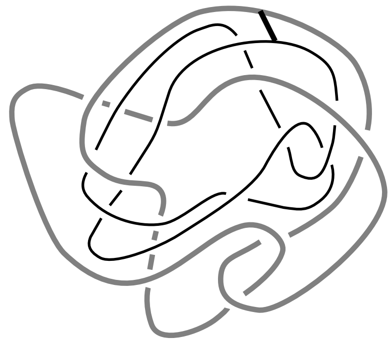


Figure 10.

Let $N(\Gamma)$ be a regular neighborhood of Γ. This is a genus n handlebody. We can assume that $N(\Gamma)$ is the union of n solid tori $N(\gamma_i)$, joined by $(n-1)$ 1-handles $N(\alpha_j)$.

Theorem (3.1). *Let Γ be a graph as above. Then $S = \partial N(\Gamma)$ is incompressible and acylindrical in $S^3 - intN(\Gamma)$. Furthermore, $M - intN(\Gamma)$ is atoroidal.*

Proof. Consider the tori $T\times\{e_i\}$, $1 \le i \le n-1$. These tori divide S^3 into $n+1$ regions, where $n-1$ of them are product regions and two of them are solid tori, namely R'_0 and R'_1. The torus $T\times\{e_i\}$ intersects Γ in one point, that is, a middle point of α_i, so $T\times\{e_i\}\cap N(\Gamma)$ consists of a disc. Let $T_i = T\times\{e_i\} - intN(\Gamma)$, for $1 \le i \le n-1$, this is a once punctured torus.

Suppose D is a compression disc for S, and suppose it intersects transversely the tori T_i. Let β be a simple closed curve of intersection between D and the collection of tori, which is innermost in D. So β bounds a disc $D' \subset D$, which is contained in a product $T\times[e_{i-1}, e_i]$, or in the solid torus R'_0 or in R'_1. If β is trivial on T_i, then by cutting D with an innermost disc lying in the disc bounded by β on T_i, we get a compression disc with fewer intersections with the $T_i's$. If β is essential on T_i, then it would be parallel to ∂T_i, or it would be a meridian of T_1 or a longitude of T_{n-1}, but then in any case, one of the curves γ_1 or γ_n will be contained in a 3-ball, which is a contradiction.

So suppose D intersects the $T_i's$ only in arcs. Let β such an arc which is outermost on D, then it cobounds with an arc $\delta \subset \partial D$ a disc D'. We can assume that β is an arc properly embedded in some T_i; if β is parallel to an arc on ∂T_i, then by cutting D with an outermost such arc lying on T_i we get another compression disc with fewer intersections with the $T_i's$, so assume that β is an essential arc on T_i. After isotoping D if necessary, we can assume that the arc δ can be decomposed as $\delta = \delta_1\cup\delta_2\cup\delta_3$, where δ_1, δ_3 lie on $\partial N(\alpha_i)$ and δ_2 lie on $\partial N(\gamma_i)$ (if δ were contained in $\partial N(\alpha_i)$, then by isotoping D we would get a compression disc intersecting T_i in a simple closed curve). Let E be a disc contained in $N(\alpha_i)$ so that $\partial E = \delta_1\cup\delta_4\cup\delta_3\cup\delta_5$, where δ_4 lies on T_i and δ_5 lies on $\partial N(\alpha_i)$. So $D'\cup E$ is an annulus, where one boundary component, i.e., $\beta\cup\delta_4$ lies on $T\times\{e_i\}$, and the other, $\delta_2\cup\delta_5$, lies on $\partial N(\gamma_i)$. If $\delta_2\cup\delta_5$ is a meridian of γ_i, then necessarily $D\cup E$ is contained in R'_0 (or in R'_1) and $\beta\cup\delta_4$ is a meridian of that solid torus. Then γ_1 (or γ_n) intersects a meridian disc of R'_0 (R'_1) in one point, which implies that it is parallel to a knot lying on the torus $T \times \{e_0\}$ ($T \times \{e_1\}$), which is a contradiction. If $\delta_2 \cup \delta_5$ is a longitudinal curve of γ_i, then this implies that γ_i is parallel to a curve on $T \times \{e_i\}$, a contradiction. If $\delta_2 \cup \delta_5$ goes more than once longitudinally on γ_i, this would only be possible for the curves γ_1 or γ_n, but then one of these curves would be a core of the solid torus R'_0 or R'_1, which is not possible. This completes the proof that S is incompressible in $S^3 - intN(\Gamma)$.

Suppose now that there is an essential annulus A in $S^3 - intN(\Gamma)$. Look at the intersection between A and the punctured tori T_i. Simple closed curves of intersection which are trivial on A, and arcs on A which are parallel to a component of ∂A are eliminated as above. So the intersection consists of a collection of essential arcs on A, or a collection of essential simple closed curves on A.

Suppose first that there are essential arcs of intersection. Let $E \subset A$ be a square determined by the arcs of intersection. So $\partial E = \epsilon_1 \cup \delta_1 \cup \epsilon_2 \cup \delta_2$, where ϵ_1, ϵ_2 are contained in different components of ∂A and δ_1, δ_2 are arcs of intersection of A with the $T'_i s$. Take the square at highest level. So δ_1, δ_2 lie on the same level T_i, and possibly $T_i = T_{n-1}$. So we can assume that ϵ_1, ϵ_2 lie on $\partial N(\alpha_i \cup \gamma_{i+1})$.

Case 1: The arcs δ_1, δ_2 are parallel on T_i, that is, they cobound a disc F in T_i.

There are two subcases, depending of the orientation of the arcs δ_1, δ_2. Give an orientation to ∂E. Suppose first that the arcs δ_1, δ_2 have the same orientation on T_i (note that the interior of F may intersect the annulus A, but it is irrelevant in this case). Then $E \cup F$ is a Möbius band, and by pushing it off T_i we get a Möbius band contained in the product $T \times [e_i, e_{i+1}]$ or in R'_1, with its boundary lying on $N(\gamma_i)$. This implies that either γ_i is a trivial knot or that it is a 2-cable of some knot, which is a contradiction.

Suppose the arcs δ_1, δ_2 have opposite orientations in T_i. If the interior of the disc F intersects A, then take another square in A, which determines a disc $F' \subset F$ with interior disjoint from A. We can form two annuli, $E \cup F$ and $(A - E) \cup F$. We will show that at least one of them is an essential annulus. Note that a core of A is homotopic to the product of a core of $E \cup F$ and a core of $(A - E) \cup F$. So if these two curves are homotopically trivial, so is the core of A. So assume one of them is incompressible, say $(A - E) \cup F$. If it is ∂-compressible then it is ∂-parallel, because S is incompressible. Then there is a ∂-compression disc for this annulus intersecting it on $(A - E)$, but this implies that the original annulus A is also ∂-compressible, a contradiction. So we get a new essential annulus with fewer intersection with the $T'_i s$.

Case 2: The arcs δ_1, δ_2 are not parallel on T_i, and the arcs ϵ_1, ϵ_2 are parallel on $\partial N(\Gamma)$.

The arcs ϵ_1, ϵ_2 must have the same orientation on $N(\alpha_i \cup \gamma_i)$, see Figure 11(a). They cobound a disc F on $\partial N(\alpha_i \cup \gamma_i)$ with $\partial F = \epsilon_1 \cup \eta_1 \cup \epsilon_2 \cup \eta_2$, where $\gamma_1, \gamma_2 \subset \partial N(\alpha_i) \cap T_i$. (Note that the disc F may intersect the arc α_{i+1}, or its interior may intersect A, but this is irrelevant in this argument). It follows that $E \cup F$ is a Möbius band whose boundary lies on T_i. This is impossible if the

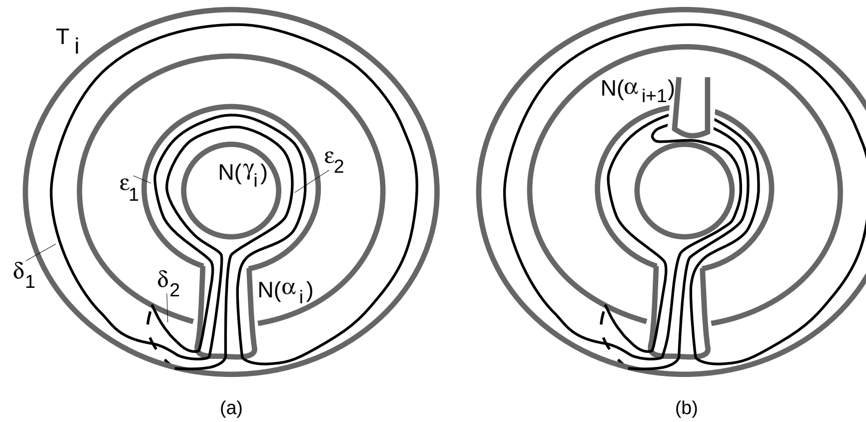


Figure 11.

band lies in a product region. If it lies in R_1', then note that the band is disjoint from the curve γ_n, but this is not possible, by hypothesis.

Case 3: The arcs δ_1, δ_2 are not parallel on T_i, and the arcs ϵ_1, ϵ_2 are not parallel on $\partial N(\Gamma)$.

Note that this case is only possible in a product region, see Figure 11(b). Forget about the arc α_{i+1}, that is, consider the square E in the complement of $N(\alpha_i \cup \gamma_i)$. Then, it is not difficult to see that one of the arcs, say ϵ_2 can be slid toward T_i. Then there is a disc, whose boundary consists of two arcs, one lying on T_i and one on $N(\gamma_i)$. By gluing to this disc a disc contained in $N(\alpha_i)$, an annulus between γ_i and $T \times e_i$ is constructed. The only possibility in this case is that the annulus goes once longitudinally on $N(\gamma_i)$, i.e., the curve γ_i is parallel to the torus $T \times e_i$, which is a contradiction.

This completes the proof in the case the annulus A is divided in squares.

Suppose now that the intersection of the annulus A with the tori $T_i's$ consists of simple closed curves which are essential on A. Take an outermost curve, say α. Then α and a component of ∂A cobound an annulus, and the component of ∂A must lie on some γ_i. This again implies that γ_i is parallel to T_i or that γ_1 or γ_n are the core of the solid torus R_0' or R_1', a contradiction.

It remains to prove that $S^3 - intN(\Gamma)$ is atoroidal. Suppose Q is an essential torus, then we can assume that it intersects the tori T_i in a collection of simple closed curves which are essential on Q, and divide Q in a collection of annuli. Take one of this annuli, say A, at highest level. If A is in a product region then it must be parallel to some T_i, and then by an isotopy we can remove two curves of intersection. So A lies on R_1'. As it is an annulus in a solid torus, it must be parallel to the boundary. If γ_n is not in this parallelism region, then an isotopy removes the intersection. If γ_n is the parallelism region, then take the annulus next to A. It must be an annulus between T_{n-1} and T_{n-2}. Continuing in this way, the only possibility is that the whole graph Γ lies inside a solid torus bounded by Q, but this is isotopic to a solid torus of the form $B \times I$, where B is an annulus in $T \times \{e_n\}$. This contradicts the choice of Γ. $\square$

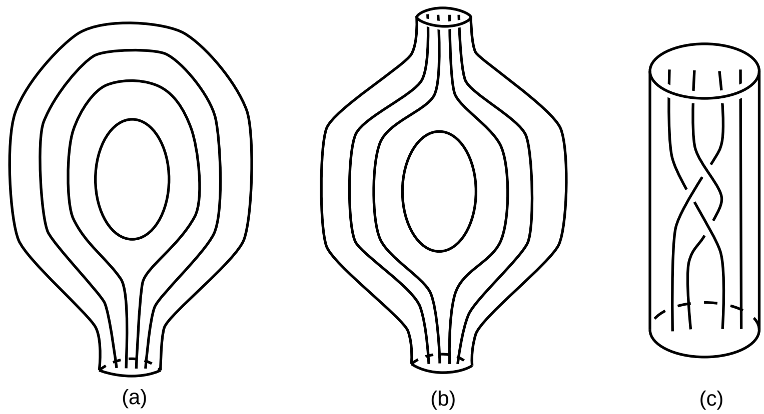


Figure 12.

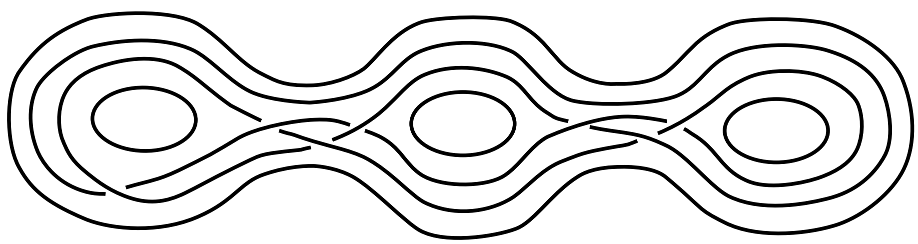


Figure 13.

Put now a knot K inside $N(\Gamma)$ in such a way that $K \cap N(\alpha_i)$, for $2 \leq i \leq n-1$, consists of four vertical arcs with a pattern like in Figure 12(c), and $k \cap N(\gamma_i)$ consists of 4 vertical arcs, going from $N(\alpha_i)$ to $N(\alpha_{i+1})$, as in Figure 12(b). Also, $K \cap N(\gamma_1)$ consists of two arcs, each having a single minimum, and $K \cap N(\gamma_n)$ consists of two arcs, each having a single maximum, as in the pattern shown in Figure 12(a). For $n = 3$, a knot K inside $N(\Gamma)$ looks like in Figure 13, where the twist is added to get a knot.

Lemma (3.2). *$S = \partial N(\Gamma)$ is acylindrical in $N(\Gamma) - K$. Furthermore, $N(\Gamma) - K$ is atoroidal.*

Proof. The proof is also an innermost disc/outermost arc argument. It is practically the same as in Lemma 2.3 of [3]. □

Theorem (3.3). *Let K and S as constructed above. K is a hyperbolic $(2,1)$-knot, tunnel number 2 knot, and S is an acylindrical surface of genus g in the complement of K.*

Proof. Note that by construction K is a $(2,1)$-knot, for it lies in $T \times I$, and it has in there exactly two maxima and two minima with respect to the projection to the factor I. It follows from Theorem (3.1) and Lemma (3.2) that S is an acylindrical surface. K is a hyperbolic knot because the complement of the surface is atoroidal and acylindrical. Finally note that the knot K has tunnel number 2; it cannot have tunnel number one, for it contains an acylindrical separating surface [19]. □

Corollary (3.4). *Given any integer $g \geq 2$, there exist infinitely many hyperbolic 3-manifolds of Heegaard genus 3 which contain an acylindrical surface of genus g.*

Proof. For each g choose a knot K as above. Do Dehn surgery on K with slope λ, such that $\Delta(\mu, \lambda) \geq 3$, where μ is a meridian of K. It follows that S remains incompressible [22], acylindrical [11], and that $M(\alpha)$ is irreducible and atoroidal [9] [10]. Then by Thurston Geometrization Theorem, $M(\alpha)$ is hyperbolic, for it is Haken and atoroidal. K has tunnel number two, which implies that $M(\alpha)$ has Heegaard genus at most 3, but it cannot have Heegaard genus 2, for it contains a separating acylindrical surface [19]. □

4. Quasi-Fuchsian surfaces of arbitrarily high genus

Let M be an irreducible orientable 3-manifold. Let K be a knot in M. Let $\mathcal{B}$ be a branched surface in M disjoint from K. (see [7] [20] for definitions and facts about branched surfaces). Denote by N a fibered regular neighborhood of $\mathcal{B}$, by $\partial_h N$ the horizontal boundary of N, and by $\partial_v N$ the vertical boundary of N, as usual.

We say that a branched surface $\mathcal{B}$ is incompressible in $M - K$ if it satisfies:

1. $\mathcal{B}$ has no discs of contact or half discs of contact.
2. $\partial_h N$ is incompressible and ∂-incompressible in $(M - K) - intN$.
3. There are no monogons in $(M - K) - intN$.

We further say that $\mathcal{B}$ is meridionally incompressible if:

4. $\partial_h N$ is meridionally incompressible, that is , there is no disc D in M, with $D \cap N = \partial D \subset \partial_h N$, so that D intersects K transversely in one point.

We further say that K is not parallel to $\mathcal{B}$ if:

5. K is not parallel to $\partial_h N$, that is, there is no an annulus A in M, with $\partial A = A_0 \cup A_1$, so that $A_0 = K$, and $A \cap N = A_1 \subset \partial_h N$

Theorem (4.1). *Let M, $\mathcal{B}$, K as above, with $\mathcal{B}$ incompressible.*

1. *Suppose $\mathcal{B}$ is meridionally incompressible. Then a surface carried with positive weights by $\mathcal{B}$ is meridionally incompressible.*
2. *If K is not parallel to $\mathcal{B}$, then K is not parallel to any surface carried with positive weights by $\mathcal{B}$.*

Then if $\mathcal{B}$ is meridionally incompressible and K is not parallel to it, any surface carried by $\mathcal{B}$ with positive weights is quasi-Fuchsian.

Proof. It is essentially the same proof as in Theorem (2.5) in [7], with the obvious modifications. □

Consider the knot K and branched surface $\mathcal{B}$ shown in Figure 14(a). Note that $\mathcal{B}$ has 4 singular curves, denoted C_1, C_2, C_3, C_4, as in Figure 14(b). Note that K is a 3-bridge knot. The knot K is just one in a collection of knots, to get more just make the knot to intersect several times the discs D_1, D_2, D_3, D_4 shown in Figure 14(b). But suppose that K intersects transversely the discs

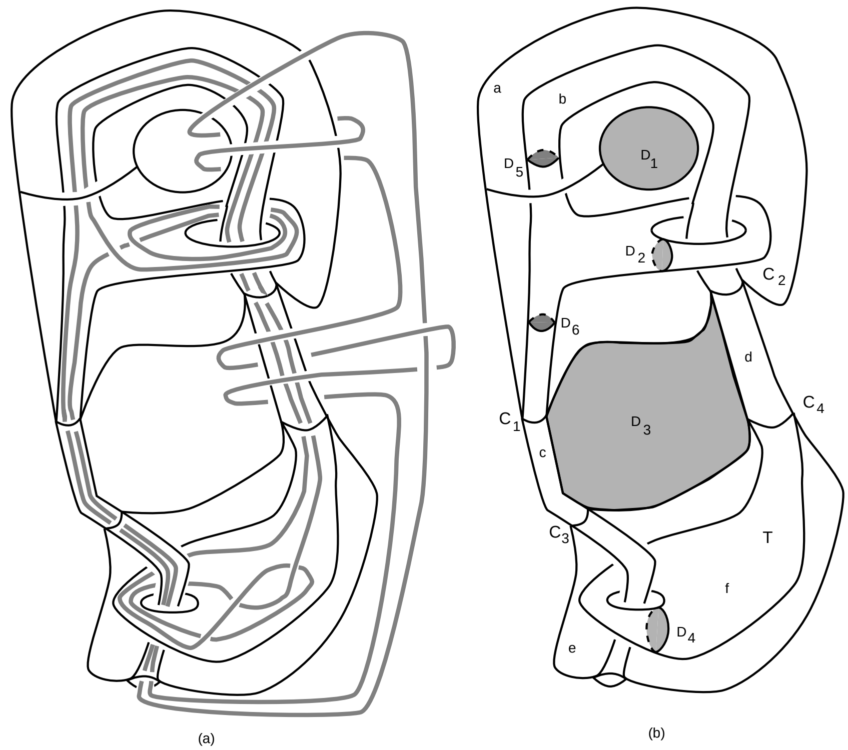


Figure 14.

D_1, D_2, D_3, D_4 in at least 2 points, that is, the minimal intersection number of the knots with the discs, when isotoping the knot in the complement of $\mathcal{B}$ is at least 2. Note that K intersects the discs D_5 and D_6 in exactly 2 points, because it is a 3-bridge knot. Suppose also that the arc of the knot lying in the solid torus T (shown in Figure 14(b)), is not parallel to ∂T; it is possible to do that, an explicit example is in Figure 14(a).

The nonsingular part of $\mathcal{B}$ has six components, whose weights (a, b, c, d, e, f) are shown in Figure 14(b). Note that if we give the weights $(1, 2n-1, 2n, 2n-2, n, n-2)$, for $n \geq 3$, then this is a collection of positive weights, which is consistent, and determines a connected surface of genus $3n$.

If a knot K is not hyperbolic then it is either a torus knot or a satellite knot. Remember that by the classical work of Schubert, a satellite 3-bridge knot must be the connected sum of 2 two-bridge knots. It is known that two-bridge knots do not contain any essential closed surface [13], and from this it follows that the only essential surfaces in the connected sum of 2 two-bridge knots are the swallow-follow tori. Also, torus knots do not contain closed essential surfaces. This implies that a 3-bridge knot which contains an essential surface of genus greater than 1 must be hyperbolic.

Theorem (4.2). *The surface* $\mathcal{B}$ *is meridionally incompressible and* K *is not parallel to it. So* K *is a hyperbolic 3-bridge knot which contains quasi-Fuchsian surfaces of arbitrarily high genus.*

Sketch of proof. Let N be a fibered neighborhood of $\mathcal{B}$. Note that $S^3 - N$ has 3 components, denoted by N_1, N_2, N_3, where say N_3 is the region that contains the knot, N_1 is the upper region, and N_2 the lower region.

Suppose that the part of $\partial_h N$ contained in N_3 is compressible or meridionally compressible, and let E be a compression or meridian compression disc. Look at the intersections between E and the discs D_1, D_3, D_5, D_6. Let γ be a simple closed curve of intersection which is innermost on E, so γ bounds a disc $E' \subset E$; suppose first that E' is disjoint from K. The curve γ also bounds a disc D' in some D_i. Suppose D' intersects K. If D' is part of D_1 or D_3, then K intersects the sphere $E' \cup D'$ several times always in the same direction, which is impossible. If D' is part of D_5 or D_6 then it must intersect K in two points, and then there is an arc of K contained in the 3-ball bounded by $E' \cup D'$. But this implies that K can be made disjoint from D_2 or D_4, or from D_3 or D_1, which is impossible by hypothesis. So D' must be disjoint from K, and then an isotopy reduces the number of intersections between E and the D_i. If E' intersects K once, then by a similar argument, D' intersects K also in a point, and then by an isotopy, we get a new compression disc with fewer intersections with the D_i. Suppose then that the intersection between E and the D_i consists only of arcs. Let γ be an arc of intersection which is outermost on E, and which bounds a disc E' disjoint from K. The arc γ also bounds a disc D' on some D_i. If K is disjoint from D', then by cutting E with an outermost disc lying on D' we get a new compression disc with fewer intersections with the D_i. If K intersects D' in one point, then it is not difficult to see that K must intersect in one point one of D_2, D_4 or D_1, which is a contradiction. So if there is such a disc E, it must be disjoint from the D_i, and by inspection it is not difficult to check that such disc does not exist. The part of $\partial_v N$ contained in N_3 consists of one annulus, corresponding to the curve C_2. Again an innermost disc/outermost arc argument shows that there is no monogon.

The part of $\partial_h N$ contained in N_1 consists of a twice punctured genus two surface; it is not difficult to check that it is incompressible. The part of $\partial_v N$ contained in N_1 consists of an annulus, corresponding to the curve C_1; it is also not difficult to check that there is no monogon. Similarly, the part of $\partial_h N$ contained in N_2 consists of a three punctured sphere and an once punctured torus, and $\partial_v N$ consists of two annuli, corresponding to the curves C_3 and C_4; again it is not difficult to check that these are incompressible and that there is no monogon.

To see that K is not parallel to $\mathcal{B}$, suppose there is an annulus A, with one boundary being K and the other on $\mathcal{B}$. Again look at the intersections between A and the discs D_i, and get that the arc of the knot that lies in the solid torus T must be parallel to ∂T, but this is not possible by the choice of such an arc. ☐

The explicit knot shown in Figure 14(a) has more interesting properties, it is a ribbon knot and it has unknotting number one, where a crossing change is located in the arc contained in the solid torus T.

Corollary (4.3). *There exist hyperbolic genus 3 closed 3-manifolds, in fact homology spheres, which contain incompressible surfaces of arbitrarily high genus, so contain infinitely many incompressible surfaces.*

Proof. Let K be a knot as in Theorem (4.2). Let $K(r)$ be the manifold obtained by performing Dehn surgery on K with slope r. If $\Delta(r, \mu) > 1$, where μ denotes a meridian of K, then $K(r)$ is irreducible by [9], and $\mathcal{B}$ remains incompressible in $K(r)$ by [22], for K is not parallel to $\mathcal{B}$. If $\Delta(r, \mu) > 2$, then $K(r)$ is atoroidal by [10]. So if $\Delta(r, \mu) > 2$, $K(r)$ is an atoroidal Haken manifold, hence it is hyperbolic. K is a tunnel number 2 knot, hence each $K(r)$ has Heegaard genus ≤ 3. Finally note that among the $K(r)$ many are homology spheres. □

Corollary (4.4). *There exist genus 2 closed 3-manifolds which contain incompressible surfaces of arbitrarily high genus, so they contain infinitely many incompressible surfaces.*

Proof. Let K be a knot as in Theorem (4.2). Let $\Sigma(K)$ denote the double cover of S^3 branched along K. As K is a 3-bridge knot, $\Sigma(K)$ has Heegaard genus 2. If S is a surface carried by $\mathcal{B}$ with positive weights, then as it is meridionally incompressible, it lifts in $\Sigma(K)$ to a (possible disconnected) incompressible surface [8]. □

Remark (4.5). It should be possible to say that the manifolds obtained in this corollary are hyperbolic; this will be the case if it is shown that the knots K do not admit a tangle decomposing sphere.

Received October 31, 2003

Final version received February 16, 2004

Instituto de Matemáticas, UNAM
Ciudad Universitaria
04510 México D.F.
México
mario@matem.unam.mx

max@matem.unam.mx

References

[1] I. Agol, *Lower bounds on volumes of hyperbolic Haken 3-manifolds*, arXiv:math.GT/9906182.

[2] C. Adams, J. Brock, J. Bugbee, et al, *Almost alternating links*, Topology Appl. **46** (1992), 151–165.

[3] C. Adams and A. W. Reid, *Quasi-Fuchsian surfaces in hyperbolic knot complements*, J. Austral. Math. Soc. (Series A) **55** (1993), 116–131.

[4] M. Boileau and R. Weidmann, *The structure of 3-manifolds with 2-generated fundamental group*, preprint, arXiv:math.GT/0111327.

[5] M. Eudave-Muñoz, *Essential meridional surfaces for tunnel number one knots*, Bol. Soc. Mat. Mexicana (3) **6** (2000), 263–277.

[6] E. Finkelstein and Y. Moriah, *Tubed incompressible surfaces in knot and link complements*, Topology Appl. **96** (1999), 153–170.

[7] W. Floyd and U. Oertel, *Incompressible surfaces via branched surfaces*, Topology **23** (1984), 117–125.

[8] C. McA. Gordon and R.A. Litherland, *Incompressible surfaces in branched coverings*, In The Smith Conjecture (New York 1979), Pure Appl. Math. 112, Academic Press (1984), 139–152.

[9] C.McA. Gordon and J. Luecke, *Only integral Dehn surgeries can yield reducible manifolds*, Math. Proc. Cambridge Philos. Soc. **102** (1987), 97–101.

[10] C.McA. Gordon and J. Luecke, *Dehn surgeries on knots creating essential tori.* I, Comm. Anal. Geom. **3** (1995), 597–644.

[11] C.McA. Gordon and Y-Q. Wu, *Annular and boundary reducing Dehn fillings*, Topology **39** (2000), 531–548.

[12] J. Hass, *Acylindrical surfaces in 3-manifolds*, Mich. Math. J. (1995), 357–365.

[13] A, Hatcher and W. Thurston, *Incompressible branched surfaces*, Invent. Math. **79** (1985), 225–246.

[14] J. Hempel, *3-manifolds as viewed from the curve complex*, Topology **40** (2001), 631–657.

[15] K. Ichihara and M. Ozawa, *Hyperbolic knot complements without closed embedded totally geodesic surfaces*, J. Austral. Math. Soc. (Series A) **68** (2000), 379–386.

[16] M. Lackenby (with an appendix by I. Agol and D. Thurston), *The volume of hyperbolic alternating link complements*, Proc. London Math. Soc. (3) **88** (2004), 204–224.

[17] H.C. Lyon, *Incompressible surfaces in knot spaces*, Trans. Amer. Math. Soc. **157** (1971), 53–62.

[18] W. Menasco, *Closed incompressible surfaces in alternating knot and link complements*, Topology **23** (1984), 37–44.

[19] W. Menasco and A. Reid, *Totally geodesic surfaces in hyperbolic link complements*, Topology '90, B. Apanosov, W. Neumann, A. Reid and L. Siebenmann, *editors*, De Gruyter, Amsterdam (1992), 215–226.

[20] U. Oertel, *Incompressible branched surfaces*, Invent. Math. **76** (1984), 385–410.

[21] W. Thurston, *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc. (N.S.) **6** (1982), 357–381.

[22] Y. Q. Wu, *Incompressibility of surfaces in surgered* 3*-manifolds*, Topology **31** (1992), 271–279.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

3-MANIFOLDS THAT ARE COVERED BY TWO OPEN BUNDLES

J. C. GÓMEZ-LARRAÑAGA, WOLFGANG HEIL, AND F. GONZÁLEZ-ACUÑA

Abstract. We obtain a list of all closed 3-manifolds that are covered by two open submanifolds, each homeomorphic to an open disk bundle over S^1, or an open I-bundle over the 2-sphere, the projective plane, the torus, or the Klein bottle.

0. Introduction

The F-category $F(M)$ of a closed 3-manifold M is the minimum number of critical points of smooth functions $M \longrightarrow R$. A lower bound for $F(M)$ is the Lusternik-Schnirelmann category $\operatorname{cat}(M)$ of M, which is a homotopy invariant and is defined to be the smallest number of sets, open and contractible in M, needed to cover M. An invariant that turns out to be equivalent to $F(M)$ is the smallest number $C(M)$ of open balls needed to cover M. Note that $2 \leq C(M), F(M), \operatorname{cat}(M) \leq 4$, and denote by $\mathcal{B}$ a connected sum of any number of S^2-bundles over S^1. Then the results about these three invariants can be summarized as follows (where $\simeq$ denotes homotopy equivalence): $F(M) = 2 \Leftrightarrow M = S^3$, $F(M) \leq 3 \Leftrightarrow M = \mathcal{B}$ (proved in [12]).

$C(M) = 2 \Leftrightarrow M = S^3$, $C(M) \leq 3 \Leftrightarrow M = \mathcal{B}$ (proved in [8]).

$\operatorname{cat}(M) = 2 \Leftrightarrow M \simeq S^3$, $\operatorname{cat}(M) \leq 3 \Leftrightarrow M \simeq \mathcal{B}$ (proved in [3]).

Generalization of these invariants were introduced by Clapp and Puppe [1] and Khimshiashvili and Siersma [9]: Let A be a closed k-manifold, $0 \leq k \leq 2$. A subset G in the 3-manifold M is A - *categorical* if the inclusion map $i : G \longrightarrow M$ factors homotopically through A. An A-*function* on M is a smooth function $M \longrightarrow R$ whose critical set is a finite disjoint union of components, each diffeomorphic to A. The A-*category* $\operatorname{cat}_A(M)$ of M is the smallest number of sets, open and A-categorical, needed to cover M. The A-*complexity* $F_A(M)$ of M is the minimum number of components of the critical set over all A-functions on M.

Then $\operatorname{cat}_{\text{point}}(M) = \operatorname{cat}(M)$, $F_{\text{point}}(M) = F(M)$, $\operatorname{cat}_{S^1}(M)$ is the *round category* of M, and $F_{S^1}(M)$ is the *round complexity* of M, studied in [9].

It is now natural to ask about minimal covers of M by open sets, each homotopy equivalent to A. In particular when A is a point, S^1, or a closed 2–manifold, consider covers of M by open disk bundles over A, i.e. open 3-balls, $\overset{\circ}{D}{}^2$-bundles over S^1, and $\overset{\circ}{I}$-bundles over surfaces. For such an open disk bundle $B(A)$ over

2000 *Mathematics Subject Classification*: 57M27.

Keywords and phrases: Lusternik-Schnirelmann category.

This work was partially supported by CONACYT grant 37352-E..

A let $C_{B(A)}(M)$ denote the minimal number of sets, each homeomorphic to $B(A)$, needed to cover M. In this paper we classify all closed 3-manifolds for which $C_{B(A)}(M) = 2$, where A is S^1, S^2, the projective plane P^2, the torus T, or the Klein bottle K. (Note that $C_{B(point)}(M) = C(M)$). The results are summarized in a table at the end of the paper. Some results are unexpected; for example the manifolds for which $C_{T\times \mathring{I}}(M) = 2$ include all lens spaces (including S^3), which can be seen as follows. Let $L_1 = l_1 \cup l_2$ be the Hopf link in S^3 and let l_i' be parallel to l_i so that $L_2 = l_1' \cup l_2'$ is a Hopf link disjoint to L_1. Then $S^3 = \left(S^3 - L_1\right) \cup \left(S^3 - L_2\right)$ is a union of two open $T \times \mathring{I}$'s. A similar construction can be made for any lens space.

1. Preliminaries

Throughout this paper we work in the PL-category. Our goal is to obtain information about closed 3-manifolds that are covered by open sets each of which is homeomorphic to the interior of a compact 3-manifold. Our main lemma shows that we can reduce the problem of a covering by two open sets to a canonical covering by two compact manifolds, each PL embedded.

(1.1) Main Lemma. *Suppose M is a closed 3-manifold covered by two open sets H_1, H_2, such that H_i is homeomorphic to the interior of a compact connected 3-manifold V_i $(i = 1, 2)$. Then M admits a covering $M = V_1 \cup V_2$ such that $\partial V_1 \cap \partial V_2 = \emptyset$ and V_1, V_2 are PL embedded.*

Proof. Using collars on ∂V_i $(i = 1, 2)$, we can write $H_i = \bigcup_{k=1}^{\infty} \operatorname{int} V_k^{(i)}$, where $V_k^{(i)} \approx V_i$, $V_k^{(i)} \subset \operatorname{int} V_{k+1}^{(i)}$, $k = 1, 2, \ldots$. The complement H_1^c of H_1 in M is a compact subspace of H_2, and it follows that $H_1^c \subset \operatorname{int} V_n^{(2)}$ for some n. Now, $\left(\operatorname{int} V_n^{(2)}\right)^c$ is a compact subspace of H_1 and hence $\left(\operatorname{int} V_n^{(2)}\right)^c \subset \operatorname{int} V_m^{(1)}$ for some m. Note that $\partial V_n^{(2)} \subset \left(\operatorname{int} V_n^{(2)}\right)^c \subset V_m^{(1)}$. Hence if we let $V_1 = V_m^{(1)}$ and $V_2 = V_n^{(2)}$ in M we obtain $M = V_1 \cup V_2$ as desired. □

By a *knot space* we mean a 3-manifold N homeomorphic to the complement of the interior of a regular neighborhood of a non-trivial knot in S^3. Note that ∂N contains a meridian curve C such that attaching a 2-handle to N with core along C yields B^3. The next lemma is well-known.

Lemma (1.1.1). *Suppose M is a compact irreducible 3-manifold.*

(i) *If M contains a 2-sided compressible torus T then either T bounds a solid torus or a knot space N in M with an essential curve of ∂N bounding a disk in $\overline{M - N}$. If T is a compressible boundary component of M then $M = D^2 \times S^1$.*

(ii) *If M contains a 2-sided compressible Klein bottle K then either K bounds a solid Klein bottle in M or M contains a 2-sided projective plane P^2. If K is a compressible boundary component then M is a solid Klein bottle.*

Proof. (i) Let $D\times[-1,1]$ be a neighborhood of a compressing disk $D = D\times\{0\}$ with $D\times[-1,1]\cap T = \partial D\times[-1,1]\cap T$. The sphere $S = (T - D\times[-1,1]\cap T)\cup D\times\{-1\}\cup D\times\{1\}$ bounds a ball B in M. If $D\cap B = \emptyset$ then $B\cup D\times[-1,1]$ is a solid torus in M bounded by T. If $D\subset B$ then $T\subset B$ such that $\partial B\cap T$ is an essential annulus of T. Hence $\overline{B - D\times[-1,1]}$ is a knot space (or a solid torus) in M bounded by T.

(ii) If we surger K as above along a compressing disk D we obtain a 2-sphere S if ∂D does not separate K. Then $B\cup D\times[-1,1]$ is a solid Klein bottle bounded by K. (The case $D\subset B$ can not happen since a Klein bottle does not imbed in a ball). If ∂D separates K into two Moebius bands then $(K - D\times[-1,1]\cap K)\cup D\times\{-1\}\cup D\times\{1\}$ gives two 2-sided P^2's in M. $\square$

Notation. By $B\tilde{\times}F$ we denote a twisted F-bundle over B, not homeomorphic to $B\times F$. In particular, $S^1\tilde{\times}D^2$ is the solid Klein bottle, $S^1\tilde{\times}S^2$ is the non-orientable S^2-bundle over S^1, and $P^2\tilde{\times}I$ is the once-punctured projective space P^3. The twisted I-bundles over a torus T and a Klein bottle K are described in the next section.

The union of two 3-manifolds N_1, N_2 glued together along boundary components is denoted by $N_1\cup_\partial N_2$.

L denotes any lens space (including S^3 and $S^1\times S^2$).

$S(2,2,n)$ denotes a Seifert fiber space over the 2-sphere with three exceptional fibers of orders $2,2,n$ $(n\geq 0)$.

The symbol $\sim$ means *homologous to.*

The symbol $\approx$ means *homeomorphic.*

2. I-bundles and (semi)-bundles over the torus and Klein bottle

Recall that an I-bundle over a surface F is *twisted* if it is not the product I-bundle $F\times I$. The twisted I-bundle $a^2\tilde{\times}I$ over the annulus a^2 is homeomorphic to the product I-bundle $m^2\times I$ over the Moebius band m^2. The twisted I-bundle $m^2\tilde{\times}I$ over m^2 is homeomorphic to the solid torus $D^2\times I$ (with m^2 embedded in $D^2\times I$ so that ∂m^2 is a $(1,2)$-curve on $\partial D^2\times S^1$).

(2.1). There is only one twisted I-bundle $T\tilde{\times}I = m^2\times S^1$ over the torus $T = S^1\times S^1$.

To see this, note that in such an I-bundle N there is a simple closed curve c on T such that the restriction of the I-bundle over c is a Moebius band. Now c cuts T into an annulus a^2 and the restriction of the I-bundle over a^2 is twisted. Hence N is the quotient $m^2\times I/(x,0)\sim(\varphi(x),1)$ for a homeomorphism φ of m^2. If φ is isotopic to the identity then $N = m^2\times S^1$. The case that φ is not isotopic to the identity can not occur since then φ would reverse the orientation of ∂m^2, which would cause ∂N to be a Klein bottle; but ∂N is a torus since it is a 2-sheeted cover of T.

(2.2). There are exactly two twisted I-bundles over the Klein bottle $K = S^1\tilde{\times}S^1$.

These can be described as follows. The restriction of such an I-bundle N over a separating simple closed curve on K splits N into two I-bundles over Moebius bands m_1^2, m_2^2, at least one of which is twisted. There are two possibilities.

(i) $N = m_1^2 \widetilde{\times} I \cup m_2^2 \widetilde{\times} I$ is a union of two solid tori along an annulus in their boundary and N can be described as a Seifert fiber space with orbit a disk and two exceptional fibers of order 2. In this case N is orientable and is denoted by $\left(K \widetilde{\times} I\right)_0$.

(ii) $N = m_1^2 \times I \cup m_2^2 \widetilde{\times} I$, where $\partial m_1^2 \times I$ is identified with an annular neighborhood of ∂m_2^2 in $\partial D^2 \times S^1 = \partial\left(m_2^2 \widetilde{\times} I\right)$. In this case ∂N is a Klein bottle, and we denote this I-bundle over K by $\left(K \widetilde{\times} I\right)_{N_0}$.

Another description of $\left(K \widetilde{\times} I\right)_{N_0}$ is obtained by cutting K into an annulus along a 2-sided non-separating circle. As for $T \widetilde{\times} I$ we obtain $\left(K \widetilde{\times} I\right)_{N_0}$ as the quotient $m^2 \times I / (x, 0) \sim (\psi(x), 1)$, where ψ is not isotopic to the identity. Viewing m^2 as a rectangle with a pair of opposite edges identified, ψ is induced by a reflection about a line mid-way between the two edges (cf [10]). Thus $\left(K \widetilde{\times} I\right)_{N_0} \approx S^1 \widetilde{\times} m^2$, the twisted m^2-bundle over S^1.

Following Hatcher [4], we call a union of two twisted I-bundles over a torus T (resp. Klein bottle K) glued together along their boundary component a torus (resp. Klein bottle) *semi-bundle.* These semi-bundles are essentially classified by the isotopy classes of their gluing maps (see e.g. [4, Thm 5.1]).

There are exactly four isotopy classes of homeomorphisms of the Klein bottle ([10]) that lead to exactly four Klein bottle-bundles over S^1, described in [6].

3. Covers by $\operatorname{int} M_1$ and $\operatorname{int} M_2$

In this and the following sections we consider a closed 3-manifold M that is covered by two open sets $\operatorname{int} M_1$, $\operatorname{int} M_2$, where M_1, M_2 are homeomorphic compact connected 3-manifolds. In light of the Main Lemma, we may assume throughout that

$$(*) \qquad M = M_1 \cup M_2,\ M_1 \approx M_2 \text{ compact}, \partial M_1 \cap \partial M_2 = \emptyset.$$

We let $Q = M_1 \cap M_2 \subset M$. Note that the boundary of each component of Q contains a component of both ∂M_1 and ∂M_2. We observe

(i) If M_1, M_2 are irreducible then $\overline{M_i - Q}$ is irreducible ($i = 1, 2$).

For a 2-sphere in $\operatorname{int}\left(\overline{M_1 - Q}\right)$ bounds a ball B in $\operatorname{int}(M_1)$. If B does not lie in $\overline{M_1 - Q}$ then B contains a component of Q, hence a component of ∂M_1, a contradiction.

(ii) If M_1, M_2 are irreducible and $M \neq S^3$ then Q is irreducible.

For a 2-sphere S in Q bounds balls $B_1 \subset M_1$, $B_2 \subset M_2$. Either $B_1 = B_2 \subset Q$, or $B_1 \cap B_2 = S$ and $M = B_1 \cup_\partial B_2 = S^3$.

(3.1) Covers by open balls and open disk bundles over S^1.

(a) If $M_i \approx B^3$ then $M = S^3$.

Proof. ∂M_2 bounds a ball B in M_1 and $M = M_2 \cup_\partial B = S^3$. □

(b) If $M_i = S^1 \times D^2$ then $M = L$.

Proof. Since M_1 does not contain a closed incompressible surface, there is a compressing disk D for ∂M_2 in M_1. If $D \subset \overline{M_1 - Q}$ then $\overline{M_1 - Q}$ is a solid torus (by Lemma (1.1.1) (i) and (3) (i)) and $M = \overline{M_1 - Q} \cup_\partial M_2$ is a lens space.

If $D \subset Q$ then, viewing a regular neighborhood of D in Q as a 2-handle $U(D)$, we get that $\overline{M_1 - Q} \cup U(D) \subset M_1$ is bounded by a 2-sphere. Hence $M = \left(\overline{M_1 - Q} \cup U(D)\right) \cup_\partial \left(\overline{M_2 - U(D)}\right)$ is a union of two balls, i.e. $M = S^3$. □

(c) If $M_i \approx S^1 \tilde{\times} D^2$ then $M = S^1 \tilde{\times} S^2$.

Proof. ∂M_2 is compressible in M_1 and M_1 does not contain a projective plane. By Lemma (1.1.1) (ii), ∂M_2 bounds a solid Klein bottle $M_1' \subset M_1$ and $M = M_1' \cup_\partial M_2 = S^2 \tilde{\times} S^1$ (see e.g. [7, 2.14]). □

(3.2) Covers by open I-bundles over S^2 or P^2.

(a) If $M_i \approx S^2 \times I$ then $M = S^3$, $S^1 \times S^2$ or $S^1 \tilde{\times} S^2$.

Proof. Let $\partial M_2 = S_0 \cup S_1 \subset \operatorname{int} M_1$.

If S_0 bounds a ball B_0 in M_1 then $B_0 \subset \overline{M_1 - Q}$ since M is closed. Now $M_2' = M_2 \cup_\partial B_0$ is a ball and $M = M_1 \cup M_2'$. The boundary S_1 of M_2' is not isotopic to a boundary sphere of M_1 (since M is closed) and hence bounds a ball B_1 in M_1, different from M_2' and $M = M_2' \cup_\partial B_1 = S^3$.

If both S_0 and S_1 are parallel to the boundary spheres of M_1 then S_0 and S_1 bound a submanifold $M_2' \approx S^2 \times I$ in M_1 and we obtain $M = \overline{M_1 - M_2'} \cup_\partial M_2'$, hence $M = S^1 \times S^2$ or $S^1 \tilde{\times} S^2$. □

(b) If $M_i = P^2 \times I$ then $M = P^2 \times S^1$.

Proof. This follows from the fact that any projective plane in M_1 is isotopic to a boundary component, hence $M \approx M_1 \cup_\partial M_2$. (Note that there is no twisted P^2-bundle over S^1). □

(c) If $M_i = P^2 \tilde{\times} I$ then $M = P^3$ or $P^3 \# P^3$.

Proof. If ∂M_2 bounds a ball B in M_1 then $M = M_2 \cup_\partial B = P^3$. Otherwise ∂M_2 is parallel in M_1 to ∂M_1 and $M \approx M_1 \cup_\partial M_2 = P^3 \# P^3$. □

(3.3) Covers by open I-bundles over $S^1 \times S^1$ and $S^1 \tilde{\times} S^1$.

Let $T = S^1 \times S^1$ and $K = S^1 \tilde{\times} S^1$.

(a) If $M_i = T \times I$ then $M = L$ or a T-bundle over S^1.

Proof. Let $\partial M_2 = T_0 \cup T_1$, $\partial M_1 = T_0' \cup T_1'$.

If T_0 is incompressible in M_1 then it is isotopic to a component of ∂M_1 and splits M_1 into two copies M_1', M_1''. Assume $T_0' \subset M_1'', T_1' \subset M_1'$. Then $T_1 \subset M_1'$, say. Then (since $T_0', T_0 \subset \partial Q$ and $T_0' \subset \operatorname{int} M_2$) it follows that M_1'' is a component of $Q \subset M_1$. The other component(s) of Q lie in M_1' and are bounded by T_1' and T_1. Since $T_1' \not\sim 0$ in M_1' there is exactly one component P of Q in M_1' bounded by T_1' and T_1. Hence $T_1 \not\sim 0$ in M_1' and Lemma (1.1.1) (i) implies that T_1 is incompressible in M_1'. Hence T_0, T_1 are isotopic in M_1 to T_0', T_1' and it follows that $M \approx M_1 \cup_\partial M_2$ is a T-bundle over S^1.

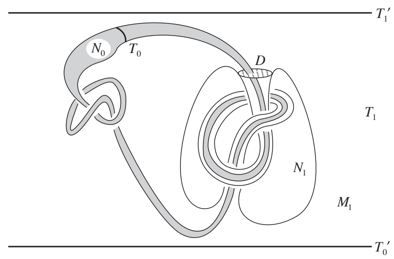


Figure 1

Now suppose that T_0, T_1 are both compressible in M_1; hence, by Lemma (1.1.1) (i), T_i bounds a solid torus or knot space N_i in M_1 (i=0,1). Now T_1 is not contained in N_0. Otherwise an arc in M_2 from a point of T_0 to a point of T_1 would be in N_0 (since T_0 separates in M_1), and it would follow that $M_2 \subset N_0 \subset M_1$, a contradiction. Similarly T_0 is not contained in N_1; hence N_0 and N_1 are disjoint. If N_0 is a solid torus then $M_2' = M_2 \cup_\partial N_0$ is a solid torus and $M = M_2' \cup_\partial N_1$. Thus if N_1 is also a solid torus, M is a lens space. If N_1 is a knot space then a meridian circle on ∂N_1 bounds a compressing disk D for M_2' in $\overline{M_1 - N_1}$ (see Figure (1)).

For a regular neighborhood $U(D)$ in $\overline{M_1 - N_1}$ we obtain $M = \overline{M_2' - U(D)} \cup_\partial \overline{N_1 \cup U(D)}$, a union of two balls, hence $M = S^3$.

The case where both N_0 and N_1 are knot spaces in M_1 can not occur. For in this case a compressing disk D for T_1 in $\overline{M_1 - N_1}$ must intersect N_0, since otherwise D would be a compressing disk for T_1 in M_2. But then an essential innermost circle component of $T_0 \cap D$ bounds a disk D' on D which would be a compressing disk for T_0 in N_0 or in M_2, a contradiction. $\square$

(b) If $M_i = K \times I$ then $M = S^1 \widetilde{\times} S^2$ or a K-bundle over S^1.

Proof. Let $\partial M_2 = K_0 \cup K_1 \subset \operatorname{int} M_1$.

If K_0 is compressible in M_1 then it bounds a solid Klein bottle V_0 in M_1 (by Lemma (1.1.1) (ii), since M_1 does not contain P^2's). The same argument as in case (a) shows that K_1 is also compressible and bounds a solid Klein bottle V_1 in M_1 such that V_0 and V_1 are disjoint. Then $M = (M_2 \cup_\partial V_0) \cup_\partial V_1$ is a union of two solid Klein bottles, hence $M = S^1 \widetilde{\times} S^2$.

If both K_0, K_1 are incompressible in M_1 then they are boundary parallel and $M = M_1 \cup M_2$ is a K–bundle over S^1. $\square$

We next consider the cases of twisted I-bundles over T and K.

Lemma (3.3.1). *Let M_i be a twisted I-bundle over T or K $(i = 1, 2)$.*

(i) *If ∂M_1 is incompressible in M_2 then $M \approx M_1 \cup_\partial M_2$ is a semi-bundle.*

(ii) *If ∂M_1 is compressible in M_2 then $M = M_2 \cup_\partial \left(S^1 \times D^2\right)$ (for $M_i = T \widetilde{\times} I$ or $\left(K \widetilde{\times} I\right)_0$), resp. $M = M_2 \cup_\partial \left(S^1 \widetilde{\times} D^2\right)$, (for $M_i = \left(K \widetilde{\times} I\right)_{N_0}$).*

Proof. If ∂M_1 is incompressible M_2 then it is parallel to ∂M_2 in M_2 and $M \approx M_1 \cup_\partial M_2$.

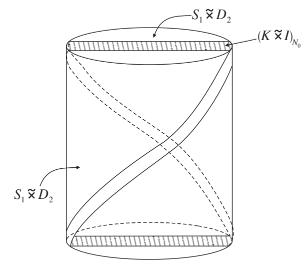


Figure 2

If ∂M_1 compresses in M_2 then it bounds a solid torus, a knot space, or a solid Klein bottle in M_2 (by Lemma (1.1.1)). It can not bound a knot space N since otherwise a meridian of ∂N would bound a compressing disk D in $\overline{M_2 - N} \subset Q$ and hence D would be a compressing disk for ∂M_1 in M_1. It follows that $M = M_2 \cup_\partial \left(S^1 \times D^2\right)$ or $M_2 \cup_\partial \left(S^1 \widetilde{\times} D^2\right)$. $\square$

(c) If $M_i = T \widetilde{\times} I$ then M is a torus semi-bundle or $M = P^2 \times S^1$ or $M = S^1 \widetilde{\times} S^2$.

Proof. By the previous lemma it suffices to consider the case where $M = M_2 \cup_\partial \left(S^1 \times D^2\right)$.

In the 2-sheeted orientable cover $\widetilde{M}$ of M, $M_2 = m^2 \times S^1$ lifts to $a^2 \times S^1 = T \times I$ and the attaching solid torus $S^1 \times D^2$ lifts to two attaching solid tori. Hence $\widetilde{M}$ is a lens space; its fundamental group is infinite, since it covers the closed non-orientable manifold M. By the classification of (orientation-reversing) fixed point free involutions on $S^1 \times S^2$ ([13], [14, Corollary]), M is as claimed. $\square$

(d) If $M_i = \left(K \widetilde{\times} I\right)_0$ then M is a Klein bottle semi-bundle or $M = P^3 \# P^3$ or $M = S(2, 2, n)$ (for any $n \geq 0$).

Proof. Again we need to consider only the case that $M = M_2 \cup_\partial \left(S^1 \times D^2\right)$. Representing M_2 as a Seifert fiber space over a disk with two exceptional fibers each of order 2 we obtain $M = S(2, 2, n)$ if the meridian ∂D^2 of the attaching solid torus is not homotopic to a fiber on ∂M_2, and $M = P^3 \# P^3$ otherwise (see e.g. [5]). $\square$

(e) If $M_i = \left(K \widetilde{\times} I\right)_{N_0}$ then M is a Klein bottle semi-bundle or $M = P^2 \times S^1$.

Proof. Considering only the case $M = M_2 \cup_\partial \left(S^1 \widetilde{\times} D^2\right)$, we represent $M_2 = S^1 \widetilde{\times} m^2$ (as in section 2) and note that ∂m^2 cuts $\partial M_2 = S^1 \widetilde{\times} \partial m^2$ into an annulus. Up to isotopy, there is only one simple closed curve on K that cuts K into an annulus ([10]). Thus there is only one way to attach $S^1 \widetilde{\times} D^2$ to M_2 : the meridian ∂D^2 of $S^1 \widetilde{\times} D^2$ must be glued to ∂m^2, and it follows that $M = \left(S^1 \widetilde{\times} m^2\right) \cup_\partial \left(S^1 \widetilde{\times} D^2\right) = S^1 \widetilde{\times} P^2 = S^1 \times P^2$. $\square$

Figure (2) shows that $P^2 \times S^1$ indeed admits a decomposition of type $\left(K \widetilde{\times} I\right)_{N_0} \cup_\partial \left(S^1 \widetilde{\times} D^2\right)$.

The following table summarizes the results.

$M = \operatorname{int} M_1 \cup \operatorname{int} M_2$

M_i	B^3	$S^1 \times D^2$	$S^1 \widetilde{\times} D^2$	$S^2 \times I$	$P^2 \times I$	$P^2 \widetilde{\times} I$
M	S^3	L	$S^1 \widetilde{\times} S^2$	S^3 $S^1 \times S^2$ $S^1 \widetilde{\times} S^2$	$P^2 \times S^1$	P^3 $P^3 \# P^3$

M_i	$T \times I$	$T \widetilde{\times} I$	$K \times I$	$\left(K \widetilde{\times} I\right)_0$	$\left(K \widetilde{\times} I\right)_{N_0}$
M	L T-bundles over S^1	$S^1 \widetilde{\times} S^2$ $P^2 \times S^1$ T-semi bundles	$S^1 \widetilde{\times} S^2$ K-bundles over S^1	$P^3 \# P^3$ $S\,(2,2,n)$ K-semi bundles (orientable)	$P^2 \times S^1$ K-semi bundles (non orientable)

Conversely, it is easy to see that each manifold in the table is a union of two open covers as indicated.

Received July 14, 2003

Final version received May 25, 2004

J. C. Gómez-Larrañaga
Centro de Investigación en Matemáticas
Apdo. Postal 402
36000 Guanajuato, Gto.
México
jcarlos@cimat.mx

Wolfgang Heil
Department of Mathematics
Florida State University
Tallahasee, FL 32306
USA
heil@zeno.math.fsu.edu

F. González-Acuña
Instituto de Matemáticas, UNAM
Ciudad Universitaria
04510 México, D.F.
México
fico@math.unam.mx

References

[1] M. Clapp and D.Puppe, *Invariants of the Lusternik-Schnirelmann type and the topology of critical sets*, Trans. Amer. Math. Soc. **298** (1986), 603–620.

[2] J.C. Gómez-Larrañaga, 3-*manifolds which are unions of three solid tori*, Manuscripta Math. **59** (1987), 325–330.

[3] J. C. Gómez-Larrañaga and F. González-Acuña, *Lusternik–Schnirelmann category of 3–manifolds*, Topology **31** (1992), 791–800.

[4] A. Hatcher, *Notes on basic 3-manifold Topology*, Lecture Notes.

[5] W. Heil, *Elementary surgery on Seifert fiber spaces*, Yokohama Math. J. **23** (1974), 135–139.

[6] W. Heil and P. Raspopović, *Pedja Dehn fillings of Klein bottle bundles*, Math. Proc. Cambridge Philos. Soc. **112** (1992), 255–270.

[7] J. Hempel, 3-Manifolds, Ann. of Math. Stud. **86** Princeton Univ. Press, Princeton, NJ, 1976.

[8] J. Hempel and D. R. McMillan, *Covering three–manifolds with open cells*, Fund. Math. **64** (1969), 99–104.

[9] G. Khimshiashvili and D. Siersma, *On minimal round functions*, preprint 1118, Department of Mathematics, University of Utrecht (1999). Bulletin of the Georgian Academy of Science, **164** (2000), 3–7.

[10] W. B. R. Lickorish, *Homeomorphisms of non-orientable two-manifolds*, Proc. Cambridge Philos. Soc. **59** (1963), 307–317

[11] Kanji Morimoto, *Some orientable 3-manifolds containing Klein bottles*, Kobe J. Math. **2** (1985), 37–44

[12] F. Takens, *The minimal number of critical points of a function on a compact manifold and the Lusternik–Schnirlemann category*, Invent. Math. **6** (1968), 197–244.

[13] Y.Tao, *On fixed point free involutions of* $S^1 \times S^2$, Osaka Math. J. **14** (1962), 145–152.

[14] J. Tollefson, *Involutions on* $S^1 \times S^2$ *and other 3-manifolds*, Trans. Amer. Math. Soc. **183** (1973), 139–152.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

SELF-COINCIDENCE OF MAPS FROM S^q-BUNDLES OVER S^n TO S^n

DACIBERG L. GONÇALVES AND DUANE RANDALL

Abstract. Let E be the total space of a sphere bundle over a sphere S^n. We investigate the problem of when a pair of maps $(f_1, f_2) : E \to S^n$ can be deformed to a coincidence free pair. Special attention is given to the case $f_1 = f_2$, *called the self-coincidence case*. We study this case from the point of view of an arbitrary deformation of f and of a small deformation of f. We show that if $(f, f) : E \to S^n$ can be deformed to a coincidence free pair by a small deformation and f is a fibre map, then the Euler number of the fibration is 2 and the fibration is fibre homotopy equivalent to the Stiefel fibration $p_1 : V_{2n+1,2} \to S^{2n}$. We study the coincidence problem (i.e. an arbitrary pair (f_1, f_2)) in more detail when the total space is one of the spaces, S^3, S^7, S^{15}, $V_{5,2}$,$V_{9,2}$. For many other cases of the domain we show that the problem can be reduced to a problem of maps either between spheres or from a complex with two cells of positive dimension into a sphere. For maps $f : S^m \to S^n$ and $m < 2n-1$ we classify the pairs (f_1, f_2) which can be deformed to a coincidence free pair. We construct maps $f_{2n} : S^{4n-1} \to S^{2n}$ for all odd $n > 1$ for which (f_{2n}, f_{2n}) can be deformed to a coincidence free pair, but not by a small deformation.

1. Introduction

Let M and N be closed manifolds of dimensions m and n, respectively. One of the major problems in coincidence theory is to determine when a given pair of maps $f_1, f_2 : M \to N$ can be deformed to a pair (g_1, g_2) which is coincidence free. The more general question is to describe the minimal (in some sense) set $coin(g_1, g_2)$ obtained among all pairs (g_1, g_2) homotopic to f_1, f_2, respectively. This problem is well understood for the cases where $dim(M) = dim(N) \geq 3$. See, for example, [14] and [7] for M, N orientable manifolds and nonorientable, respectively. For some specific families of manifolds, see [4],[5], [18] and [8], and for the cases where M is a CW-complex and $dim(M) = dim(N)$, see [1], [6]. The case where $m > n$ is more subtle. Several aspects of this problem have been considered in [9], [2], [12], [3], [11], [19]. In [2] the coincidence problem for the case $f_1 = f_2$ was analyzed via homotopy theory. Also U. Koschorke in [12] has considered the same case under the hypothesis that $m < 2n-2$ (which we call *the stable case*).

This work uses the approach and some of the main results of [2]. For a given map f, we investigate whether the pair (f, f) can be deformed to a coincidence free map pair, and also if there exists a small deformation. This latter question

2000 *Mathematics Subject Classification*: Primary: 55M20; Secondary: 55R25, 55S35.
Keywords and phrases: Self-coincidence, obstruction, fibrations, Whitehead product, Barratt-Puppe sequence.

is interesting in its own right and often appears in analysis. Let us point out that such phenomenon does not occur in self-coincidence of codimension zero, where the two concepts are equivalent. In [2], a homotopy approach is developed to study these two questions. For the benefit of the reader, we now state the main results from [2] used here. From [2], we have the following notation: given any two arbitrary maps, $f_1, f_2 : E \to B$, we say that f_1 *can be separated from* f_2, *or* (f_1, f_2) *is homotopy disjoint,* in symbols $f_1 \wr f_2$, if there exist $q_1, q_2 : E \to B$ such that q_1, q_2 are homotopic to f_1 and f_2, respectively, and $q_1(x) \neq q_2(x)$ for all x; otherwise, we write $f_1 \not\wr f_2$. Given $p : E \to B$, denote by $\Gamma_p \subset E \times B$ the graph of p. *A pair of maps* $(f, f) : E \to B$ *is homotopy disjoint by small deformation* in [2] if $N\Gamma_f$, the normal bundle of Γ_f, admits a nowhere-zero cross section. It follows from the above definitions that if a pair (f, f) is homotopy disjoint by a small deformation, then it is homotopy disjoint. *If the pair is homotopy disjoint by small deformation, then we simply say that the pair can be deformed to be coincidence free by a small deformation.* Proposition (2.11) of [2] identifies the primary obstruction to deforming a pair (f, f) to coincidence free and affirms:

Proposition (1.1). *If B is a compact, connected n-dimensional manifold and $f : E \to B$, then the primary obstruction to lift (f, f) in*

$$\begin{array}{ccc} & & B \times B - \Delta \\ & & \downarrow \\ E & \xrightarrow{(f,f)} & B \times B \end{array}$$

by deformation, is the f^*-image of the primary obstruction to lift (id, id) in

$$\begin{array}{ccc} & & B \times B - \Delta \\ & & \downarrow \\ B & \xrightarrow{(id,id)} & B \times B \end{array}$$

by deformation. The latter is the twisted Euler class of B, i.e. $= \chi(B) \cdot \mu_B$, where $\chi(B)$ is the Euler characteristic of B and $\mu_B = \pi_1 B$-twisted fundamental class.

For a pair of maps (f_1, f_2) into a sphere, Proposition (2.10) of [2] affirms:

Proposition (1.2). *If $B = S^n$, then $f_1 \wr f_2$ implies that $A \circ f_2$ is homotopic to f_1, where A is the antipodal map on S^n.*

Let τ_B be the tangent bundle of the differentiable manifold B, $S(\tau_B)$ the sphere bundle and $q : S(\tau_B) \to B$ the projection map.

In order to have a small deformation, Proposition (2.13) and (2.16) of [2] state, respectively:

Proposition (1.3). *The map $f : E \to B$ admits a lift to $S(\tau_B)$ if and only if (f, f) is homotopy disjoint by small deformation.*

Proposition (1.4). *If $f : E \to B$ is a map where B is a smooth manifold then the map $f : E \to B$ admits a lift to $S(\tau_B)$ if and only if the horizontal tangent bundle $f^*(\tau_B)$ over E has a nowhere-zero cross section. Therefore we conclude that (f, f) is homotopy disjoint by small deformation if and only if the horizontal tangent bundle $f^*(\tau_B)$ over E has a nowhere-zero cross section.*

In this work we consider the coincidence problem for maps $f : E \to S^n$, where E is a S^q-bundle over S^n, and for maps between spheres. Special attention is given to the case of self-coincidence. This paper is divided into 3 additional sections. In section 2 we prove some generalities and reduce the problem to a problem of maps either between spheres or from a two cell complex into a sphere. In section 3 we study the case where the total space E is an S^{2n-1}-sphere bundle over S^{2n}. In section 4 we consider maps between spheres S^m to S^{2n} for $m = 4n - 1$.

2. Generalities and preliminary results

Let E be the total space of an S^q-bundle over S^n and $f : E \to S^n$ be a map. In this section we first study the primary obstruction to make (f, f) coincidence free in the case where f is a fibre map. Then for f an arbitrary map and $q \neq n-1$, we reduce our problem to a problem either of maps between spheres, or of maps from a complex with two cells of positive dimension into a sphere.

The coincidence problem has a simple answer for n odd. Consider $f_1, f_2 : X \to S^n$ a pair of maps where X is an arbitrary space.

Lemma (2.1). *Let n be odd. Then the pair (f_1, f_2), for $f_1, f_2 : X \to S^n$, is homotopy disjoint if and only if $[f_1] = [f_2]$. In this case (f_1, f_1) is homotopy disjoint by a small deformation.*

Proof. From Proposition (1.2) we must have $[f_2] = [A \circ f_1]$, where A denotes the antipodal map on S^n. Since A is homotopic to the identity for n odd, the first part follows. For the last part it suffices to compose f_1 with a small perturbation of the identity of S^n which is fixed point free. □

From now on, we restrict ourselves to the case where the target is a sphere of even dimension. Let $p : E \to S^{2n}$ be a fibre map.

Proposition (2.2). *The trangression homomorphism $\Delta : H_{2n}(S^{2n}) \to H_{2n-1}(S^q)$ of the Serre Spectral sequence of the sphere bundle $S^q \to E \to S^{2n}$ is the trivial homomorphism if $q \neq 2n - 1$ and multiplication by an integer $l \in Z$ for a fixed choice of generators if $q = 2n - 1$.*

Proof. Since the fibre has homology only in dimension $0, q$, the result follows immediately from the Serre spectral sequence for dimensional reasons. □

Proposition (2.3). *The group $H^{2n}(E; Z)$ is isomorphic to Z if $q \neq 2n - 1$, $2n$. It is isomorphic to $Z + Z$ for $q = 2n$ and to Z/l if $q = 2n - 1$. Furthermore, $p^* : H^{2n}(S^{2n}; Z) \to H^{2n}(E; Z)$ is either the identity if $H^{2n}(E; Z)$ is isomorphic to Z, the natural projection $Z \to Z/l$ if $q = 2n - 1$, or an inclusion $Z \to Z + Z$ if $q = 2n$.*

Proof. This is a consequence of the Gysin sequence, where l above is the Euler number of the bundle. □

The integer l in the proposition above is the Euler number of the bundle. Now we will show that $p \not\pitchfork p$ in all cases where $l \neq \pm 1, \pm 2$.

Theorem (2.4). *We have $p \not\pitchfork p$ if either $q \neq 2n - 1$, or $q = 2n - 1$ and $l \neq \pm 1, \pm 2$.*

Proof. By Proposition (2.2) we have that $\chi(S^{2n})\cdot p^*(i_{2n}) \neq 0$. By Proposition (1.1), the result follows. ☐

Let $f : E \to S^{2n}$ be an arbitrary map and $q \neq 2n-1$. We will show that the problem of whether $f \mid\!\wr\, f$ either by a deformation or by a small deformation is equivalent to the same separation problem about maps between spheres or maps from a two cell complex into a sphere.

THEOREM (2.5). *Let $f : E \to S^{2n}$ where E is the total space of an S^q-sphere bundle over S^{2n} and $q \neq 2n-1$. Then we have:*

a) *For $q < 2n-1$, the map f factors through the cofibre $L \approx E/S^q$ of the inclusion $S^q \hookrightarrow E$ by a map $f' : L \to S^{2n}$ and L has the homotopy type of a two cell complex with cells in dimension $2n$ and $2n+q$. Then the primary obstruction in dimension $2n$ to have $f \mid\!\wr\, f$ can be identified with the primary obstruction in dimension $2n$ to have $f' \mid\!\wr\, f'$ either by a deformation or by a small deformation, through the isomorphism induced by the projection $E \to L$ in cohomology in dimension $2n$. This primary obstruction is zero if and only if ${f'}^*(\iota_{2n}) = 0$. When the primary obstruction to deform (f', f') coincidence free vanishes, f' factors through a map $f'' : S^{2n+q} \to S^{2n}$ where S^{2n+q} is the quotient of E by the $2n$-skeleton. Similarly, the higher obstruction to have $f'' \mid\!\wr\, f''$ (by small deformation) can be identified with the only obstruction to have $f'' \mid\!\wr\, f''$ (by small deformation), respectively.*

b) *For $q > 2n-1$, let $s : S^{2n} \to E$ be any section to the sphere bundle fibration $p : E \to S^{2n}$. Suppose $f \circ s$ is essential for a map $f : E \to S^{2n}$. Then $f \not\mid\!\wr\, f$.*

Proof. Part a). The existence of the factorization f' follows because $\pi_q(S^{2n})$ is zero for $q < 2n$. So let $f' : L \to S^{2n}$ be a factorization. Since the projection $E \to L$ induces an isomorphism in cohomology in dimensions $2n$ and $2n+q$, it follows that the primary obstruction to deforming (f, f) to be coincidence free vanishes if and only if the primary obstruction to deforming (f', f') to be coincidence free vanishes. For the case of a small deformation, a similar argument applies using Proposition (1.3). If the primary obstruction vanishes, then we must have ${f'}^*(\iota_{2n}) = 0$. So, by the Barratt-Puppe sequence, there is a second factorization, and the proof follows in the same fashion.

b) Let $g : E \to S^{2n}$ be any map homotopic to f. The coincidence number of the pair $(f \circ s, g \circ s)$ on S^{2n} equals the Lefschetz number $L(f \circ s, g \circ s) = 2\,\mathrm{degree}\,(f \circ s) \neq 0$. Thus f and g have a coincidence, so f cannot be separated from itself. ☐

From now on, we consider the cases where E is a S^{2n-1}-bundle over S^{2n}. The cases where the total space is not a sphere will be treated in the next section, and the remaining cases will be treated in the last section.

3. Self-coincidence of S^{2n-1}-bundles over S^{2n}

In this section we analyze the coincidence problem for maps defined on the above bundles, for $\ell \neq \pm 1$. The cases $\ell = \pm 1$ are considered in the next section. First we consider the self-coincidence problem for fibre maps $p : E \to S^{2n}$. Then we consider the coincidence problem for an arbitrary pair of maps $f_1, f_2 : E \to S^{2n}$.

Let $p : E \to S^{2n}$ be an $S^{2n-1}-$bundle.

For the fibration $p_1 : V_{2n+1,2} \to S^{2n}$ we have $p_1 \,|\!\wr\, p_1$ by a small deformation (see Proposition (1.3)). We show that for $\ell = \pm 2$ this is the only such case, up to fibre homotopy equivalence. More precisely:

Proposition (3.1). *We have that $p \,|\!\wr\, p$ by a small deformation if and only if the fibration $p : E \to S^{2n}$ is fibre homotopy equivalent to the fibration $p_1 : V_{2n+1,2} \to S^{2n}$.*

Proof. From Proposition (1.3), we have that $p \,|\!\wr\, p$ by a small deformation if and only if there is a map $\tilde{f} : E \to V_{2n+1,2}$ which makes the diagram below homotopy commutative

$$\begin{array}{ccc} & & V_{2n+1,2} \\ & & \downarrow \\ E & \xrightarrow{p} & S^{2n} \end{array}$$

Since p_1 is a fibration, we can replace $\tilde{f}$ by another map homotopic to $\tilde{f}$ which makes the diagram commutative. Denote also by $\tilde{f}$ this new map. Thus $\tilde{f}$ is a fibre-preserving map. This map, restricted to the fibre S^{2n-1}, is homotopic to either the identity or its negative, since the trangression of both fibrations is multiplication by 2. By the homotopy long exact sequence of the two fibrations and the induced homomorphism, it follows that $\tilde{f}$ is a fibre homotopy equivalence. $\square$

Remark (3.2). The above result shows, in particular, the following: Given a map $f : V_{2n+1,2} \to S^{2n}$ such that (f, f) is homotopy disjoint by small deformation, then either f is not in the homotopy class of a fibre map or it is in the class of the fibre map $p_1 : V_{2n+1,2} \to S^{2n}$.

Now we consider arbitrary maps, but we restrict our domains to the manifolds $V_{2n+1,2}$. To describe the set of homotopy classes of maps $[V_{2n+1,2}, S^{2n}]$, the following lemma is useful.

Lemma (3.3). *Let $L \approx V_{2n+1,2}/S^{2n-1}$ be the cofibre of the inclusion $i : S^{2n-1} \hookrightarrow E$. Then L has the homotopy type of $S^{2n} \vee S^{4n-1}$.*

Proof. The manifold $V_{2n+1,2}$ is the sphere bundle of the tangent bundle of the sphere S^{2n}. Since the tangent bundle is stably trivial, the Stiefel manifold $V_{2n+1,2}$ is stably reducible. That is, the gluing map $g : S^{4n-2} \to K$ is stably trivial, where K denotes the $2n$-skeleton of $V_{2n+1,2}$. Thus the composite of g with the projection $K \to K/S^{2n-1}$ is stably trivial. As the map g is already in the stable range, the result follows. $\square$

In the next Proposition we will use the Barratt-Puppe sequence. A general reference for the sequence is [15, Chapter II, Prop. (2.48)].

Proposition (3.4). *We have a short exact sequence of sets*

$$0 \to \pi_{4n-1}(S^{2n}) \to [V_{2n+1}, S^{2n}] \to Z/2 \to 0$$

.

Proof. Let K denote the $2n$-skeleton of $V_{2n+1,2}$. This complex has a cell structure of the form $S^{2n-1} \cup_2 e^{2n}$ where the characteristic map $2 : S^{2n-1} \to S^{2n-1}$ is a map of degree 2. The Barratt-Puppe sequence is

$$\cdots \to S^{4n-2} \to K \to V_{2n+1,2} \to S^{4n-1} \to \Sigma K \to \Sigma V_{2n+1,2} \to \dots$$

and produces

$$\dots \to [\Sigma V_{2n+1,2}, S^{2n}] \to [\Sigma K, S^{2n}] \to [S^{4n-1}, S^{2n}] \to [V_{2n+1,2}, S^{2n}] \to [K, S^{2n}] \to \dots$$

The cell structure of K, $S^{2n-1} \cup_2 e^{2n}$, yields $[K, S^{2n}] = H^{2n}(K, Z) \approx Z/2$, where the first equality follows from the Hopf Classification Theorem. In order to obtain the desired result, it suffices to show that the map $[V_{2n+1,2}, S^{2n}] \to [K, S^{2n}]$ is surjective and the map $[\Sigma K, S^{2n}] \to \pi_{4n-1}(S^{2n})$ is trivial. Since $V_{2n+1,2}$ is the total space of a fibration $p_1 : V_{2n+1,2} \to S^{2n}$, this map restricted to K sends the generator of $H^{2n}(S^{2n}, Z)$ to a generator of $H^{2n}(K, Z)$. So the map $[V_{2n+1,2}, S^{2n}] \to [K, S^{2n}]$ is surjective.

In order to prove that the map $[\Sigma K, S^{2n}] \to \pi_{4n-1}(S^{2n})$ is trivial , it suffices to show that $[\Sigma V_{2n+1,2}, S^{2n}] \to [\Sigma K, S^{2n}]$ is surjective. First let us consider the Barratt-Puppe sequence

$$\cdots \to S^{2n} \to S^{2n} \to \Sigma K \to S^{2n+1} \to S^{2n+1} \to \dots$$

Then we obtain the long sequences when we take homotopy classes into S^{2n}. By a routine argument we deduce that $[S^{2n+1}, S^{2n}] \to [\Sigma K, S^{2n}]$ is an isomorphism.

Call L the cofibre of the inclusion $S^{2n-1} \hookrightarrow V_{2n+1,2}$. So we have the map from $V_{2n+1,2}$ to L and a commutative diagram:

$$\begin{array}{cccccccccccc}
\dots \to & K \to & V_{2n+1,2} \to & S^{4n-1} \to & \Sigma K \to & \Sigma V_{2n+1,2} \to \dots \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
\dots \to & S^{2n} \to & L \to & S^{4n-1} \to & S^{2n+1} \to & \Sigma L \to \dots
\end{array}$$

and a commutative diagram of homomorphisms:

$$\begin{array}{ccc}
\dots \to [\Sigma L, S^{2n}] \to & [S^{2n+1}, S^{2n}] \to \dots \\
\downarrow & \downarrow \\
\dots \to [\Sigma E, S^{2n}] \to & [\Sigma K, S^{2n}] \to \dots
\end{array}$$

We have seen that the map $[S^{2n+1}, S^{2n}] \to [\Sigma K, S^{2n}]$ is an isomorphism. But $[\Sigma L, S^{2n}] \to [\Sigma S^{2n+1}, S^{2n}]$ is surjective, since the complex L is a wedge of two spheres by Lemma (3.3), and the result follows. □

Now we state a Proposition which is similar to the Proposition (2.12) in [2]. For let us consider the Barratt-Puppe sequence of a cofibration

$$\dots \to A \to B \to C(h) \xrightarrow{j} \Sigma A \to \dots$$

where $h : A \to B$. Consider the action of the group $[\Sigma A, \ \ ]$ on the set $[C(h), \ \ ]$. This action is provided from a pinching map $pin : C(h) \to C(h) \vee \Sigma A$ which arises in the study of the Barratt-Puppe sequence.

Proposition (3.5). *Let $\alpha_1, \alpha_2 \in [\Sigma A, X]$ and $[f_1], [f_2] \in [C(h), X]$. If, for $i = 1, 2$, obstructions θ_1^i and θ_2^i are defined for making (α_1, α_2) and (f_1, f_2) homotopy disjoint, respectively, then for $\alpha_1 \# [f_1], \alpha_2 \# [f_2]$ and $i = 1, 2$, we also have $\theta^i(\alpha_1 \# [f_1], \alpha_2 \# [f_2]) = j^i(\theta_1^i(\alpha_1, \alpha_2)) + \theta_2^i([f_1], [f_2])$.*

Proof. Consider the diagram:

$$\begin{array}{ccccccc} & & & & & & X \times X - \Delta \\ & & & & & & \downarrow \\ C(h) & \xrightarrow{\Delta} & C(h) \times C(h) & \xrightarrow{pin \times pin} & (\Sigma A \vee C(h)) \times (\Sigma A \vee C(h)) & \xrightarrow{\theta} & X \times X \end{array}$$

where $\theta = (\alpha_1 \vee f_1) \times (\alpha_2 \vee f_2)$, and also the diagram:

$$\begin{array}{ccccc} C(h) & \xrightarrow{\Delta \circ pin \times pin} & (\Sigma A \vee C(h)) \times (\Sigma A \vee C(h)) & \rightarrow & X \times X \\ \downarrow & & \uparrow & & \uparrow \\ \Sigma A \vee C(h) & \xrightarrow{\Delta \vee \Delta} & (\Sigma A \times \Sigma A) \vee (C(h) \times C(h)) & \rightarrow & (X \times X) \vee (X \times X) \end{array}$$

The obstruction to deforming the pair $(\alpha_1 \# [f_1], \alpha_2 \# [f_2])$ to be coincidence free is given by the obstruction to lift the composite of the horizontal maps in the first diagram. The commutativity of the second diagram gives us the desired formula. $\square$

With respect to the action of $\pi_{4n-1}(S^{2n})$ on $[V_{2n+1,2}, S^{2n}]$ we denote $I([f]) = isotropy[f] = \{\alpha \in \pi_{4n-1}(S^{2n}) | \alpha \# [f] = [f]\}$. Consider the Z_2-action on $[V_{2n+1,2}, S^{2n}]$ given by composing a map f with the antipodal map on S^{2n}. Call $i^{\#} : [V_{2n+1,2}, S^{2n}] \to [K, S^{2n}]$ the induced map from the inclusion $i : K \hookrightarrow E$. Recall that $[K, S^{2n}]$ is a group which is isomorphic to $Z/2$. Now we have a necessary condition to have (f, f) homotopy disjoint.

Proposition (3.6). *Let $f : V_{2n+1,2} \to S^{2n}$. The primary obstruction to have (f, f) homotopy disjoint is always zero. If $i^{\#}([f]) = 0$ then f factors through a map $f' : S^{4n-1} \to S^{2n}$, and the secondary obstruction to deforming (f, f) to be coincidence free coincides with the secondary obstruction for f'. If $i^{\#}([f]) \neq 0$ then $[f] = \alpha \# [p]$, and the secondary obstruction to make (f, f) coincidence free is identified with the one for the class $\alpha \in \pi_{4n-1}(S^{2n})$.*

Proof. This follows from Proposition (3.5) and the fact that (p, p) is homotopy disjoint. $\square$

The behaviour of the antipodal map with respect to the action of $\pi_{4n-1}(S^{2n})$ is given as follows:

Proposition (3.7). $[A \circ (\alpha \# h)] = [(A \circ \alpha)] \# [(A \circ h)]$, *where A is the antipodal map. For h either the constant map or the fibre map $p_1 : V_{2n+1,2} \to S^{2n}$ we have $[A \circ h] = [h]$.*

Proof. The proof is straightforward. $\square$

Now we come to the main result. Let $f_1, f_2 : V_{2n+1,2} \to S^{2n}$, where $f_1 = \alpha_1 \# [h_1]$ and $[f_2] = \alpha_2 \# [h_2]$ for $\alpha_i \in \pi_{4n-1}(S^{2n})$ and h_i either the constant map or the fibre map $p_1 : V_{2n+1,2} \to S^{2n}$. For n either 2 or 4 we write $[\alpha_i] = (r_i H + s_i v)$, where H is the Hopf map $S^{4n-1} \to S^{2n}$ for $n = 2, 4$ and v is a generator of the torsion part of $\pi_{4n-1}(S^{2n})$. Otherwise we write $[\alpha_i] = r_i[\iota_{2n}, \iota_{2n}] + v_i$, where $[\iota_{2n}, \iota_{2n}]$ denotes the Whitehead product and v_i a torsion element (see [16] and the next section for more details).

THEOREM (3.8). *Let $f_1, f_2 : V_{2n+1,2} \to S^{2n}$ be two maps as above.*

a) *If we have $f_1 \wr\!\wr f_2$, then $[h_1] = [h_2]$*

b) *For $n \neq 1, 2, 4$, if $f_1 \wr\!\wr f_2$, then $r_1 = r_2$ and $v_1 = -v_2$. In particular, for $f \wr\!\wr f$, v_1 is an element of torsion 2.*

c) *Let $n = 2$. In the case where $[h_1] = [h_2] = 0$ the two maps can be made coincidence free if and only if $r_1 = r_2$ and $r_1 + s_1 + s_2$ is divisible by 12. In the case where $[h_1] = [h_2] = [p]$ the two maps can be made coincidence free if and only if $(r_1 - r_2)H + (r_1 + s_1 + s_2)v \in I([p])$.*

d) *Let $n = 4$. In the case where $[h_1] = [h_2] = 0$ the two maps can be made coincidence free if and only if $r_1 = r_2$ and $r_1 + s_1 + s_2$ is divisible by 120. In the case where $[h_1] = [h_2] = [p]$ the two maps can be made coincidence free if and only if $(r_1 - r_2)H + (r_1 + s_1 + s_2)v \in I([p])$.*

Proof. Part a) follows from Proposition (3.4). Part b) follows from Proposition (4.1) and Proposition (3.5). Part c) and d) follows from Proposition (3.5) and from Proposition (4.3) and Corollary (4.4) in the next section. □

Remark (3.9). The obstruction to have the pair $(\alpha\#p, \alpha\#p)$ coincidence free is the pullback of the obstruction to have (α, α) coincidence free, by the induced homomorphism of the projection $V_{2n+1,2} \to S^{4n-1}$, since the obstruction to have (p, p) coincidence free is zero. This together with Proposition (4.3) implies that for $n = 2$ the elements of the form $rH + sv$ where $r + 2s$ are not divisible by 12, do not belong to $I([p])$. Similarly, using Proposition (4.3), for $n = 4$ the elements of the form $rH + sv$ where $r + 2s$ are not divisible by 120, do not belong to $I([p])$.

4. Coincidence of maps between spheres

In this section we discuss the coincidence problem for maps $f_1, f_2 : S^m \to S^n$ between spheres. For the particular case of self-coincidence, we consider the question from the point of view of deformations and small deformations.

We present some results when $m \leq 2n - 1$. More precisely we characterize in terms of the Whitehead product and the torsion elements of the homotopy group which pairs (f_1, f_2) are homotopy disjoint. In particular, (f_n, f_n) is homotopy disjoint for all n, where $f_n = [\iota_n, \iota_n]$. Then we consider the special cases $S^3 \to S^2$, $S^7 \to S^4$ and $S^{15} \to S^8$. At the end we give examples showing maps (f, f) homotopy disjoint but not by a small deformation. These examples show that the bounds given in [2, Lemma (2.14)] or [12, Theorem (2.2)] are sharp and confirm the comment in [2] immediately after the definition of homotopy disjoint deformation.

From the discussion in section 2, we can consider the target an even sphere S^{2n}.

PROPOSITION (4.1). *The pair (f_1, f_2) for $f_1, f_2 : S^m \to S^{2n}$ is homotopy disjoint if and only if:*

a) *For $m < 4n - 1$ the class $[f_1] = -[f_2]$ in $\pi_m(S^{2n})$; in the case $[f_1] = [f_2]$, $[f_1]$ has order 2.*

b) *For $m = 4n - 1$ and $n \neq 1, 2, 4$, the classes $[f_1], [f_2]$ are of the form $[f_1] = k_1[\iota_{2n}, \iota_{2n}] + \alpha_1$, $[f_2] = k_2[\iota_{2n}, \iota_{2n}] + \alpha_2$, where $[\iota_{2n}, \iota_{2n}]$ is the Whitehead product, the α_i's are torsion elements, and $k_1 = k_2$ and $\alpha_1 + \alpha_2 = 0$. In* *particular, if* $[f] = k[\iota_{2n}, \iota_{2n}] + r\alpha$ *then* (f, f) *is homotopy disjoint if and only if* α *is an element of order* ≤ 2 *and* k *is arbitrary.*

Proof. By Proposition (1.2) we must verify when f_2 is homotopic to $A \circ f_1$. For part a), since $m < 4n - 1$, we have that $A \circ f_1 = -f_1$ and the result follows.

For part b), from the formula in [17] or [16] for the composition of a map with the sum of two others, it follows that:

$$0 = [(I + A) \circ f_1] = [f_1] + [A \circ f_1] + [\iota_{2n}, -\iota_{2n}]H([f_1]) =$$

$$[f_1] + [A \circ f_1] - 2k_1[\iota_{2n}, \iota_{2n}],$$

since the Hopf invariant of the Whitehead product is 2 and the Hopf invariant of a torsion element is zero. Therefore $[A \circ f_1] = k_1[\iota_{2n}, \iota_{2n}] - \alpha_1$. From the equality $A \circ f_1 = f_2$ the result follows. □

Remark (4.2). If $f : S^m \to S^{2n}$ has order 2 and $m < 4n - 2$, either [2, Lemma (2.14)] or [12, Theorem (2.2)] implies that (f, f) is homotopy disjoint by a small deformation. It is natural to ask what happens in the cases $m = 4n - 2$ where $[f]$ has order 2. We will show at the end that these results are sharp.

Now we consider the cases where the domain is one of the spheres S^3, S^7,or S^{15}, and the target is S^2, S^4, or S^8, respectively. These cases correspond to $l = \pm 1$. The case of pairs of maps (f_1, f_2) where $f_i \in [S^3, S^2]$ is quite simple. The pair (f_1, f_2) is homotopy disjoint if and only if $[f_1] = [f_2]$. Further, the pair (f, f) is homotopy disjoint by a small deformation for an arbitrary map f. We leave the details to the reader. In order to make the exposition more clear we deal first with the case where the total space is the sphere S^7 and then the case where the total space is S^{15}. From [16] we have that $\pi_7(S^4) = Z + Z/12$, where a generator of the summand Z can be taken as the class of the Hopf map and a generator of the torsion part as v, where v is the suspension of v', for v' a generator of $\pi_6(S^3)$.

Let $f_1 = r_1 H + s_1 v$ and $f_2 = r_2 H + s_2 v$.

PROPOSITION (4.3). *For* $f = rH + sv \in \pi_7(S^4)$ *we have that* $[(A \circ f)] = rH - (r + s)v$. *Furthermore, the pair of maps* $(f_1, f_2) : S^7 \to S^4$ *has the property that* $f_1 \not\wr f_2$ *if and only if* $r_1 = r_2 = r$ *and* $s_1 + s_2 + r$ *is divisible by 12.*

Proof. The proof is the same as the proof of Proposition (4.1) part b), where we use the fact that the Hopf invariant of the Hopf map is 1. Namely,

$$0 = [(I + A) \circ f_1] = [f_1] + [A \circ f_1] + [I, A] \circ H(f_1) = [f_1] + [A \circ f_1] - (2H - v)r_1.$$

Since $I + A$ is trivial, we get

$$[A \circ f_1] = -[f_1] + 2r_1 H - r_1 v = -r_1 H - s_1 v + 2r_1 H - r_1 v = r_1 H + (-r_1 - s_1)v.$$

From the equality $A \circ f_1 = f_2 = r_2 H + s_2 v$, we have $r_1 = r_2$ and $(s_2 + s_1 + r_1)v = 0$. Therefore $r_1 = r_2 = r$ and $r + s_1 + s_2$ is divisible by 12, and the result follows. □

COROLLARY (4.4). *A map* $f : S^7 \to S^4$ *has the property that* $f \not\wr f$ *by small deformation if and only if* $r + 2s$ *is divisible by 12, where* $f = rH + sv$.

Proof. By the previous result we have that (f, f) is homotopy disjoint if and only if $2s + r$ is divisible by 12. By [2, Corollary (2.15)] the result follows. □

Corollary (4.5). *The kernel of the transgression* Δ_7 *of the fibration* $p : V_{5,2} \to S^4$ *is generated by* $2h - v$ *and* $6v$.

Proof. By Corollary (4.4) the kernel of the transgression are the elements of $\pi_7(S^4)$ of the form $rH + sv$ where $r + 2s$ is divisible by 12. A simple divisibility argument shows that these elements are generated by $2H - v$ and $6v$, and the result follows. □

Remark (4.6). See [13] for results related with above Corollary.

Now we consider the case of S^{15}. From [16] we have that $\pi_{15}(S^8) = Z + Z/120$, where a generator of the summand Z can be take as the class of the Hopf map and a generator of the torsion part as v, where v is the suspension of v', for v' a generator of $\pi_{14}(S^7)$. Let f be of the form $rH + sv$.

Proposition (4.7). *Let* $f = rH + sv \in \pi_{15}(S^8)$. *Then we have that* $[(A \circ f)] = rH - (r + s)v$. *Further, the pair of maps* $(f_1, f_2) : S^{15} \to S^8$ *has the property that* $f_1 \mathrel{|\!\wr} f_2$ *if and only if* $r_1 = r_2 = r$ *and* $s_1 + s_2 + r$ *is divisible by 120.*

Proof. The proof is the same as the proof of Proposition (4.1) part *b*) where we use the fact that the Hopf invariant of the Hopf map is 1. Namely,

$$0 = [(I + A) \circ f_1] = [f_1] + [A \circ f_1] + [I, A] \circ H(f_1) = [f_1] + [A \circ f_1] - (2H - v)r_1.$$

Since $I + A$ is trivial, we get

$$[A \circ f_1] = -[f_1] + 2r_1H - r_1v = -r_1H - s_1v + 2r_1H - r_1v = r_1H + (-r_1 - s_1)v.$$

Form the equality $A \circ f_1 = f_2 = r_2H + s_2v$ it follows that $r_1 = r_2$ and $(s_2 + s_1 + r_1)v = 0$. Therefore $r_1 = r_2 = r$ and $r + s_1 + s_2$ is divisible by 120, and the result follows. □

Corollary (4.8). *A map* $f : S^{15} \to S^8$ *has the property that* $f \mathrel{|\!\wr} f$ *by small deformation if and only if* $r + 2s$ *is divisible by 120, where* $f = rH + sv$.

Proof. By the previous result we have that (f, f) is homotopy disjoint if and only if $2s + r$ is divisible by 120. By [2, Corollary (2.15)] the result follows. □

Corollary (4.9). *The kernel of the transgression* Δ_{15} *of the fibration* $p : V_{9,2} \to S^8$ *is generated by* $2h - v$ *and* $60v$.

Proof. By Corollary (4.8) the kernel of the transgression are the elements of $\pi_{15}(S^8)$ of the form $rH + sv$ where $r + 2s$ is divisible by 120. A simple divisibility argument shows that these elements are generated by $2H - v$ and $60v$, and the result follows. □

Remark (4.10). See [13] for results related with above Corollary.

Now we consider the difference between deformation and small deformation. We construct maps $f_{2n} : S^{4n-1} \to S^{2n}$ for all odd $n > 1$ for which (f_{2n}, f_{2n}) is homotopy disjoint, but not homotopy disjoint by a small deformation.

Theorem (4.11). *Let* ξ_{2n} *denote the bundle over* S^{4n-1} *induced from the tangent bundle* $\tau(S^{2n})$ *by the Whitehead square* $[\iota_{2n}, \iota_{2n}] \in \pi_{4n-1}(S^{2n})$ *for* $n \geq 2$. *Then* ξ_{2n} *admits a non-zero section for all even* n, *but does not admit a non-zero section for all odd* $n > 1$.

Proof. The Whitehead product $[\iota_{2n-1}, \eta_{2n-1}]$ has order 2 for all odd $n > 1$ by [10, Lemma (3.5)], and is trivial for all even n. Moreover, James proved that $\partial([\iota_{2n}, \iota_{2n}]) = [\iota_{2n-1}, \eta_{2n-1}]$ in the homotopy exact sequence for $V_{2n+1,2} \to S^{2n}$ by (6.1) and (6.2) of [10]. Consequently, $\partial\iota_{4n-1} = [\iota_{2n-1}, \eta_{2n-1}]$ in the homotopy exact sequence for $S(\xi_{2n}) \to S^{4n-1}$ by naturality. □

THEOREM (4.12). *Let $f_{2n} : S^{4n-1} \to S^{2n}$ be any representative for $[\iota_{2n}, \iota_{2n}]$ with $n > 1$. Then (f_{2n}, f_{2n}) is homotopy disjoint for all $n > 1$. Moreover, (f_{2n}, f_{2n}) is homotopy disjoint by a small deformation if and only if n is even.*

Proof. By Proposition (4.1) above, (f_{2n}, f_{2n}) is homotopy disjoint for all n. Moreover, (f_{2n}, f_{2n}) is homotopy disjoint by a small deformation if and only if n is even by Theorem (4.11) and Proposition (1.4). □

Our last result affirms that homotopy disjointness does not imply homotopy disjoint by a small deformation even for $m = 4n - 2$.

PROPOSITION (4.13). *Let $f : S^{30} \to S^{16}$ represent σ_{16}^2. Then (f, f) is homotopy disjoint, but not homotopy disjoint by a small deformation.*

Proof. σ_{16}^2 has order 2 in $\pi_{30}(S^{16})$ by [16]. Moreover, $\partial(\sigma_{16}^2) = 2\sigma_{15}^2$, where σ_{15}^2 has order 4 in $\pi_{29}(S^{15})$ by [16]. Here ∂ denotes the homotopy boundary operator in the fibration $S^{15} \to V_{17,2} \to S^{16}$, where $\partial\iota_{16} = 2\iota_{15}$. By Proposition (4.1), (f, f) is homotopy disjoint, but not homotopy disjoint by a small deformation by Proposition (1.4). □

Acknowledgments

This work started during the visit of the first author to Loyola University, New Orleans, in April, 2000. He would like to thank the Topology group at Loyola University for the invitation and their hospitality. The visit was supported by the FAPESP-Fundação de Amparo a Pesquisa do Estado de São Paulo.

Received July 02, 2003

Final version received April 07, 2004

D. L. GONÇALVES
DEPARTAMENTO DE MATEMÁTICA - IME-USP
CAIXA POSTAL 66281 - AG. CIDADE DE SÃO PAULO
CEP: 05311-970 - SÃO PAULO - SP
BRASIL
dlgoncal@ime.usp.br

D. RANDALL
DEPARTMENT OF MATHEMATICS AND
COMPUTER SCIENCE - LOYOLA UNIVERSITY
6363 ST CHARLES AVENUE
NEW ORLEANS LA 70118
USA
randall@loyno.edu

References

[1] L. D. Borsari and D. L. Gonçalves, *A Van Kampen Type theorem for coincidence*, Top. Appl. **101**, (2000), 149–160.

[2] A. Dold and D. L. Gonçalves, *Self-coincidence of fibre maps*, Osaka J. Math., to appear.

[3] D. Dimovski and R. Geoghegan, *One-parameter fixed point theory*, Forum Math. **2** (1990), 125–154.

[4] D. L. Gonçalves, *Coincidence Reidemeister classes on nilmanifolds and nilpotent fibrations*, Top. Appl. **83** (1998), 169–186.

[5] D. L. Gonçalves, *The coincidence Reidemeister classes of maps on nilmanifolds*, Topol. Methods Nonlinear Anal. **12** (1998), 375–386.

[6] D. L. Gonçalves, *Coincidence theory for maps from a complex to a manifold.* Top. Appl. **92** (1999), 63–77.

[7] D. Gonçalves and J. Jezierski, *The Lefschetz coincidence formula on non-orientable manifolds*, Fund. Math. **153** (1997), 1–23.

[8] D. Gonçalves and P. Wong, *Nilmanifolds are Jiang-type spaces for coincidences*, Forum Math. **13** (2001), 133–141.

[9] A. Hatcher and F. Quinn, *Bordism invariants of intersections of submanifolds*, Trans. Amer. Math. Soc. **200** (1974), 327–344.

[10] I. M. James, *Products on Spheres*, Mathematika **6** (1959), 1–13.

[11] J. Jezierski, *One codimensional Wecken type theorems*, Forum Math. **5** (1993), 421–439.

[12] U. Koschorke, *Self-coincidence in higher codimensions*, J. Reine und Angew. Math., to appear.

[13] G. F. Paechter, *The groups $\pi_r(V_{n,m})$ I*, Quart. J. Math. **7** (1956), 249–268.

[14] H. Schirmer, *Mindestzahlen von Koinzidenzpunkte*, J. Reine und Angew. Math. **194** (1955), 21–39.

[15] R. M. Switzer, Algebraic Topology-Homotopy and Homology, Springer-Verlag, Berlin 1975.

[16] H. Toda, Composition methods in the homotopy groups of the spheres, Ann. of Math. Stud. **49** Princeton Univ. Press, Princeton, NJ, 1962.

[17] G. Whitehead, Elements of Homotopy Theory. Springer Verlag, New York, 1978.

[18] P. Wong, *Reidemeister number, Hirsch rank, coincidences on polycyclic groups and solvmanifolds*, J. Reine Angew. Math. **524** (2000), 185–204.

[19] A. Wyler, *Sur certaines singularités d'aplications de variétés topologiques*, Comment. Math. Helv. **42** (1067), 28–48.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

ARTIN GROUPS, 3-MANIFOLDS AND COHERENCE

C. MCA. GORDON

Dedicated to Fico on the occasion of his 60th birthday.

Abstract. Following work of Droms [D] and Hermiller and Meier [HM], we show that the Artin group $A\Gamma$ associated with a labeled graph Γ is a 3-manifold group if and only if each component of Γ is either a tree, or a triangle with each edge labeled 2.

1. Introduction

By a *labeled graph* we shall mean a finite (non-empty) graph Γ, without loops or multiple edges, each of whose edges is labeled by an integer greater than or equal to 2. Let the vertices of Γ be $s_1, s_2, \ldots, s_n$, and let the label on an edge with endpoints s_i and s_j be $m_{ij} \geq 2$. Define $\langle ab \rangle^m$ to be the word $abab\ldots$ of length m. Then the *Artin group* $A\Gamma$ associated with the labeled graph Γ is the group with generators $s_1, s_2, \ldots, s_n$, and relations $\langle s_i s_j \rangle^{m_{ij}} = \langle s_j s_i \rangle^{m_{ij}}$, one for each edge of Γ. In particular, if $m_{ij} = 2$ then the generators s_i and s_j commute. Note also that if Γ is the disjoint union of graphs Γ_1 and Γ_2 then $A\Gamma \cong A\Gamma_1 * A\Gamma_2$.

A *3-manifold group* is a group that is isomorphic to $\pi_1(M)$ for some (connected) 3-manifold M. Note that we do not assume that M is orientable, or compact, or without boundary. Taking a connected sum shows that if G_1 and G_2 are 3-manifold groups then so is $G_1 * G_2$. If Γ is a tree, then $A\Gamma$ is the fundamental group of the complement of a link L in S^3, where L is a connected sum of $(2, m)$ torus links; see [Bru], [HM]. Thus $A\Gamma$ is a 3-manifold group. If Γ is a triangle with each edge labeled 2, then $A\Gamma \cong \mathbb{Z}^3 \cong \pi_1(T^3)$ is also a 3-manifold group. In this note we confirm the suspicion of Hermiller and Meier [HM, p.143] that these are the only connected graphs whose Artin groups are 3-manifold groups.

Theorem (1.1). *For an Artin group $A\Gamma$ the following are equivalent.*
(1) $A\Gamma$ *is a 3-manifold group.*
(2) $A\Gamma$ *is virtually a 3-manifold group.*
(3) *Each component of* Γ *is either a tree, or a triangle with each edge labeled 2.*

The equivalence of (1) and (3) was proved by Droms [D] in the case of *right-angled* Artin groups, or *graph groups*, that is, when all labels are 2, and by Hermiller and Meier [HM] in the case when all labels are even.

Theorem (1.1) is proved in Section 2. The main tool used is the fact that 3-manifold groups are coherent [Sc].

2000 *Mathematics Subject Classification*: 20F36, 57M99.
Keywords and phrases: Artin groups; 3-manifold groups; coherence.
Partially supported by TARP grant 003658-0519-2001.

In Section 3 we make some additional remarks about coherence. In particular we show that $\mathrm{Aut}(F_2)$ and the braid group B_4 are incoherent, although neither has a subgroup of the form $F_2 \times F_2$. The latter fact for B_4 was originally proved by Akimenkov [A], using different methods.

2. Artin groups and 3-manifolds

Recall that a group is *coherent* if every finitely generated subgroup is finitely presented.

The following is proved in [HM, Proposition 5.7(ii)].

LEMMA (2.1) (Hermiller and Meier). *Let Γ be a cycle of length at least 4. Then $A\Gamma$ is incoherent.*

If Γ is a labeled graph, we shall say that Γ is of *infinite* or *finite type* according as the Coxeter group corresponding to the Artin group $A\Gamma$ is infinite or finite. We will use (p, q, r) to denote a triangle with edge labels p, q and r. The triangles of finite type are then $(2, 2, m)$, $(2, 3, 3)$, $(2, 3, 4)$ and $(2, 3, 5)$.

If Γ is a triangle, the simplicial complex K_0 defined in [CD2] is either a triangle or a 2-simplex, according as Γ is of infinite or finite type. The Main Conjecture of [CD1] and [CD2] therefore holds for $A\Gamma$, by [CD1]. The following lemma is then a consequence of [CD2, Corollary 1.4.2 and Corollary 2.2.5].

LEMMA (2.2) (Charney and Davis). *Let Γ be a triangle.*
(i) *If Γ is of infinite type then $A\Gamma$ has geometric dimension 2 and $\chi(A\Gamma) = 1$.*
(ii) *If Γ is of finite type then $A\Gamma$ has geometric dimension 3 and $\chi(A\Gamma) = 0$.*

For the three triangles $(2, 3, 6)$, $(2, 4, 4)$ and $(3, 3, 3)$ of Euclidean type, (i) also follows from the descriptions of $A\Gamma$ given in [Sq1].

LEMMA (2.3). *Let Γ be a triangle of infinite type. Then $A\Gamma$ is incoherent.*

Proof. Let $\varphi : A\Gamma \to \mathbb{Z}$ be the epimorphism that maps each generator s_i to 1. By [Me, Proposition 5.1 and Corollary 5.3] $K = \ker \varphi$ is finitely generated. Now $A\Gamma$ has geometric dimension 2 (Lemma (2.2)), and hence K has geometric dimension ≤ 2. Suppose K were finitely presented. Then K would be of type FP [Bro, p.199], and so $\chi(K)$ would be defined. We would then have $\chi(A\Gamma) = \chi(K)\chi(\mathbb{Z}) = 0$ [Bro, p.250], [St2], (compare [G]), contradicting Lemma (2.2). $\square$

In [W], Wall asked whether a group of the form $F *_C F'$, where F and F' are free and C has finite index in F and F', is coherent. This was answered negatively by Gersten [G], who showed that the double of a free group of rank ≥ 2 along a subgroup of finite index ≥ 3 is always incoherent. We remark that Lemma (2.3) also provides examples, which are not doubles, since Squier has shown [Sq1] that $A(2, 3, 6)$ and $A(3, 3, 3)$ can each be expressed as a free product with amalgamation $F *_C F'$, where $\operatorname{rank} F = 4$, $\operatorname{rank} F' = 3$, and C has index 2 in F and index 3 in F'.

LEMMA (2.4). *$A(2, 3, 3)$ and $A(2, 3, 4)$ are incoherent.*

Proof. Since $A(2, 3, 4)$ embeds in $A(2, 3, 3)$ (as a subgroup of finite index) [La], it is enough to show that $A(2, 3, 4)$ is incoherent. One way to do this is to use the fact that $A(3, 3, 3)$ embeds in $A(2, 3, 4)$ [KP], together with Lemma (2.3).

Another argument is that the commutator subgroup A' of $A(2,3,4)$ is finitely generated but, since $H_2(A') \cong \mathbb{Z}^\infty$, not finitely presented [Sq2]. □

Note that $A(2,2,m) \cong A(m) \times \mathbb{Z}$, where $A(m)$ is the Artin group of a single edge with label m. Since $A(m)$ is a 3-manifold group, it is coherent by [Sc] (it is also easy to show this directly), and hence $A(2,2,m)$ is coherent.

LEMMA (2.5). *$A(2,2,m)$, $m > 2$, and $A(2,3,5)$ are not virtually 3-manifold groups.*

Proof. Let G be a finitely generated group, with an epimorphism $\varphi : G \to \mathbb{Z}$ such that $\ker\varphi$ is finitely generated, and let H be a subgroup of G of finite index. Then φ induces an epimorphism $\psi : H \to \mathbb{Z}$, where $\ker\psi$ has finite index in $\ker\varphi$. Now suppose that H is a 3-manifold group. Since H is finitely generated, it is the fundamental group of a compact 3-manifold [Sc]. Therefore, since $\ker\psi$ is finitely generated, by [St1] $\ker\psi$ is a *2-manifold group*, i.e. it is either free or the fundamental group of a closed surface. Hence $\ker\varphi$ is virtually a 2-manifold group.

Suppose $A(2,2,m) \cong A(m) \times \mathbb{Z}$ is virtually a 3-manifold group. Then by the above discussion $A(m)$ has a subgroup B of finite index such that B is either free or the fundamental group of a closed orientable surface. Since $A(m)$ is the fundamental group of a compact, orientable, irreducible 3-manifold M whose boundary consists of tori, $\chi(A(m)) = \chi(M) = \frac{1}{2}\chi(\partial M) = 0$. Hence $\chi(B) = 0$, implying that B is isomorphic to either $\mathbb{Z}$ or $\mathbb{Z}\times\mathbb{Z}$. But if $m > 2$, this contradicts the fact that $A(m)$ contains a non-abelian free group.

Now let $A = A(2,3,5)$, and let $\varphi : A \to \mathbb{Z}$ be abelianization. Then $A' = \ker\varphi$ is finitely generated by [Me]. Suppose A is virtually a 3-manifold group. Then there is a 2-manifold subgroup B of A' of finite index. By a standard transfer argument, $H_2(B;\mathbb{Q}) \to H_2(A';\mathbb{Q})$ is surjective. But $\dim H_2(A';\mathbb{Q}) = 7$ [Sq2], whereas $\dim H_2(B;\mathbb{Q}) \leq 1$. □

Proof of Theorem (1.1). Clearly (1) implies (2) and (3) implies (1); we must show that (2) implies (3).

A subgraph Γ_0 of Γ is *full* if every edge of Γ whose vertices are in Γ_0 is an edge of Γ_0. We recall the basic fact [Le] that if Γ_0 is a full subgraph of a labeled graph Γ then the homomorphism $A\Gamma_0 \to A\Gamma$ induced by the inclusion map $\Gamma_0 \subset \Gamma$ is injective.

Let Γ be a connected labeled graph, and suppose that $A\Gamma$ is virtually a 3-manifold group. By [Sc], 3-manifold groups, and hence virtual 3-manifold groups, are coherent. Also, a subgroup of a virtual 3-manifold group is clearly a virtual 3-manifold group. It follows from Lemma (2.1) that Γ is *chordal*, i.e. has no full subgraph that is a cycle of length ≥ 4. By Lemmas (2.4) and (2.5), any triangle in Γ has all labels equal to 2. If Γ is not a tree or a $(2,2,2)$ triangle, then Γ has a full subgraph Γ_0 of one of the forms shown in Figure 1 (where all unlabeled edges are understood to have label 2); see [D].

In cases (i) and (ii), $A\Gamma_0 \cong A \times \mathbb{Z}$, where in case (i) $A \cong \mathbb{Z}^3$, and in case (ii) A is the Artin group of a tree with two edges, each labeled 2. Since $A\Gamma_0$ is virtually a 3-manifold group by assumption, A has a subgroup B of finite index that is either free or the fundamental group of a closed orientable surface. Since

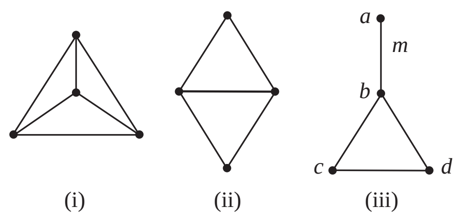


Figure 1

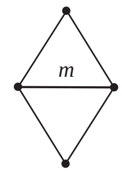


Figure 2

$\chi(A) = 0$, we have $\chi(B) = 0$, and hence $B \cong \mathbb{Z}$ or $\mathbb{Z} \times \mathbb{Z}$, which is clearly impossible.

In case (iii), m odd, let $\varphi : A\Gamma_0 \to \mathbb{Z}$ be the epimorphism defined by $\varphi(a) = \varphi(b) = 1$, $\varphi(c) = \varphi(d) = 0$. By [Me], $\ker\varphi$ is finitely generated. Hence $\ker\varphi$ has a subgroup B of finite index that is either free or the fundamental group of a closed orientable surface. Since $\ker\varphi$ contains $\langle c, d\rangle \cong \mathbb{Z} \times \mathbb{Z}$, we must have $B \cong \mathbb{Z}\times\mathbb{Z}$. But this contradicts the fact that $\ker\varphi$ also contains the commutator subgroup of $A(m)$, which is a non-abelian free group.

In case (iii), m even, define $\varphi : A\Gamma_0 \to \mathbb{Z}$ by $\varphi(b) = 1$, $\varphi(a) = \varphi(c) = \varphi(d) = 0$, as in [HM]. Then $\ker\varphi$ is finitely generated and contains both $\mathbb{Z}\times\mathbb{Z}$ and a non-abelian free group, giving a contradiction as before. $\square$

3. Coherence

It is natural to ask which Artin groups $A\Gamma$ are coherent. For graph groups, i.e. when all edge labels are 2, this has been answered by Droms [D]: $A\Gamma$ is coherent if and only if Γ is chordal. For the general case, it is necessary to be able to answer the following question.

Question (3.1). *Is $A(2,3,5)$ coherent?*

If at most one of p, q, r is even, the homomorphism $\varphi : A(p,q,r) \to \mathbb{Z}$ in the proof of Lemma (2.3) is abelianization, so that proof shows that if (p,q,r) is of infinite type then the commutator subgroup of $A(p,q,r)$ is finitely generated but not finitely presented. However, as pointed out in [Sq2], the commutator subgroup of $A(2,3,5)$ is finitely presented (the same argument applies to $A(2,3,3)$).

If $A(2,3,5)$ is incoherent, one can show that an Artin group $A\Gamma$ is coherent if and only if Γ is chordal, every complete subgraph of Γ with 3 or 4 vertices has at most one edge label > 2, and Γ has no full subgraph of the form shown in Figure 2, where $m > 2$ and unlabeled edges are understood to have label 2. If $A(2,3,5)$ is coherent, the characterization would be more complicated.

Let F_n denote the free group of rank n. A popular way of showing that a group is incoherent is to show that it has a subgroup isomorphic to $F_2 \times F_2$, which is well-known to be incoherent; see e.g. [G]. For example, $\mathrm{Aut}(F_3)$ has such a subgroup [FP], and hence $\mathrm{Aut}(F_n)$ is incoherent for $n \geq 3$. For $n = 2$ we have

Theorem (3.2). (1) $\mathrm{Aut}(F_2)$ *is incoherent.*
(2) $F_2 \times F_2$ *does not embed in* $\mathrm{Aut}(F_2)$.

Let B_n denote the n-strand braid group. Note that B_n is coherent for $n \leq 3$. Since $A(2,3,3) \cong B_4$, we see by Lemma (2.4) that B_n is incoherent for $n \geq 4$. Also, since the center of $F_2 \times F_2$ is trivial, by Part (2) of Theorem (3.2) and the proof of Part (1) below we recover the result of Akimenkov [A] that $F_2 \times F_2$ does not embed in B_4. It follows (see the proof of Lemma (2.4)) that the incoherent groups $A(2,3,4)$ and $A(3,3,3)$ also do not contain an $F_2 \times F_2$. (It is shown in [Ma] that $F_2 \times F_2$ does embed in B_n for $n \geq 5$.)

Proof of Theorem (3.2). (1)Let $Z(B_4)$ denote the center of B_4; then $B_4/Z(B_4)$ is isomorphic to an index 2 subgroup of $\mathrm{Aut}(F_2)$ [DFG]. Now $A(3,3,3)$ embeds in B_4 [KP], and since it is a free product with amalgamation of the form $F_4 *_{F_7} F_3$ [Sq1], it has trivial center, and hence embeds in $\mathrm{Aut}(F_2)$. Since $A(3,3,3)$ is incoherent by Lemma (2.3), $\mathrm{Aut}(F_2)$ is also incoherent.

(2) There is a short exact sequence

$$1 \to F_2 \xrightarrow{i} \mathrm{Aut}(F_2) \xrightarrow{\pi} GL(2,\mathbb{Z}) \to 1 \ ,$$

where $i(g)$ is conjugation by g. Mapping $GL(2,\mathbb{Z})$ onto $PSL(2,\mathbb{Z}) \cong \mathbb{Z}_2 * \mathbb{Z}_3$ gives the related sequence

$$1 \to \ker\rho \to \mathrm{Aut}(F_2) \xrightarrow{\rho} \mathbb{Z}_2 * \mathbb{Z}_3 \to 1 \ ,$$

where $\ker\pi$ has index 4 in $\ker\rho$. In particular, $\ker\rho$ is virtually free.

Suppose $H < \mathrm{Aut}(F_2)$, where $H = \langle\alpha,\beta\rangle \times \langle\gamma,\delta\rangle = H_1 \times H_2 \cong F_2 \times F_2$. We claim that either $\rho(H_1) = 1$ or $\rho(H_2) = 1$. For, if not, then writing $\bar{\alpha} = \rho(\alpha)$ etc., we may assume that $\bar{\alpha} \neq 1 \neq \bar{\gamma}$. By the Kurosh Subgroup Theorem, any abelian subgroup of $\mathbb{Z}_2 * \mathbb{Z}_3$ is cyclic. Therefore $\langle\bar{\alpha},\bar{\gamma}\rangle = \langle x\rangle$, $\langle\bar{\alpha},\bar{\delta}\rangle = \langle y\rangle$, say. Then we have $1 \neq \bar{\alpha} = x^p = y^q$ for some $p,q \in \mathbb{Z}$. It follows that $\langle x,y\rangle$ has a non-trivial center, and hence, again by the Kurosh Subgroup Theorem, $\langle x,y\rangle = \langle z\rangle$, say. Therefore $\langle\bar{\gamma},\bar{\delta}\rangle < \langle z\rangle$ is cyclic, implying that $(\ker\rho) \cap H_2 \neq 1$. Similarly $(\ker\rho) \cap H_1 \neq 1$. This gives $\mathbb{Z} \times \mathbb{Z} < \ker\rho$, contradicting the fact that $\ker\rho$ is virtually free.

We may assume, then, without loss of generality, that $\rho(H_2) = 1$. Then $\rho|H_1$ is injective, otherwise we would have $\mathbb{Z} \times H_2 < \ker\rho$, again contradicting the fact that $\ker\rho$ is virtually free. It follows that $\pi|H_1$ is injective. Also, $H_2 < \ker\rho$, and therefore $(\ker\pi) \cap H_2$ has finite index in H_2. Hence $(\ker\pi) \cap H_2 = i(G)$, where $G < F_2$ has rank ≥ 2.

Note that since F_2 has trivial center, $\varphi \in \mathrm{Aut}(F_2)$ commutes with $i(g)$, $g \in F_2$, if and only if $\varphi(g) = g$. Hence if $\varphi \in H_1$, then $G < \mathrm{Fix}(\varphi)$. Since rank $G \geq 2$, it follows from the Scott Conjecture [BH] that $\mathrm{Fix}(\varphi)$ has rank 2. Furthermore, by [CT] there is a basis a,b of F_2 such that $\varphi(a) = a$, $\varphi(b) = ba^n$. Hence $\pi(\varphi) \in GL(2,\mathbb{Z})$ has trace 2. But since $\pi(H_1)$ is a free group of rank 2, this is a contradiction. □

Received July 14, 2003

Final version received January 28, 2004

Department of Mathematics
The University of Texas at Austin
1 University Station C1200
Austin, TX 78712-0257
USA
gordon@math.utexas.edu

References

[A] A. M. Akimenkov, *Subgroups of the braid group B_4*, Math. Notes Acad. Sci. USSR **50** (6), (1991), 1211–1218.

[BH] M. Bestvina and M. Handel, *Train tracks and automorphisms of free groups*, Ann. of Math. (2) **135** (1992), 1–53.

[Bro] K. S. Brown, Cohomology of Groups, Graduate Texts in Mathematics **87**, Springer-Verlag, New York, 1982.

[Bru] A. M. Brunner, *Geometric quotients of link groups*, Topology Appl. **48** (1992), 245–262.

[CD1] R. Charney and M. Davis, *The $K(\pi,1)$-problem for hyperplane complements associated to infinite reflection groups*, J. Amer. Math. Soc. **8** (1995), 597–627.

[CD2] R. Charney and M. Davis, *Finite $K(\pi,1)$s for Artin groups*, Ann. of Math. Studies **138**, Prospects in Topology, ed. F. Quinn, Princeton University Press, Princeton, New Jersey, 1995, 110–124.

[CT] D. J. Collins and E. C. Turner, *All automorphisms of free groups with maximal rank fixed subgroups*, Math. Proc. Camb. Phil. Soc. **119** (1996), 615–630.

[D] C. Droms, *Graph groups, coherence, and three-manifolds*, J. Algebra **106** (1987), 484–489.

[DFG] J. L. Dyer, E. Formanek and E. K. Grossman, *On the linearity of automorphism groups of free groups*, Arch. Math. **38** (1982), 404–409.

[FP] E. Formanek and C. Procesi, *The automorphism group of a free group is not linear*, J. Algebra **149** (1992), 494–499.

[G] S. M. Gersten, *Coherence in doubled groups*, Communications in Algebra **9**(18) (1981), 1893–1900.

[HM] S.M. Hermiller and J. Meier, *Artin groups, rewriting systems and three-manifolds*, J. Pure Appl. Algebra **136** (1999), 141–156.

[KP] R. P. Kent IV and D. Peifer, *A geometric and algebraic description of annular braid groups*, Int. J. Alg. Comp. **12** (2002), 85–97.

[La] S. S. F. Lambropoulou, *Solid torus links and Hecke algebras of $\mathcal{B}$-type*, Proceedings of the Conference on Quantum Topology, ed. D.N. Yetter, World Scientific, River City, New Jersey, 1994, 225–245.

[Le] H. van der Lek, *The homotopy type of complex hyperplane complements*, Ph.D. thesis, University of Nijmegen, Nijmegen, 1983.

[Ma] T. A. Makanina, *Occurrence problem for braid groups B_{n+1} with $n+1 \geq 5$*, Math. Notes Acad. Sci. USSR **29** (1), (1981), 16–17.

[Me] J. Meier, *Geometric invariants of Artin groups*, Proc. London Math. Soc. (3) **74** (1997), 151–173.

[Sc] G. P. Scott, *Finitely generated 3-manifold groups are finitely presented*, J. London Math. Soc. (2) **6** (1973), 437–440.

[Sq1] C. G. Squier, *On certain 3-generator Artin groups*, Trans. Amer. Math. Soc. **302** (1987), 117–124.

[Sq2] C. G. Squier, *The homological algebra of Artin groups*, Math. Scand. **75** (1995), 5–43.

[St1] J.R. Stallings, *On fibering certain 3-manifolds*, Topology of 3-Manifolds and Related Topics, ed. M.K. Fort, Jr., Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962, 95–100.

[St2] J. R. Stallings, *An extension theorem for Euler characteristics of groups*, preprint.

[W] C. T. C. Wall, *Problems*, in Homological and Combinatorial Techniques in Group Theory, London Math. Soc. Lecture Note Series **36**, Cambridge University Press, 1979.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

POSITIVE HEEGAARD DIAGRAMS

JOHN HEMPEL

Abstract. Every (compact, orientable) 3-manifold, M, can be represented by a positive Heegaard diagram: a closed, oriented surface S together with a pair (X, Y) of compact 1-manifolds in S whose components serve as attaching curves for the 2-handles of the two sides of a Heegaard splitting for M and for which the oriented intersection number of X with Y is $+1$ at each point. Such a diagram is completely determined by the two permutations of the intersection points of X with Y given by flowing from one point to the next in the positive direction along X and Y respectively. Montesinos observed that these permutations also describe the (monodromy) representation of M as a branched cover of the 3-sphere branched over a certain universal graph. In this paper we study 3-manifolds in terms of the combinatorics related to the corresponding permutations. We derive "moves" sufficient to connect any two permutational representations of the same 3-manifold, give a procedure for generating all positive diagrams of a given genus in terms of a finite number of "carrier" graphs, and analyze the lattice of branched covers over the associated universal graph – including an explanation of why "good" properties of 3-manifolds proliferate upwards in this lattice.

0. Introduction

There are many examples known of *universal branch sets* — graphs Γ in the 3-sphere $\mathbb{S}^3$ with the property that every closed, oriented 3-manifold, M, is a finite sheeted branched covering $p : M \to \mathbb{S}^3$ branched over Γ [M], [HLM].

For such a branched covering of degree d we have a representation

$$\varphi : \pi_1(\mathbb{S}^3 - \Gamma) \to S_d$$

to the symmetric group on d symbols given by the action of the fundamental group (by path lifting) on the fiber over a regular point and called the *monodromy* of the covering. We multiply permutations from left to right; so we write them acting on the right as exponents:

$$i \mapsto i^\sigma$$

Two such coverings are equivalent if and only if their monodromy representations are conjugate. The monodromy in turn determines an index d subgroup:

$$\varphi^{-1}(\mathrm{Stab}(1)) \subset \pi_1(\mathbb{S}^3 - \Gamma)$$

and is conjugate to the action (by right multiplication) on the cosets of this subgroup.

2000 *Mathematics Subject Classification*: 57N10, 57M12.
Keywords and phrases: Heegaard diagram, positive diagram, universal branch set.

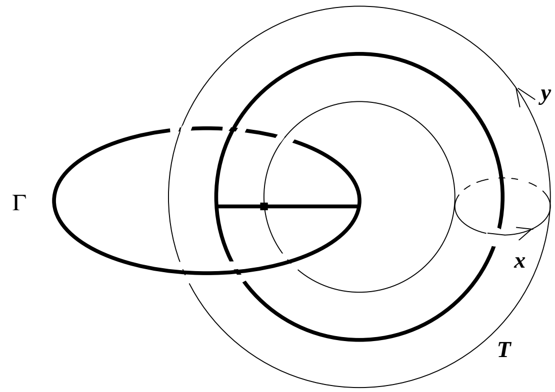


Figure 1. Universal branch set

The set of all coverings branched over a fixed Γ forms a lattice in which

$$\{p_1 : M_1 \to \mathbb{S}^3\} \succeq \{p_2 : M_2 \to \mathbb{S}^3\}$$

if there is a factorization $p_1 = p \circ p_2$ for a branched covering $p : M_1 \to M_2$. This lattice is isomorphic to the lattice of conjugacy classes of finite index subgroups of $\pi_1(\mathbb{S}^3 - \Gamma)$. For Γ universal this lattice provides a description of all 3-manifolds which I find to be challenging.

In this paper we will concentrate on the graph Γ shown in Figure 1. This graph was introduced by Montesinos [M] who showed it to be universal and derived some of its properties. There are two important advantages of this universal graph:

First of all $\pi_1(\mathbb{S}^3 - \Gamma)$ is a free group freely generated by the elements x, and y indicated in Figure 1. Thus for any pair $\sigma, \tau \in S_d$ of permutations we have a branched covering

$$\hat{M}_{\sigma,\tau} \text{ with monodromy given by } x \mapsto \sigma, \quad y \mapsto \tau.$$

Second there is, as noted in [M], a correspondence between the branched coverings of 3-manifolds with this universal branch set and "positive" Heegaard diagrams. This is the starting point of this paper. We will use this correspondence in section 2 to give necessary and sufficient conditions that two pairs $(\sigma_1, \tau_1), \quad (\sigma_2, \tau_2)$ determine the same 3-manifold. The answer will be in the form of a set of combinatorially defined "moves" which generate the associated equivalence on pairs of permutations. This provides a "calculus" (but not a solution) for the enumeration and classification problems for 3-manifolds which is easy to implement — I will provide a Macintosh program doing this on request.

Actually much of 3-manifold theory has a pleasant interpretation in the combinatorial setting of permutations and we will develop some of this here.

In section 3 we describe a way of generating all positive Heegaard diagrams: for each genus g there is a finite number of families of genus g positive diagrams with each family being "carried" by a fixed configuration from which the corresponding families of permutation pairs can be readily determined.

In section 4 we show how the lattice structure is reflected in the combinatorial setting. This is significant as many “good” properties proliferate upwards in this lattice – see Proposition (4.3).

We stress that, since Γ is not a manifold, the branched covers $\hat{M}_{\sigma,\tau}$ will, in general, only be pseudo 3-manifolds. The links of vertices lying over the singular vertices of Γ will be surfaces (which cover S^2 with three branch points). $\hat{M}_{\sigma,\tau}$ will be a 3-manifold precisely when it has zero Euler characteristic. This can be calculated in terms of σ and τ as follows. For $\alpha, \beta, \dots \in S_d$ let

$$c(\alpha, \beta, \dots) = \text{ number of orbits in } S_d \text{ of } gp(\alpha, \beta, \dots),$$

the group generated by $\alpha, \beta, \dots$. In particular for a single element α, $c(\alpha)$ is the number of cycles of α. The Euler characteristic is [M]

$$\chi(\hat{M}_{\sigma,\tau}) = d + c(\sigma, \tau\sigma\tau^{-1}) + c(\tau, \sigma\tau\sigma^{-1}) - c(\sigma) - c(\tau) - c([\sigma, \tau])$$

Thus the condition is given by:

Proposition (0.1). [M] *$\hat{M}_{\sigma,\tau}$ is a 3-manifold if and only if*

$$c(\sigma) + c(\tau) + c([\sigma, \tau]) = d + c(\sigma, \tau\sigma\tau^{-1}) + c(\tau, \sigma\tau\sigma^{-1}).$$

We stress that we are not assuming that σ, τ is a transitive pair; thus $\hat{M}_{\sigma,\tau}$ need not be connected; for $\hat{M}_{\sigma,\tau}$ will have $c(\sigma, \tau)$ components. They will all inherit an orientation from a fixed orientation of $\mathbb{S}^3$, and so we get a well defined compact 3-manifold, $M_{\sigma,\tau}$, obtained by removing an open regular neighborhood of the (finite) set of non-manifold points and then taking the connected sum of the components of the result.

This generality is advantageous because the natural notion of equivalence for permutation pairs passes through non transitive pairs and corresponds to equivalence of the $M_{\sigma,\tau}$'s (not the $\hat{M}_{\sigma,\tau}$) —and in particular provides a means of detecting a connected sum decomposition or recognizing $\mathbb{S}^3$ (see Example (2.7). While it is true that every (compact, oriented, with no 2-sphere boundary components) 3-manifold is represented as $M_{\sigma,\tau}$ for a transitive pair (σ, τ), this may disguise the nature of the manifold— it may be better to represent it as $M_{\sigma,\tau}$ for some non transitive pair. Some of the material in section 1 is an extension of the results of [M] to this more general setting. However what is called $M_{\sigma,\tau}$ in [M] is what we have chosen to call $\hat{M}_{\sigma,\tau}$.

We will primarily be interested in the case in which $M_{\sigma,\tau}$ is a closed 3-manifold, i.e. when σ, τ satisfy the condition of Proposition (0.1) which we will refer to as the *closed condition.* Some of this material is useful (with some complications) for compact 3-manifolds with boundary and we will keep this generality when convenient, but will revert to the closed case when the complications become distracting.

We will frequently use the construction of splitting a manifold M along a codimension one, properly embedded, 2-sided submanifold N. *M split along N* will denote a manifold M_1 homeomorphic to $M - N \times (-1, 1)$ for an appropriate product neighborhood $N \times [-1, 1]$ of N in M and is characterized by the property that there are disjoint subsets N_+ and N_- in ∂M_1 and a map $f : M_1 \to M$ which takes $M_1 - N_+ \cup N_-$ homeomorphically onto $M - N$ and takes each of

N_+ and N_- homeomorphically onto N. Often we will suppress mention of this homeomorphism and simply identify the appropriate subsets of M_1 and M.

1. Diagrams

By a *(Heegaard) diagram* we mean an ordered triple $(S; X, Y)$ where S is a closed, orientable, connected surface and X and Y are compact 1-manifolds in S which are in relative general position and for which no component of $S - X \cup Y$ is a "bigon" — a disk whose boundary is the union of an arc in X and an arc in Y. This definition is more liberal than some. For example it allows X (or Y) to have "superfluous curves" — i.e. some subset of components of X could bound a planar surface in S. The reason for this is the correspondence (Proposition (1.2) below) between diagrams and pairs of permutations under which the diagram for $M_{\sigma,\tau}$ associated with (σ, τ) will most likely have superfluous curves. When we are just interested in the topological structure of $M_{\sigma,\tau}$, and not the particular branched covering, we will eliminate superfluous curves.

An *oriented diagram* is one in which S, X, and Y are all given specific orientations.

A diagram gives rise to a *(Heegaard) splitting* of a 3-manifold M obtained by adding 2-handles to $S \times [-1, 1]$ along the curves of $X \times \{-1\}$ and $Y \times \{1\}$ and then adding 3-handles along all resulting 2-sphere boundary components. If the diagram is oriented, we take the corresponding orientation of M to be the one for which S is oriented from the X-side. We use the term *efficient diagram* for one such that no proper sub diagram determines the same 3-manifold. There can be no superfluous curves in an efficient diagram. A diagram represents a closed 3-manifold if and only if both $S - X$ and $S - Y$ are planar. We call such a diagram *closed*. In an efficient diagram for a closed, connected 3-manifold $S - X$ and $S - Y$ will be connected and X and Y will each have $g = g(S)$ components.

A *positive diagram* is an oriented diagram in which the intersection number $\langle X, Y\rangle_p$ of X with Y is $+1$ at each point $p \in X \cap Y$.

Every compact, oriented 3-manifold with no 2-sphere boundary components is represented by a positive diagram. One can start with an arbitrary diagram for the manifold and eliminate negative crossings by adding trivial handles. The new curves associated with the trivial handle can be oriented so as to introduce only positive crossings. One would expect that the minimal genus of a positive diagram for a given 3-manifold would be much greater than the minimal genus among all diagrams. We have worked this out for Seifert fibered 3-manifolds [H]. Here the *Heegaard genus* and the *positive Heegaard genus* (somewhat surprisingly) turn out to be generically the same with the difference never more than two.

Given (positive) diagrams $D_i = (S_i; X_i, Y_i)$, i = 1,2, we can form a (oriented) connected sum

$$D_1\#D_2 = (S_1\#S_2; X_1 \cup X_2, Y_1 \cup Y_2)$$

of the two diagrams by choosing the "summing" disks $E_i \subset S_i - (X_i \cup Y_i)$. This certainly involves some choices, but also leads to other complications when dealing with manifolds with boundary. If D_i determines the manifold M_i then $D_1\#D_2$ determines a manifold M which is homeomorphic to a (possibly proper) subset of $M_1\#M_2$ — the difference could be one or two 2-handles. The condition for equality is that at least one of A_1, A_2 and one of B_1, B_2 be planar where A_i (B_i) is the component of $S_i - X_i$ ($S_i - Y_i$) containing E_i. If one of the M_i is closed this is automatic. Consequently we will only allow taking the connected sum of two diagrams when at least one of them represents a closed 3-manifold. In this case a (any) connected sum of D_1 and D_2 represents $M_1 \# M_2$.

We use the term *isomorphism* to denote equivalence in the category appropriate to the objects to which it is applied. In most cases this should be clear: for oriented manifolds it will mean orientation preserving homeomorphism, etc. For positive diagrams, we spell it out. Isomorphism will be the smallest equivalence relation containing all homeomorphisms of ordered triples which preserves all orientations and modulo which allowable oriented connected sum is well defined.

Given a pair (σ, τ) of permutations which is either transitive or satisfies the closed condition we get a positive diagram

$$D(\sigma, \tau) = (S_{\sigma,\tau}; X_{\sigma,\tau}, Y_{\sigma,\tau}) \text{ for } M_{\sigma,\tau}$$

as follows. Consider the branched covering $p : \hat{M}_{\sigma,\tau} \to \mathbb{S}^3$. Observe the torus T in Figure 1 which meets Γ in a single point, and note that $(T; x, y)$ is a positive diagram for $\mathbb{S}^3$. We pull this back via p to get $S = p^{-1}(T)$, $X = p^{-1}(x)$, $Y = p^{-1}(y)$. The positivity condition follows because x meets y at a single point (which we choose as base point) with $+1$ intersection. Note that $\hat{M}_{\sigma,\tau}$ can be constructed by adding 2-handles to $S \times [-1, 1]$ along the curves of $X \times \{-1\}$ and $Y \times \{1\}$ and then adding cones over the resulting boundary components.

If (σ, τ) is a transitive pair, we let $D(\sigma, \tau) = (S; X, Y)$. If not then, by hypothesis, $\hat{M}_{\sigma,\tau}$ is a closed 3-manifold each component of which contains a single component of S which determines a positive diagram for that component. Then we get $D(\sigma, \tau)$ by taking a (allowable) connected sum of these diagrams.

Observe that the components of X (respectively Y) are in one to one correspondence with the cycles of σ (respectively τ), the fiber over the base point is $p^{-1}(x \cap y) = X \cap Y$, and σ (τ) is given by flowing along X (Y) in the positive direction from one intersection point to the next. The components of $S - (X \cup Y)$ are in one to one correspondence with the cycles of $[\sigma, \tau]$: each is an open disk which is a cyclic branched cover, via p, of $T - (x \cup y)$ of degree the length of the cycle. If the pair is not transitive then the components of $S_{\sigma,\tau} - N(X \cup Y)$ will be planar (and will depend on "where" the connected sum is done) and their boundary curves will be in one to one correspondence with the cycles of $[\sigma, \tau]$. In any event we have:

Proposition (1.1). *For $\sigma, \tau \in S_d$ either a transitive pair or one satisfying the closed condition the surface $S_{\sigma,\tau}$ has genus*

$$g_{\sigma,\tau} = c(\sigma, \tau) + 1/2(d - c([\sigma, \tau])).$$

Proof. As in the above discussion $S = p^{-1}(T)$ is a branched cover of T branched over a single point whose inverse image has $c([\sigma, \tau])$ points. By the Riemann-Hurwitz formula $\chi(S) = c([\sigma, \tau]) - d$. The conclusion follows from the facts that S has $c(\sigma, \tau)$ components whose connected sum is $S_{\sigma,\tau}$. □

We use the above conditions to reconstruct σ and τ from the diagram. Given any positive diagram $D = (S; X, Y)$ flow along X and Y respectively defines permutations

$$\sigma(D), \tau(D) \in S_d; \quad d = \#(X \cap Y)$$

which depend on labeling the points of $X \cap Y$ by the integers $1, 2, \ldots, d$, but are well defined up to conjugacy. If, as in the genus one diagram for $S^2 \times S^1$, $X \cap Y = \emptyset$, we make the convention that $(\sigma(D), \tau(D))$ is the empty pair. We also make the convention that the diagram corresponding to the empty pair is the genus 0 diagram $(S^2; \emptyset, \emptyset)$ for $\mathbb{S}^3$.

These constructions are essentially inverse; the discrepancy is entirely due to $S^2 \times S^1$ summands as we explain in the following two propositions. From now on it will be simpler to restrict to closed manifolds. We say a diagram $D = (S; X, Y)$ is *good* if:

1. It is closed i.e. $S - X$ and $S - Y$ are planar,
2. Each simple closed curve in $S - (X \cup Y)$ separates S, and
3. Neither X nor Y has *isolated components* — components of X which don't meet Y or vice versa.

Note that in a diagram satisfying 1. and 2. an isolated component would be (separating and) superfluous.

A *good* diagram is called *very good* if $X \cup Y$ is connected — equivalently if each component of $S - X \cup Y$ is an open 2-cell. It is easy to see that every good diagram is a connected sum of very good diagrams. The summands are unique up to isomorphism. Each is a regular neighborhood of a component of $X \cup Y$ with its boundary components capped off with 2-cells. We regard two good diagrams as equivalent if they have isomorphic very good summands. Clearly equivalent diagrams represent the same manifold.

The above observations are summarized in

PROPOSITION (1.2). *The correspondence $(\sigma, \tau) \mapsto D(\sigma, \tau)$ induces a one to one correspondence between the set of conjugacy classes of pairs of permutations satisfying the closed condition and the set of equivalence classes of good positive diagrams. The inverse correspondence is $D \mapsto (\sigma(D), \tau(D))$.*

We could add copies of the genus 1 splitting of $S^2 \times S^1$ to $D(\sigma, \tau)$ to get a positive diagram for $M_{\sigma,\tau} \# S^2 \times S^1 \# \ldots$ which determines the same pair of permutations. The converse is true:

PROPOSITION (1.3). *If $D = (S; X, Y)$ is a positive diagram of genus g for a closed 3-manifold M then $(\sigma(D), \tau(D))$ satisfies the closed condition and M is isomorphic to $M_{\sigma(D),\tau(D)} \#(g - g_{\sigma(D),\tau(D)}) S^2 \times S^1$.*

Proof. First suppose that D is a good diagram. Then by the above we see that D and $D(\sigma(D), \tau(D))$ are equivalent and M is isomorphic to $M_{\sigma(D),\tau(D)}$.

If D is not good, there is a simple closed curve $J \subset S - X \cup Y$ which does not separate S. Since M is closed, J must bound disks on both sides of S. The union of these disks is a non- separating 2-sphere in M. Splitting the pair (M, S) along this 2-sphere and capping off the boundary components gives a 3-manifold M_1 represented by a positive diagram D_1 of genus $g - 1$ with M isomorphic to $M_1 \# S^2 \times S^1$ and $\sigma(D_1), \tau(D_1)$ the same as $\sigma(D), \tau(D)$. So we can complete the proof by induction on $g - g(\sigma, \tau)$. ∎

For $D(\sigma, \tau) = (S; X, Y)$ we have noted that the orbits of σ (respectively τ) are in one to one correspondence with the components of X (respectively Y). It is convenient to identify the integers $1, 2, \ldots, d$ with the points of $X \cap Y$ which they index. So for a cycle $(i_1, i_2, \ldots, i_k)$ of σ there is a corresponding component of X which contains $i_1, i_2, \ldots, i_k$ in the indicated order.

We can identify the components of $S - X$ (and of $S - Y$) in a similar manner. For C a component of $S - X$ we distinguish between the *topological boundary* $Bd(C) = \overline{C} \cap \overline{(X - C)}$ and the *combinatorial boundary* $\partial C = Bd(\overline{C})$. Each is the union of components of X, but some components of X may have C on both sides and so lie in $Bd(C) - \partial C$. If Z is a component of X then we say that Z is *positively oriented* by C if either $Z \subset Bd(C) - \partial C$ or $Z \subset \partial C$ and its orientation agrees with that induced from C. We let

$$\mathcal{O}(C) = \{i \in X \cap Y : \text{the component of } X \text{ containing } i \text{ is positively oriented by } C\}.$$

Proposition (1.4). *For a transitive pair (σ, τ) the function $\mathcal{O}$ gives a one to one correspondence between the components of $S_{\sigma,\tau} - X_{\sigma,\tau}$ and the orbits of $gp(\sigma, \tau\sigma\tau^{-1})$. Similarly there is a one to one correspondence between the components of $S_{\sigma,\tau} - Y_{\sigma,\tau}$ and the orbits of $gp(\tau, \sigma\tau\sigma^{-1})$.*

Proof. Let $S = S_{\sigma,\tau}$, etc. If D is a component of $S - (X \cup Y)$ then D is a cyclic branched cover of the "square" $T - (x \cup y)$. In particular $Bd(D)$ is connected. A component of $S - X$ is obtained by (maximally) pasting together such $D's$ along edges in Y. At a point $i \in X \cap Y$ the pair (X, Y) gives a local coordinate system for S, and $i \in \mathcal{O}(C)$ exactly when the first and second quadrants lie in C. The sets $\mathcal{O}(C)$ partition $\{1, 2, \ldots, d\}$ and are in one to one correspondance with the components C. If the first quadrant at $i \in X \cap Y$ lies in a component D of $S - X \cup Y$, then the first quadrant at i^σ lies in a component D' of $S - X \cup Y$ sharing an edge in Y with D, and the second quadrant at $i^{\tau\sigma\tau^{-1}}$ lies in D. Thus an orbit of $gp\{\sigma, \tau\sigma, \tau^{-1}\}$ lies in a single $\mathcal{O}(C)$. By moving from one component of $S - X \cup Y$ to an adjacent one shairing an edge in Y we see that any two points of $\mathcal{O}(C) \cap D$ lie in the same orbit of $gp(\sigma, \tau\sigma\tau^{-1})$. $\square$

A Heegaard diagram for a closed 3-manifold always gives rise to a pair of "dual" presentations for its fundamental group. When (σ, τ) is an efficient representation of the closed 3-manifold $M_{\sigma,\tau}$, these presentations can be written down as follows. Order the cycles $\sigma_1, \sigma_2, \ldots, \sigma_g$ of σ and $\tau_1, \tau_2, \ldots, \tau_g$ of τ. For $i = 1, 2, \ldots, d$ let $s(i)$ be the index of the cycle of σ containing i. The presentation $\mathcal{P}_X$ has generators $a_1, a_2, \ldots, a_g$ dual to the components of X and for each cycle $(i_1, i_2, \ldots, i_r)$ of τ a relation

$$a_{s(i_1)} a_{s(i_2)} \ldots a_{s(i_r)} = 1$$

The presentation $\mathcal{P}_Y$ is similarly obtained by interchanging the roles of σ and τ.

From the above it should be clear that the *incidence matrix*

$$P = (\#(\sigma_i \cap \tau_j))$$

is a presentation matrix for $H_1(M_{\sigma,\tau})$.

2. Equivalence of the $M_{\sigma,\tau}$

The theorem of Reidemeister and Singer [R],[S] asserts that any two Heegaard splittings of a given 3-manifold are *stably equivalent* — become equivalent after adding some trivial handles to each. The purpose of this section is to understand this theorem in the context of positive diagrams and to derive a set of moves on pairs (σ, τ) of permutations which generate the relation of isomorphism of the corresponding 3-manifolds $M_{\sigma,\tau}$. The moves are of five types. Each involves replacing (σ, τ) by a pair (σ_1, τ_1) as described below. Some of these moves are described in terms of adding and/or deleting some elements in certain cycles of σ and τ. We will always assume that the initial permutations are on the symbols $\{1, 2, \ldots, d\}$ and, without saying so, that the new elements are renumbered to be $\{1, 2, \ldots,$ (new) $d\}$. We assume throughout this section that (σ, τ) satisfies the closed condition.

0. Superfluous cycle deletion. Let s be a cycle of σ and x the corresponding component of $X = X_{\sigma,\tau}$. We say that s (respectively x) is a *superfluous cycle (curve)* if $(S; X - x, Y)$ is still a (positive) diagram for $M_{\sigma,\tau}$. This will be the case if and only if the components of $S - X$ on opposite sides of x are distinct. Proposition (1.4) then tells us how to recognize superfluous cycles. Move **0** is:

Delete all the elements in a cycle $(i_1, i_2, \ldots, i_k)$ of σ for which i_1 and ${i_1}^{\tau^{-1}}$ (and hence i_j and ${i_j}^{\tau^{-1}}$) are in different orbits of $gp(\sigma, \tau\sigma\tau^{-1})$ to get $(\sigma_1, \tau_1) \in S_{d-k}$.

Similarly one can delete a superfluous cycle of τ.

The result of deleting superfluous cycles changes the manifolds only to the extent allowed by Proposition (1.3):

Proposition (2.1). *If (σ_1, τ_1) is obtained from a pair (σ, τ) which satisfies the closed condition by deleting some superfluous cycles, then (σ_1, τ_1) satisfies the closed condition and $M_{\sigma,\tau} \cong M_{\sigma_1,\tau_1} \# (g_{\sigma,\tau} - g_{\sigma_1,\tau_1})(S^2 \times S^1)$*

The following example illustrates how this can happen.

Example (2.2).

$$\sigma = (1\,2)(3\,4)(5\,6)(7\,8\,9\,10), \quad \tau = (1\,3\,9\,5\,8)(2\,7)(4\,10\,6)$$

represents $S^2 \times S^1$. If we delete the superfluous cycle $(7\,8\,9\,10)$ from σ we get

$$\sigma_1 = (1\,2)(3\,4)(5\,6), \quad \tau_1 = (1\,3\,5)(2)(4\,6)$$

which represents $\mathbb{S}^3$. Observe that after deleting the superfluous curve corresponding to this cycle, the resulting diagram is no longer good.

We say that (σ, τ) is an *efficient representation* of a closed 3-manifold M if $M \cong M_{\sigma,\tau}$ and $D(\sigma, \tau)$ is an efficient diagram — equivalently

$$c(\sigma) = c(\tau) = c(\sigma, \tau) + 1/2(d - c([\sigma, \tau])).$$

and yet equivalently

$$c(\sigma, \tau) = c(\sigma, \tau\sigma\tau^{-1}) = c(\tau, \sigma\tau\sigma^{-1})$$

If we start with any representation (σ, τ) of M and delete superfluous cycles in some order, we will get to an efficient representation (σ_1, τ_1) which also represents

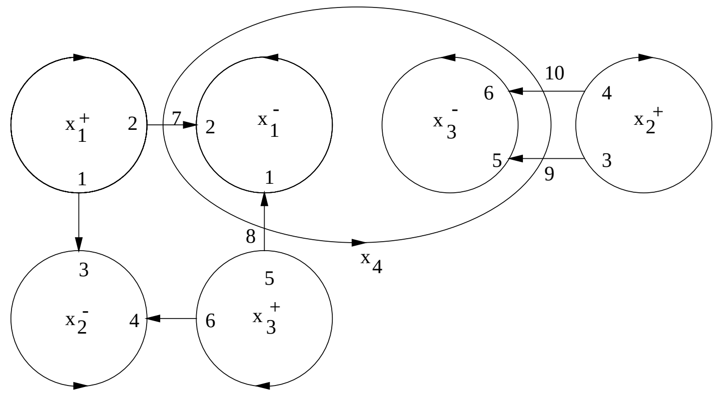


Figure 2. Example 2.2

M unless $g_{\sigma_1,\tau_1} < g_{\sigma,\tau}$. If we "add" in $g_{\sigma,\tau} - g_{\sigma_1,\tau_1}$ copies (on distinct symbols) of the efficient representation

$$\sigma = (1\,2\,3)(4\,5\,6), \quad \tau = (1\,3\,6\,4)(2\,5)$$

of $S^2 \times S^1$ (because there is a trivial handle and the first homology is $\mathbb{Z}$), we get an efficient representation of M.

We will never have to consider the inverse move of inserting a superfluous cycle: in what follows we will establish the equivalence between efficient representations of a 3-manifold by moves which keep efficient representations at every stage. This is the reason we have introduced it first — to keep it separate from the other moves.

I. Elementary equivalence.

A. *Conjugation*:

$$(\sigma_1, \tau_1) = (\mu\sigma\mu^{-1}, \mu\tau\mu^{-1}); \mu \in S_d.$$

B. *Inversion*:

$$(\sigma_1, \tau_1) = (\sigma^{-1}, \tau^{-1}).$$

C. *Exchange*:

$$(\sigma_1, \tau_1) = (\tau, \sigma).$$

It should be clear that each of these induces an isomorphism between the associated manifolds. The exchange will reverse orientation of the splitting surfaces while reversing the sides of the splittings.

It is interesting to note that exchange is not actually needed. This is implicit in the use of the Reidemeister-Singer Theorem which allows one stably to interchange the two sides of a Heegaard splitting by an orientation preserving automorphism of the underlying manifold even though this is not in general possible to do without stabilization. However it seems better to keep it as a basic move. In particular, it allows us to get by with stating only the "σ-side" version of the moves.

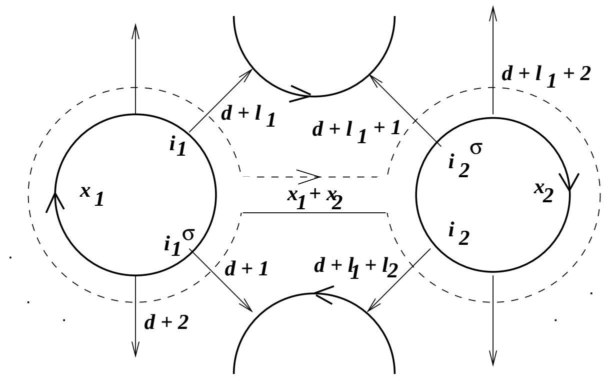


Figure 3. Sum of cycles

II. Trivial handle insertion/cancellation.

A. *Insertion*: Add a fixed point to both σ and τ:

$$(\sigma_1, \tau_1) = (\sigma \centerdot (d+1), \tau \centerdot (d+1))$$

B. *Cancellation*: Delete a common fixed point of σ and τ to get $(\sigma_1, \tau_1) \in S_{d-1}$.

It should be clear that these moves correspond to adding or removing a canceling pair of handles to the associated positive diagrams.

The rationale for the next move is as follows. We consider the diagram $D(\sigma, \tau) = (S; X, Y)$ and look for two components of X which can be piped together in an orientation preserving manner without creating any additional intersections. Doing this and pushing the result off of X gives a new curve which can be added to X as a superfluous curve (it doesn't separate S as it has positive intersection number with some component of Y). Either of the original curves is now superfluous and can be eliminated. The process gives a new diagram D_1 for the same splitting. We then take $(\sigma_1, \tau_1) = (\sigma(D_1), \tau(D_1))$ which represents the same manifold.

For a transitive representation, the two curves contain intervals in the boundary of some component C of $S - (X \cup Y)$. The new curve is obtained by taking parallel copies of these curves and piping them together in C. This can be done only if C lies on the positive side of both curves (case A) or on the negative side of both curves (case B). For a non transitive representation we could also choose to construct the splitting $D(\sigma, \tau)$ in such a way that two components of X corresponding to cycles of σ in different orbits of $gp(\sigma, \tau)$ are adjacent. The following describes the combinatorics for detecting these and in writing down the resulting permutations. Refer to Figure 3 which illustrates case A.

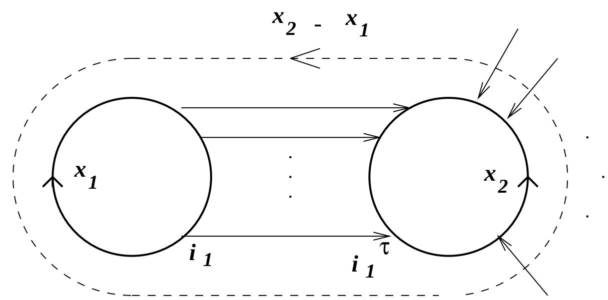


Figure 4. Difference of cycles

III. Adding two cycles of σ.

A. *On the positive side.* Choose distinct cycles s_1, s_2 (with lengths ℓ_1, ℓ_2) of σ which either share elements with some cycle of $[\sigma, \tau]$ or are in different orbits of $gp(\sigma, \tau)$. Choose elements i_1 in s_1, i_2 in s_2 such that, in the first case i_1 and i_2 are in the same cycle of $[\sigma, \tau]$. Then add the cycle

$$(d+1, d+2, \ldots, d+\ell_1+\ell_2)$$

to σ to get $\sigma' \in S_{d+\ell_1+\ell_2}$. Insert $d+n$ between

$${i_1}^{\sigma^n} \text{ and } {i_1}^{\sigma^n\tau} \text{ for } 1 \leq n \leq \ell_1$$

in the appropriate cycle of τ and insert $d+\ell_1+n$ between

$${i_2}^{\sigma^n} \text{ and } {i_2}^{\sigma^n\tau} \text{ for } 1 \leq n \leq \ell_2$$

to get $\tau' \in S_{d+\ell_1+\ell_2}$. Then delete the now superfluous cycle s_1 from (σ', τ') to get $\sigma_1, \tau_1 \in S_{d+\ell_2}$. We refer to this process as *replacing s_1 by s_1+s_2 on the positive side (at $i_1 \in s_1, i_2 \in s_2$).*

B. *On the negative side.* Choose distinct cycles s_1 and s_2 of σ and elements i_1 in s_1 and i_2 in s_2 such that either i_1 and i_2 are both in some cycle of $[\sigma, \tau^{-1}]$ or s_1 and s_2 are in different orbits of $gp(\sigma, \tau)$. Then add the cycle

$$(d+1, d+2, \ldots, d+\ell_1+\ell_2)$$

to σ to get $\sigma' \in S_{d+\ell_1+\ell_2}$. Insert $d+n$ between

$${i_1}^{\sigma^n\tau^{-1}} \text{ and } {i_1}^{\sigma^n} \text{ for } 1 \leq n \leq \ell_1$$

in the appropriate cycle of τ and insert $d+\ell_1+n$ between

$${i_2}^{\sigma^n\tau^{-1}} \text{ and } {i_2}^{\sigma^n} \text{ for } 1 \leq n \leq \ell_2$$

to get $\tau' \in S_{d+\ell_1+\ell_2}$. Then delete the now superfluous cycle s_1 from (σ', τ') to get $\sigma_1, \tau_1 \in S_{d+\ell_2}$. We say (σ_1, τ_1) is obtained by *replacing s_1 by s_1+s_2 on the negative side.*

The next move involves finding a pair x_1, x_2 of components of X so that the curves of Y which leave (or enter) x_1 on a particular side form a parallel family all leading to (coming from) x_2. We tube together these curves along a neighborhood of this parallel family to get a new curve which meets Y positively and which we exchange with x_2 to get a new diagram. This is described at the level of permutations as follows. Refer to Figure 4.

IV. Subtracting one cycle of σ from another.

A. *On the positive side.* Choose a cycle s_1 of σ which contains exactly one element i_1 which is not fixed by $[\sigma, \tau]$ and such that $i_1{}^{\tau}$ lies in a cycle s_2 of σ different from s_1 (if $s_2 = s_1$ we would immediately see that $M_{\sigma,\tau}$ has a lens space summand). It follows that $i^{\tau} \in s_2$ for every $i \in s_1$, and that $\ell_2 = \ell(s_2) \geq \ell_1 = \ell(s_1)$. In fact we must have $\ell_2 > \ell_1$; otherwise the curve we get from tubing the components of X is a non separating curve lying in $S - X \cup Y$. Delete the image, under τ, of the elements of s_1 from the cycles of σ and τ containing them to get $\sigma_1, \tau_1 \in S_{d-\ell_1}$. We say (σ_1, τ_1) is obtained by *replacing* s_2 *by* $s_2 - s_1$ *on the positive side.*

B. *On the negative side.* Choose a cycle s_1 of σ which contains exactly one element i_1 which is not fixed by $[\sigma, \tau^{-1}]$ and such that $i_1{}^{\tau^{-1}}$ lies in a cycle s_2 of σ different from s_1. It follows that $i^{\tau^{-1}} \in s_2$ for every $i \in s_1$, and that $\ell_2 = \ell(s_2) > \ell_1 = \ell(s_1)$. Delete the the image, under τ^{-1}, of the elements of s_1 from the cycles of σ and τ containing them to get to get $\sigma_1, \tau_1 \in S_{d-\ell_1}$. We say (σ_1, τ_1) is obtained by *replacing* s_2 *by* $s_2 - s_1$ *on the negative side.*

Note that "positive/negative side" in cycle subtraction refers to the positive side of (the curve corresponding to s_1); so $s_1 - s_2$ on positive/negative side will lie on the positive/negative side of s_1 (and on the negative/positive side of s_2). With this (arbitrary) convention the inverse of $s_1 \to s_1' = s_2 + s_1$ on the positive/negative side is $s_1' \to s_1' - s_2'$ on the positive/negative side. Also note that the two move sequence of replacing s_2 by $s_2 - s_1$ and then replacing s_1 by $s_1 + (s_2 - s_1)$ accomplishes the replacement of s_1 by $s_2 - s_1$.

Since each of the moves II — IV is based on describing a new good, positive diagram for the same splitting, we have

THEOREM (2.3). *Suppose (σ, τ) satisfies the closed condition and that (σ', τ') is obtained from (σ, τ) by a sequence of moves of types I through IV. Then (σ', τ') satisfies the closed condition and*

$$M_{\sigma,\tau} \cong M_{\sigma',\tau'}$$

Now we consider the converse. Suppose that $D = (S; X, Y)$ is a positive diagram for a closed 3-manifold M, and z is an simple closed curve in $S - X$ which meets Y positively at each point and which separates the two copies of some component x_1 of X in S cut open along X. We say that the diagram

$$D^* = (S; X - x_1 \cup z, Y)$$

is obtained from D by replacing x_1 by z.

LEMMA (2.4). *If (σ, τ) is an efficient representation for a closed 3-manifold M and D^* is obtained from $D = D(\sigma, \tau)$ by replacing a component x_1 of $X = X_{\sigma,\tau}$ by a curve $z \subset S - X$ as above, then there is a sequence of moves of types III and IV taking (σ, τ) to $(\sigma(D^*), \tau(D^*))$*

Proof. D is a good diagram and z does not separate S; so $z \cap Y \neq \emptyset$. $S - X^*$ is connected and planar. It follows that D^* is a good positive diagram for $M = M_{\sigma,\tau}$. Let P, Q be the components of S split along $X \cup z$ with $\chi(P) \geq \chi(Q)$. We show, by induction on $-\chi(P)$ that there is a sequence of moves of type III and IV taking (σ, τ) to $(\sigma(D^*), \tau(D^*))$.

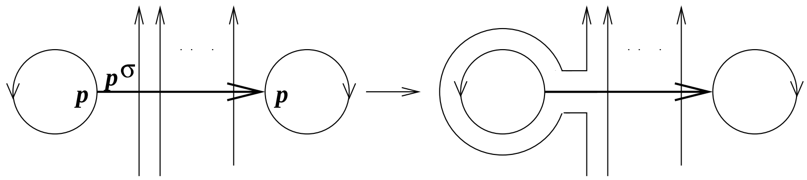


Figure 5

If $\chi(P) = 0$ then D and D^* are isomorphic. If $\chi(P) = -1$, then ∂P consists of z, x_1 and some $x_i \neq x_1$. By positivity not all pairs of components of ∂P can be joined by arcs of $Y \cap P$ nor can any such arc have ends in the same component of ∂P. If there are no arcs of $Y \cap P$ joining x_1 to x_i, then $z = x_1 + x_i$ and replacing x_1 by z is a type III move. If there are no arcs of $Y \cap P$ joining x_i to z, then $z = x_1 - x_i$, and we have a type IV move. In the third case we can replace x_i by $x_i - x_1 = z$ and then replace x_1 by $x_1 + z = x_i$.

so assume $\chi(P) \leq -2$. The components of P split along Y have vertices in ∂P and edges coming alternately from Y and $X \cup z$. By positivity the number of edges is a multiple of four. If any component of ∂P only had edges in squares, P would be an annulus. Thus at least one such region R has at least eight edges with one edge e in x_1. Count around ∂R in a fixed direction to the fourth edge from e to an edge f in ∂P such that the orientations on e and f induced from $X \cup z$ are both the same or both opposite the orientation induced from R. We pipe together the components of ∂P containing e and f along an arc in R then push into R to get an oriented 1-manifold $w \subset \operatorname{Int}(R)$ which meets Y positively at every point. If e and f are in different components of ∂P, then w has a single component which splits P to two regions of larger (negative) Euler characteristic and we can apply induction to replace x_1 by w and then w by z by a sequence of moves. If e and f lie in the same component of ∂P, then w has two components. At least one is not parallel to a component of ∂P ; otherwise $\chi(P) = 2$, and can be used as above to complete the proof by induction. ☐

Lemma (2.5). *Each of the following moves applied to a pair (σ, τ) satisfying the closed condition can be accomplished by a sequence of moves of types I through IV:*

1. **Easy handle cancellation.** *Delete all elements in a cycle of σ which contains a fixed point p of τ.*
2. **Easy handle insertion.** *Add a cycle $(d+1, d+2, \ldots, d+\ell)$ to σ and define $(d+1)^\tau = d+1$. The elements $d+2, \ldots, d+\ell$ will be inserted in existing cycles of τ. The choice for the position of $d+2$ can be arbitrarily specified.*
3. **Meiosis.** *Choose a cycle $(i_1, i_2, \ldots, i_r, \ldots, i_s)$ of σ and replace it by the two cycles $(i_1, \ldots, i_r, d+1)(d+2, i_{r+1}, \ldots, i_s)$. Then add the cycle $(d+3, d+4)$ to σ and add the two cycles $(d+1, d+3)(d+2, d+4)$ to τ.*

Proof. For the first move we replace the cycle t of τ which contains p^σ by the difference $t - (p)$. This reduces the length of the cycle of σ containing p by one (see Figure 5). We repeat this process until p is fixed by both σ and τ. Then we remove the corresponding trivial handle.

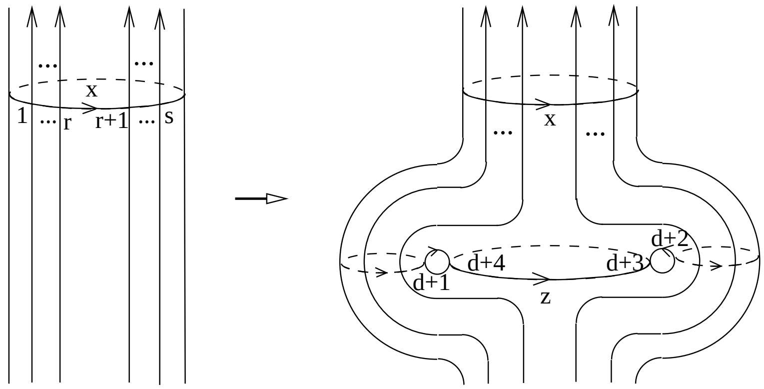


Figure 6. Meiosis

The second move is the inverse of the first. We introduce a fixed point, $d+1$, to both σ and τ. This corresponds to adding a trivial handle which we may suppose is done in a component of $S-(X\cup Y)$ whose boundary contains the interval in Y from q to q^τ. Then replace the cycle of τ containing q by its sum with $(d+1)$. The effect is to insert $d+2$ between q and q^τ in this cycle of τ and to change $(d+1)$ to $(d+1, d+2)$ as a cycle of σ. We can continue to add $(d+1)$ to cycles of τ subject to the rules. We will have no need to be more specific than this in the applications of this move.

The geometry behind the third move is as follows. The given cycle of σ corresponds to a component x of X which bounds a disk E on the X side of the splitting. Take a properly embedded arc a in E which separates the two indicated subsets of ∂E. We get a new splitting for the same manifold by removing a neighborhood of a from one side of the splitting and adding it to the other. This neighborhood splits E into two new meridional disks for the X side of this new splitting, but we need a new meridional disk for the Y side. This introduces a negative intersection which we correct with a a trivial handle. This is accomplished by two easy handle insertions (see Figure 6) corresponding to adding two fixed points to τ followed by replacing x by a curve z aĺa Lemma (2.4). □

Theorem (2.6). *Let (σ,τ) and (σ',τ') be efficient representations of isomorphic closed 3-manifolds. Then (σ,τ) can be transformed into (σ',τ') by a finite sequence of moves of types I through IV.*

Proof. By assumption $M_{\sigma,\tau}$ and $M_{\sigma',\tau'}$ are isomorphic. Take the corresponding positive diagrams $D(\sigma,\tau)=(S;X,Y), D(\sigma',\tau')=(S';X',Y')$. The associated splittings are stably equivalent [R], [S]. Since adding trivial handles is realized by type II moves, there is no loss in assuming that they are already equivalent. Using this equivalence we identify S' with S in a 3-manifold $M=U\cup V$ where U and V are handlebodies with $U\cap V=\partial U=\partial V=S$ and with the components of X and of X' (respectively Y and Y') bounding disks in U (respectively V). In particular $(S;X,Y)$ and $(S;X',Y')$ are positive diagrams for the same splitting of M.

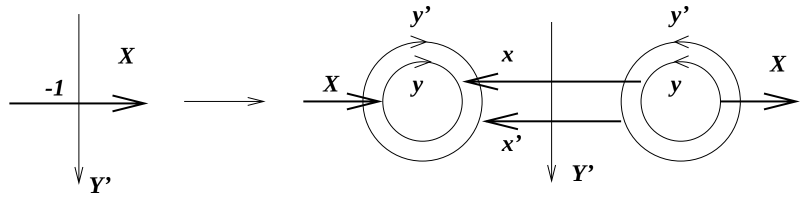


Figure 7

We may assume that all curves are in general position and that X (respectively X') meets Y' (respectively Y) positively at each point. To see this we note that a negative intersection between (say) X and Y' can be eliminated by a trivial handle addition which is simultaneously a easy handle insertion for both systems. This is depicted in Figure 7 which shows replacing X by $X \cup x$, X' by $X' \cup x'$, Y by $Y \cup y$, and Y' by $Y' \cup y'$.

Now suppose that $X \cap X' = Y \cap Y' = \emptyset$. Note that in this case each component of X' meets Y and so on; otherwise we get a non separating simple closed curve of S lying in $S - X \cup Y$. I claim that there are orderings $x_1, x_2, \ldots, x_g$ of the components of X and ${x_1}', {x_2}', \ldots, {x_g}'$ of the components of X' such that for each $i = 0, \ldots, g$

$$D_i = (S; X_i, Y); X_i = {x_1}' \cup \cdots \cup {x_i}' \cup x_{i+1} \cup \cdots \cup x_g$$

is a good positive diagram. Since we are dealing with efficient representations, the curves of X (and those of X') represent a maximal set of linearly independent elements of $H_1(S; \mathbb{Z})$. We merely need to preserve this property for each X_i. This can be done by the (linear algebra) replacement theorem of Steinitz.

Now D_{i+1} is obtained from D_i by replacing x_{i+1} by ${x_{i+1}}'$. By Lemma (2.4) this can be done by type III and IV moves. Repeating this argument for the "Y side" completes the proof in case $X \cap X' = Y \cap Y' = \emptyset$.

In general we take sets E, E' (respectively F, F') whose components are properly embedded disks in U (respectively V) bounded by the components of X, X' (respectively Y, Y') and such that the components of $E \cap E'$ (respectively $F \cap F'$) are properly embedded arcs and induct on the total number of these arcs.

We have already considered the initial case; so suppose (say) that $E \cap E' \neq \emptyset$. Choose a component a of $E \cap E'$. We get a new splitting which corresponds to a Meiosis associated to a on each of the two systems. This corresponds to eliminating a component of $E \cap E'$. One must check that no new intersections need to be introduced and that we preserve the condition that X (X') meets Y' (Y) positively. Lemma (2.5) and induction then complete the proof. $\square$

Example (2.7). See Figure 8.

$$\sigma = (1\,2\,3\,4\,5\,6)(7\,8\,9\,10\,11\,12\,13\,14\,15)\ \tau = (1\,3\,13\,9\,6\,7\,4\,14\,10)(2\,12\,8\,5\,15\,11)$$

is an efficient pair satisfying the closed condition whose degree cannot be reduced by any elementary move. Replacing t_1 by $t_1 + t_2$ (at $0 \in t_1, 15 \in t_2$) gives

$$\sigma' = (1\,2\,17\,3\,4\,5\,20\,6)(7\,8\,19\,9\,10\,11\,16\,12\,18\,13\,14\,15\,21)$$

$$\tau' = (1\,16\,17\,18\,19\,20\,21\,3\,13\,9\,6\,7\,4\,14\,10)(2\,12\,8\,5\,15\,11)$$

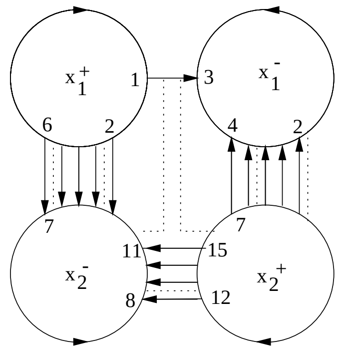


Figure 8. Example 2.7

Replacing s'_2 by $s'_2 - s'_1$ (positive side) gives

$$\sigma'' = (1\,2\,11\,3\,4\,5\,13\,6)(7\,12\,8\,9\,10)\ \tau'' = (1\,11\,12\,13\,3\,8\,6\,4\,9)(2\,7\,5\,10)$$

Replacing s''_2 by $s''_2 - s''_1$ (positive side) gives

$$\sigma''' = (1\,8\,2)(3\,7\,4\,5\,6)\ \tau''' = (1\,7\,8\,4\,2\,5)(3\,6)$$

Replacing s'''_2 by $s'''_2 - s'''_1$ (positive side) gives

$$\sigma^{iv} = (1\,5\,2)(3\,4)\ \tau^{iv} = (1\,5\,2)(3\,4)$$

This is a connected sum of diagrams representing $L_{2,1}\#L_{3,1}$.

In general for σ a p-cycle and q prime to p

$$(\sigma, \sigma^q) \text{ represents the lens space } L_{p,q}$$

We leave the following as an exercise.

Proposition (2.8). *If the incidence matrix of an efficient pair (σ,τ) satisfying the closed condition has an entry p which is the only non zero entry of its row or column, then $M_{\sigma,\tau}$ has a lens space summand $L_{p,q}$ for some q.*

3. Carriers for Positive Diagrams

In this section we show how to decompose the positive diagrams into families which are generated by a single graph, called a *carrier* for the family, from which the family and its associated permutation pairs can be readily computed. We then show how to generate all the carriers. Some authors use the term *Whitehead graph*, but this tends to have a more group theoretic interpretation which may not stress the embedding of the graph in $\mathbb{S}^3$; so we prefer to keep the concepts separate.

To motivate this construction, we start with a positive diagram $(S; X, Y)$ of genus $g \geq 2$. We assume that the diagram is efficient; so X and Y each have g components and that $X \cup Y$ is connected – otherwise we immediately recognize a connected sum of lower genus diagrams. We cut open S along X and collapse the resulting boundary curves to points. This changes S to a 2-sphere and Y to a bipartite graph in S^2 with g source vertices (corresponding to the positive sides of the components of X, g sink vertices, and edges (corresponding to the components of $Y - X$ each directed from a source vertex to a sink vertex. The

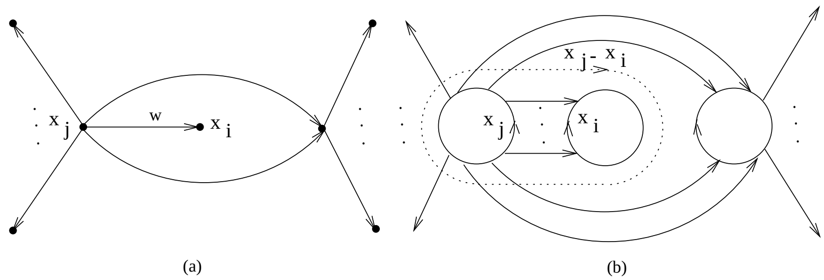


Figure 9

complementary regions will have an even number of edges whose orientations alternate as one traverses the boundary of the region. The regions with two edges (*bigons*) will occur in *stacks*: maximal sequences of bigons each sharing an edge with the next one. We collapse each stack of bigons to a single directed edge to which we assign a weight: the number of original edges which were identified to this edge. The condition $g \geq 2$ prevents us from collapsing S^2 to a graph. Thus we continue to get a bipartite graph in S^2 with g source and g sink vertices. The complementary regions will be $2i$-gons, $i \geq 2$. If some complementary region is a $2i$-gon with $i \geq 3$, we can add an edge in this region directed from a source vertex to a sink vertex three edges distant on the boundary of the region. This arc splits the region into a square and a $2(i-1)$-gon. We associate zero weight to the added edge. By repeating this operation we may assume that all complementary regions are squares. We say that the given positive diagram is *carried by* the resulting graph.

We want to make one additional simplification which will be useful in enumerating carriers. The complementary squares need not have distinct edges. This can happen in just one way — see Figure 9a. There must be an edge with a vertex of order one. Since we are assuming that $X \cup Y$ is connected and $g \geq 2$ the weight on this edge and at least one other weight on an edge at the opposite vertex must be positive. Now back the picture to the level of S cut open along X (Figure 9b.) .

The curve, denoted x_i, corresponding to the order one vertex cannot be paired with the curve, x_j, corresponding to the other end of this edge; as these vertices have different weight sums. Thus we can replace x_j by $x_j - x_i$. The degree $d = \#(X \cap Y)$ is reduced to $d - \#(Y \cap x_i)$. Since the degree cannot be reduced indefinitely, there must result after a finite repetition of the above operation a positive diagram equivalent to the original one which is carried by a graph with no vertices of order one.

We now formalize this by defining a *genus g carrier* to be a connected, bipartite graph $C \subset S^2$ with

• g source vertices and g sink vertices, with

• each component of $S^2 - C$ a disk having boundary the union of four distinct edges of C.

• A bijective pairing $p : V_+(C) \to V_-(C)$ between the source and sink vertices of C, and

• a weight $w(e) \geq 0$ assigned to each edge e of C so that for each $v \in V_+(C)$ we have

$$\sum\{w(e) : v \in e\} = \sum\{w(e) : p(v) \in e\}.$$

Proposition (3.1). *A genus g carrier has*
$4g-4$ edges, and
$2g-2$ complementary regions.

Proof. Since the complementary regions are squares, $E = 2F$.
Also, $F = \beta_0(S^2 - C)$ which by duality is the same as $1 + \beta_1(C) = 2 - \chi(C) = 2 - 2g + E$. □

The preceding discussion has established

Proposition (3.2). *Each efficient, positive diagram of genus $g \geq 2$ is equivalent, via type IV moves, to one which is either carried by a genus g carrier or is a connected sum of lower genus diagrams.*

To reconstruct a diagram from a carrier, we remove neighborhoods of the vertices to obtain a $2g$-punctured sphere S'. We replace each edge e of $E(C)$ by $w(e)$ parallel arcs to obtain a properly embedded 1-manifold $Y' \subset S'$. We identify paired components of $\partial S'$ to obtain an oriented surface S in such a way that Y' maps to a closed 1-manifold $Y \subset S$ and we denote the image of $\partial S'$ by X. We have naturally induced orientations on S, X, and Y so that X meets Y with $+1$ intersection at each point of $X \cap Y$.

The ambiguity in this construction is the amount of *twist* used in gluing the paired components of $\partial S'$. We find it convenient to "mark" one corner of some complementary region of $S^2 - C$ at each vertex to indicate the *zero-twist* gluing.

To determine the corresponding permutations we label the initial points of the components of Y' in some convenient way. Then the cycle σ_i of σ corresponding to the component x_i of X is read from the associated "positive" boundary component of S'. This way σ is determined independent of the twisting.

If τ_0 is the permutation determined by Y for the zero-twist gluing, then the permutation corresponding to Y with twist coordinates $(t_1, \ldots, t_g)$ will be:

$$\tau_0 \sigma_1^{t_1} \ldots \sigma_g^{t_g}.$$

It is not hard, for small genus, to determine the possible carrier graphs (see remark below), and eliminate the pairings which do not admit non trivial weight solutions. In this way we can prove:

Example (3.3). There is just one genus two carrier (Figure 10). The genus two positive diagrams are determined by

$$\sigma = (1, 2, \ldots, q+p)(q+p+1, q+p+2, \ldots, q+2p+r) = \sigma_1 \sigma_2$$

$$\tau = (q+1, q+p+1)(q+2, q+p+2) \ldots (q+p, q+2p) \sigma_1^{t_1} \sigma_2^{t_2}$$

for some $p, q, r \in \mathbb{Z}_+, t_1, t_2 \in \mathbb{Z}$.

There are ten genus three carriers. They are based on all pairings of the two graphs of Figure 11 (modulo the obvious symmetries).

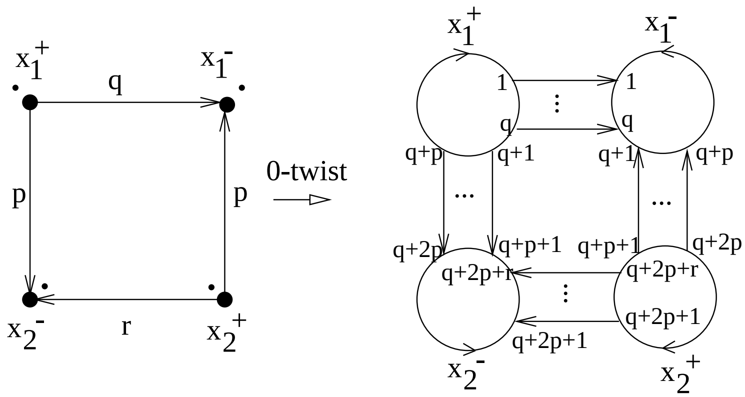


Figure 10. Genus two carrier

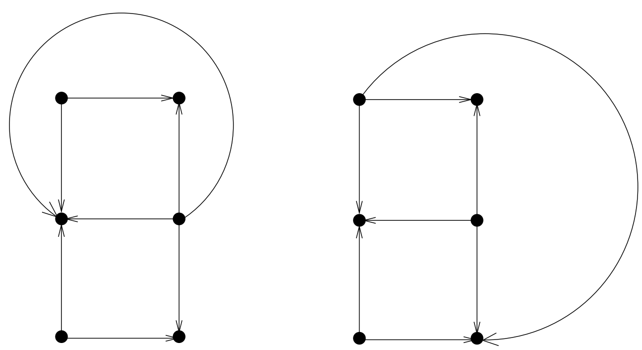


Figure 11. Genus three carriers

Remark. If $C \subset S^2$ is a carrier graph, then there is a branched cover $p : S^2 \to S^2$ branched over three points – all the oriented edges of C are identified and diametrically opposite points on each square are identified. So the source vertices map to one branch point, the sink vertices to the second, and the centers of squares to the third. This process is reversible:

Proposition (3.4). *The genus g carrier graphs are determined by the fixed point free, transitive pairs $\mu, \nu \in S_{4g-4}$ with $c(\mu) = c(\nu) = g$ and $\mu\nu$ a product of $2g-2$ 2-cycles.*

4. The lattice structure

Remember that we have a partial order $\hat{M}_{\tilde{\sigma},\tilde{\tau}} \succeq \hat{M}_{\sigma,\tau}$ if there is a factorization $\tilde{p} = p \circ q$:

$$\hat{M}_{\tilde{\sigma},\tilde{\tau}} \xrightarrow{q} \hat{M}_{\sigma,\tau} \xrightarrow{p} \mathbb{S}^3$$

of branched coverings. At the level of permutations this is given by:

Proposition (4.1). *Let $\tilde{\sigma}, \tilde{\tau} \in S_{\tilde{d}}$, $\sigma, \tau \in S_d$. Then $\hat{M}_{\tilde{\sigma},\tilde{\tau}} \succeq \hat{M}_{\sigma,\tau}$ if and only if $\tilde{d} = \lambda d$ for some $\lambda \in \mathbb{Z}$ and there is a λ to one map $\psi : \{1, 2, \ldots, \tilde{d}\} \to$*

$\{1,2,\ldots,d\}$ *such that*

$$(i^{\tilde{\sigma}})^{\psi} = (i^{\psi})^{\sigma} \text{ and } (i^{\tilde{\tau}})^{\psi} = (i^{\psi})^{\tau}$$

for all i.

Proof. We have identifications of the fibers $\tilde{p}^{-1}(x_0)$ with $\{1,2,\ldots,\tilde{d}\}$ and $p^{-1}(x_0)$ with $\{1,2,\ldots,d\}$. So ψ will be induced by q to give necessity.

Conversely, given ψ, we note that the corresponding unbranched coverings (of $\mathbb{S}^3 - \Gamma$) are determined by the subgroups $\tilde{\varphi}^{-1}(\mathrm{Stab}(1)) \subset \varphi^{-1}(\mathrm{Stab}(1))$ of $\pi_1(\mathbb{S}^3 - \Gamma)$. So the existence of q follows. $\square$

Note that with the above proposition it is easy to construct lots of branched covers of a given $\hat{M}_{\sigma,\tau}$. However even if the given space is a closed 3-manifold ($\hat{M}_{\sigma,\tau} = M_{\sigma,\tau}$), the total space will most likely be a non-manifold. We are most interested in determining the true (finite sheeted) covers of a manifold. This is given by

Proposition (4.2). *Let σ,τ and $\tilde{\sigma},\tilde{\tau}$ be transitive pairs with $\hat{M}_{\sigma,\tau} = M_{\sigma,\tau}$ a closed 3-manifold and $\hat{M}_{\tilde{\sigma},\tilde{\tau}} \succeq M_{\sigma,\tau}$.*

Suppose that $c(\tilde{\sigma}) = \lambda c(\sigma), c(\tilde{\tau}) = \lambda c(\tau)$, and $c([\tilde{\sigma},\tilde{\tau}]) = \lambda c([\sigma,\tau])$ for some $\lambda \in \mathbb{Z}$. Then $q : \hat{M}_{\tilde{\sigma},\tilde{\tau}} \to M_{\sigma,\tau}$ is a true covering map, and $\hat{M}_{\tilde{\sigma},\tilde{\tau}} (= M_{\tilde{\sigma},\tilde{\tau}})$ is a closed 3-manifold.

Proof. Note that the map $\psi : \{1,2,\ldots,\tilde{d}\} \to \{1,2,\ldots,d\}$ maps each cycle of $\tilde{\sigma},\tilde{\tau}$, or $[\tilde{\sigma},\tilde{\tau}]$ onto a cycle of σ,τ, or $[\sigma,\tau]$. So we always have $c(\tilde{\sigma}) \leq \lambda c(\sigma)$, $c(\tilde{\tau}) \leq \lambda c(\tau)$, and $c([\tilde{\sigma},\tilde{\tau}]) \leq \lambda c([\sigma,\tau])$, and the condition of the proposition is equivalent to asserting that for each cycle of σ,τ, or $[\sigma,\tau]$ there are λ cycles of $\tilde{\sigma},\tilde{\tau}$, or $[\tilde{\sigma},\tilde{\tau}]$ projecting one-to-one to it.

We note that q will always be a covering on a neighborhood in $S_{\tilde{\sigma},\tilde{\tau}}$ of $\tilde{X} \cup \tilde{Y}$. We get $S_{\tilde{\sigma},\tilde{\tau}}$ from this neighborhood by filling in disks whose boundaries correspond to the cycles of $[\tilde{\sigma},\tilde{\tau}]$. The condition $c([\tilde{\sigma},\tilde{\tau}]) = \lambda c([\sigma,\tau])$ makes q a homeomorphism on the curves – hence on the disks they bound.

Similarly we see that q will be a homeomorphism on the 2-handles of $M_{\tilde{\sigma},\tilde{\tau}}$ to those of $M_{\sigma,\tau}$. The links of the vertices of $M_{\sigma,\tau}$ are 2-spheres and we have seen that q is a covering map on the inverse image of these 2-spheres. Hence the components of their inverse images are 2-spheres mapping homeomorphically by q and so q will take the 3-cells they bound homeomorphically as well. $\square$

The following illustrates that "good" properties for closed 3-manifolds tend to proliferate upwards in the lattice of branched coverings over Γ.

Proposition (4.3). *Let $\hat{M}_{\tilde{\sigma},\tilde{\tau}} = M_{\tilde{\sigma},\tilde{\tau}}$ and $\hat{M}_{\sigma,\tau} = M_{\sigma,\tau}$ be closed 3-manifolds with $M_{\tilde{\sigma},\tilde{\tau}} \succeq M_{\sigma,\tau}$. If $M_{\sigma,\tau}$ has a finite sheeted true cover which has positive first betti number or which contains a closed, 2-sided incompressible surface then so does $M_{\tilde{\sigma},\tilde{\tau}}$.*

Proof. By hypothesis there is a finite sheeted true covering $r : M_{\sigma^*,\tau^*} \to M_{\sigma,\tau}$ whose total space has one of the "good" properties mentioned. We have a pullback diagram:

$$\begin{tikzcd} M_{\bar\sigma,\bar\tau} \arrow[r] \arrow[d] & M_{\sigma^*,\tau^*} \arrow[d, "r"] \\ M_{\tilde\sigma,\tilde\tau} \arrow[r, "q"'] & M_{\sigma,\tau} \end{tikzcd}$$

with monodromy determined by

$$\bar\sigma, \bar\tau \in Aut(S)$$

$$\text{where } S = \{(i,j) \in \{1,\dots,\tilde d\} \times \{1,\dots,d^*\}) : q(i) = r(j)\}$$

(we are identifying points with their labels), and

$$\bar\sigma = \tilde\sigma \times \sigma^*|S, \bar\tau = \tilde\tau \times \tau^*|S.$$

We wish to show that $\hat M_{\bar\sigma,\bar\tau} \to M_{\tilde\sigma,\tilde\tau}$ is a true covering and that (any component of) the total space has the corresponding "good" property. Each cycle $\bar\sigma_i$ of $\bar\sigma$ covers (in the sense of proposition (4.1) cycles $\tilde\sigma_j$ of $\tilde\sigma$ and σ^*_k of σ^* which cover the same cycle σ_m of σ. The lengths of these cycles satisfy $\ell(\bar\sigma_i) = lcm(\ell(\tilde\sigma_j, \ell(\sigma^*_k))$. But $\ell(\sigma{*}_k) = \ell(\sigma_m)$; since r is a true covering, and $\ell(\sigma_m)|\ell(\tilde\sigma_j)$ by (4.1); so $\ell(\bar\sigma_i) = \ell(\tilde\sigma_j)$. The same argument applies to cycles of $\bar\tau$ and of $[\bar\sigma,\bar\tau]$ to give the conditions of proposition (4.2).

Now $\pi_1(M_{\bar\sigma,\bar\tau})$ maps to a subgroup of finite index in $\pi_1(M_{\sigma^*,\tau^*})$. So if $\beta_1(M_{\sigma^*,\tau^*}) > 0$ then $\beta_1(M_{\bar\sigma,\bar\tau}) > 0$.

If M_{σ^*,τ^*} contains a closed, 2-sided incompressible surface F, we pull it back to a closed 2-sided surface $\bar F \subset M_{\bar\sigma,\bar\tau}$. If $\bar F$ does not compress completely we are done. If it does, then using the fact that $\pi_1(F) \to \pi_1(M_{\sigma^*,\tau^*})$ is monic, we see that $\pi_1(\bar F) \to \pi_1(F)$ factors through a free group. So $\bar F \to F$ factors, up to homotopy, through a 1-complex. This is impossible since $H_2(\bar F) \to H_2(F)$ is not zero. ☐

In the next proposition we use the presentation $\mathcal{P}_X$ of $\pi_1(M_{\sigma,\tau})$ described at the end of section 1. In particular we have an indexing $\sigma_1,\dots,\sigma_g$ of the cycles of σ, corresponding generators $a_1,\dots,a_g$ for $\pi_1(M_{\sigma,\tau})$, and $s(i)$ denotes the index of the cycle of σ containing i.

Proposition (4.4). *Let $\sigma,\tau \in S_d$ be an efficient presentation of a closed 3-manifold and let $\rho : \tilde M \to M_{\sigma,\tau}$ be a λ-sheeted covering map with monodromy $\mu : \pi_1(M_{\sigma,\tau}) \to S_\lambda$.*

Then $\tilde M \cong M_{\tilde\sigma,\tilde\tau}$ where $\tilde\sigma,\tilde\tau \in S_{\lambda d} \cong Aut(\{1,\dots,d\} \times \{1,\dots,\lambda\})$ are given by:

$$\tilde\sigma : (i,j) \mapsto (i^\sigma, j)$$

$$\tilde\tau : (i,j) \mapsto (i^\tau, j^{\mu(a_{s(i)})})$$

Proof. We have branched covers $p : M_{\sigma,\tau} \to \mathbb{S}^3$ and $\tilde p = p \circ \rho : \tilde M \to \mathbb{S}^3$ branched over Γ and the associated diagrams $D(\sigma,\tau) = (S; X, Y)$ and $D(\tilde\sigma,\tilde\tau) = (\tilde S; \tilde X, \tilde Y) = (\rho^{-1}(S); \rho^{-1}(X), \rho^{-1}(Y))$. The components of X and Y lift homeomorphically to components of $\tilde X$ and $\tilde Y$ respectively. $S - X$ is connected and each component of $\rho^{-1}(S - X)$ projects homeomorphically via ρ. Thus we can regard μ as permuting these components – in terms of some labeling. Since a_i

is represented by a curve dual to x_i, a lift of this curve which begins in the component of $\tilde{S} - \tilde{X}$ labeled j ends in the component labeled $j^{\mu(a_i)}$.

We have a labeling of the points of $X \cap Y$ consistent with σ, τ. We label the points of $\tilde{X} \cap \tilde{Y}$ as follows. A point $z \in \tilde{X} \cap \tilde{Y}$ is labeled (i, j) provided that $\rho(z)$ is labeled i and the component of $\tilde{S} - \tilde{X}$ lying on the negative side of the component of $\tilde{X}$ containing z is labeled j. It should be clear that in terms of this labeling the permutation given by flowing along $\tilde{X}$ ($\tilde{Y}$) is $\tilde{\sigma}$ ($\tilde{\tau}$). $\square$

Remarks. 1. The above propositions give an alternate way of thinking about the finite representations of $\pi_1(M_{\sigma,\tau})$. A finite representation is determined by a transitive pair $\tilde{\sigma}, \tilde{\tau} \in S_{\lambda d}$, for some λ, satisfying the conditions of (4.1) and (4.2). The representation (to S_λ) is recovered by reversing the proof of (4.4). We note that the corresponding cover $M_{\tilde{\sigma},\tilde{\tau}} \to M_{\sigma,\tau}$ will be regular if and only if the group generated by $\tilde{\sigma}$ and $\tilde{\tau}$ has order λ.

2. It is straightforward, using Proposition (4.4), to determine the maximal abelian cover of $M_{\sigma,\tau}$ (if finite). Successive applications determine the coverings corresponding to the derived series of $\pi_1(M_{\sigma,\tau})$. I have found this quite effective in determining whether a given fundamental group is finite or not. The possible finite fundamental groups of 3-manifolds are known [Mi], [L], and from a comparison of the quotients of the derived series of the given example with those of the known examples one can either guarantee that the group is infinite (no comparison) or determine its order if finite. Of course one would need an effective procedure for settling the triviality problem for these group presentations to make this algorithmic.

Received June 19, 2003

Final version received March 26, 2004

Rice University
Houston, Tx 77005
USA
hempel@math.rice.edu

References

[H] J. Hempel, *Positive diagrams for Seifert fibered spaces*, Topology Appl. **117** (2002), 319–334.

[HLM] H. M. Hilden, M. T. Lozano, and J. M. Montesinos, *On knots that are universal*, Topology **24** (1985), 499–504.

[M] J. M. Montesinos, *Representing 3-manifolds by a universal branching set*, Math. Proc. Camb. Phil. Soc. **94** (1983), 109–123.

[MSY] W. Meeks, L. Simon, and S. T. Yau, *Embedded minimal surfaces, exotic spheres, and manifolds with positive Ricci curvature*, Ann. of Math. **116** (1982), 621–659.

[Mi] J. Milnor, *Groups which act on S^n without fixed points*, Amer. J. Math. **79** (1957), 623–630.

[L] R. Lee, *Semicharacteristic classes*, Topologie **12** (1973), 183–199.

[R] K. Reidemeister, *Zur dreidimensionalen Topologie*, Abh. Math. Sem. U. Hamburg **9** (1933), 189–194.

[S] J. Singer, *Three-dimensional manifolds and their Heegaard diagrams*, Trans. Amer. Math. Soc. **35** (1933), 88-111.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

THE VARIETY OF CHARACTERS IN $\mathrm{PSL}_2(\mathbb{C})$

MICHAEL HEUSENER AND JOAN PORTI

Abstract. We study some basic properties of the variety of characters in $\mathrm{PSL}_2(\mathbb{C})$ of a finitely generated group. In particular we give an interpretation of its points as characters of representations. We construct 3-manifolds whose variety of characters has arbitrarily many components that do not lift to $\mathrm{SL}_2(\mathbb{C})$. We also study the singular locus of the variety of characters of a free group.

1. Introduction

The varieties of representations and characters have many applications in 3-dimensional topology and geometry. The variety of $\mathrm{SL}_2(\mathbb{C})$-characters has been intensively studied since the seminal paper of Culler and Shalen [CS], but for many applications it is more convenient to work with $\mathrm{PSL}_2(\mathbb{C})$ instead of $\mathrm{SL}_2(\mathbb{C})$ (see [BZ] and [BMP] for instance). The purpose of this note is to study some basic properties of the variety of characters in $\mathrm{PSL}_2(\mathbb{C})$. Most of the results of invariant theory that we use can be found in any standard reference (e.g. [KSS], [Kra], [PV]).

Throughout this paper, Γ will denote a finitely generated group.

Definition (1.1). *The set of all representations of* Γ *in* $\mathrm{PSL}_2(\mathbb{C})$ *is denoted by* $R(\Gamma)$ *and it is called the variety of representations of* Γ *in* $\mathrm{PSL}_2(\mathbb{C})$.

The variety of representations $R(\Gamma)$ has a natural structure as an affine algebraic set over the complex numbers given as follows: the group $\mathrm{PSL}_2(\mathbb{C})$ is algebraic (see Section 2). Given a presentation $\Gamma = \langle \gamma_1, \ldots, \gamma_s \mid (r_i)_{i\in I}\rangle$ we have a natural embedding:

$$\begin{array}{rcl} R(\Gamma) & \to & \mathrm{PSL}_2(\mathbb{C}) \times \cdots \times \mathrm{PSL}_2(\mathbb{C}) \\ \rho & \mapsto & (\rho(\gamma_1), \ldots, \rho(\gamma_s)) \end{array}$$

and the defining equations are induced by the relations. This structure can be easily seen to be independent of the presentation. In fact using the isomorphism $\mathrm{PSL}_2(\mathbb{C}) \cong \mathrm{SO}_3(\mathbb{C})$, $R(\Gamma)$ has a structure of an affine set (see Lemma (2.2.1)).

The action of $\mathrm{PSL}_2(\mathbb{C})$ on $R(\Gamma)$ by conjugation is algebraic. The quotient $R(\Gamma)/\mathrm{PSL}_2(\mathbb{C})$ may be not Hausdorff and it is more convenient to consider the algebraic quotient of invariant theory, because $\mathrm{PSL}_2(\mathbb{C})$ is reductive.

Definition (1.2). *The variety of* $\mathrm{PSL}_2(\mathbb{C})$*-characters* $X(\Gamma)$ *is the quotient* $R(\Gamma)//\mathrm{PSL}_2(\mathbb{C})$ *of invariant theory.*

2000 *Mathematics Subject Classification*: 57M50, 57M05, 20C15.

Keywords and phrases: representation spaces; variety of characters; $\mathrm{PSL}_2(\mathbb{C})$.

This definition means that $X(\Gamma)$ is an affine algebraic set together with a regular map $t\colon\ R(\Gamma)\to X(\Gamma)$ which induces an isomorphism

$$t^*\colon\ \mathbb{C}[X(\Gamma)]\to\mathbb{C}[R(\Gamma)]^{\mathrm{PSL}_2(\mathbb{C})}$$

(i.e. the regular functions on $X(\Gamma)$ are precisely the regular functions on $R(\Gamma)$ invariant by conjugation). We will use the notation $R(M)=R(\pi_1 M)$ and $X(M)=X(\pi_1 M)$ if M is a path-connected topological space.

In this paper we study the basic properties of $X(\Gamma)$.

First we explain the name "variety of characters": given a representation $\rho\colon\ \Gamma\to\mathrm{PSL}_2(\mathbb{C})$, its character is the map

$$\begin{array}{rccc}\chi_\rho\colon & \Gamma & \to & \mathbb{C}\\ & \gamma & \mapsto & \mathrm{tr}^2(\rho(\gamma))\end{array}$$

Theorem (1.3). *There is a natural bijection between $X(\Gamma)$ and the set of characters of representations $\rho\in R(\Gamma)$. This bijection maps every $t(\rho)\in X(\Gamma)$ to the character χ_ρ.*

In many cases the representations of $R(\Gamma)$ lift to $\mathrm{SL}_2(\mathbb{C})$, for instance if Γ is a free group. In such a case, $X(\Gamma)$ is just a quotient of the usual variety of characters in $\mathrm{SL}_2(\mathbb{C})$ (See Proposition (4.2.2)). This quotient is the definition already used in [Bur90], [HLM1],[HLM2] and [Ril84] for 2-bridge knot exteriors. The explicit computation for the figure eight knot exterior is found in [GM].

There are cases where representations do not lift to $\mathrm{SL}_2(\mathbb{C})$, for instance the holonomy representation of an orientable hyperbolic 3-orbifold with 2 torsion. The next result proves that there are manifolds with arbitrarily many components of characters that do not lift.

Theorem (1.4). *For every n, there exist a compact irreducible 3-manifold M with ∂M a 2-torus such that $X(M)$ has at least n irreducible one dimensional components whose characters do not lift to* $\mathrm{SL}_2(\mathbb{C})$.

In Section 2 we prove Theorem (1.3). In Section 3 we study the fiber of the projection $t\colon\ R(\Gamma)\to X(\Gamma)$, introducing the different notions of irreducibility. Section 4 is devoted to the study of lifts of representations and the proof of Theorem (1.4). In the last section we determine the singular set of $X(\Gamma)$ when $\Gamma\cong F_n$ is the free group of rank $n\geq 3$.

2. Invariants of $\mathrm{PSL}_2(\mathbb{C})$

Before proving Theorem (1.3) we quickly review some basic notions of algebraic geometry and invariant theory (that the reader may prefer to skip and go directly to the proof in Subsection 2.3). For details see [KSS], [Kra] or [PV].

(2.1) Basic notions of invariant theory. A closed algebraic subset $Z\subset\mathbb{C}^N$ is called *affine*. We denote by $\mathbb{C}[Z]$ the ring of regular functions on Z. An algebraic group G that acts algebraically on Z acts naturally on $\mathbb{C}[Z]$ via $gf(z):=f(g^{-1}z)$. We denote by $\mathbb{C}[Z]^G$ the ring of invariant functions, i.e. functions $f\in\mathbb{C}[Z]$ for which $gf=f$ for all $g\in G$.

The group G is called *reductive* if it has the following property: for each finite dimensional rational representation $\rho\colon\ G\to\mathrm{GL}(V)$ and every G-invariant subspace $W \subset V$ there exist a complementary G-invariant subspace $W' \subset V$, i.e. $V = W' \oplus W$.

If Z is affine and G is reductive, then the ring $\mathbb{C}[Z]^G$ is finitely generated. The affine set Y such that $\mathbb{C}[Y] \cong \mathbb{C}[Z]^G$ is called the *algebraic quotient* and it is denoted by $Z//G$.

We shall use the following properties of reductive groups:

- By Maschke's theorem, finite groups are reductive.
- More generally, let $G \subset \mathrm{GL}_n(\mathbb{C})$ be a linear algebraic group. The group G is reductive if there is a Zariski-dense subgroup $K \subset G$ which is compact in the classical topology. It follows that $\mathrm{GL}_n(\mathbb{C})$, $\mathrm{SL}_n(\mathbb{C})$, $\mathrm{O}_n(\mathbb{C})$, $\mathrm{SO}_n(\mathbb{C})$ and $\mathrm{Sp}_n(\mathbb{C})$ are reductive.
- Let G be a reductive linear algebraic group. Let Y and Z be varieties on which G acts and let $f\colon X \to Y$ be a G-invariant regular map. If $f^*\colon \mathbb{C}[Y] \to \mathbb{C}[X]$ is surjective then $f^*(\mathbb{C}[Y]^G) = \mathbb{C}[X]^G$ holds.

(2.2) Algebraic structure of $\mathrm{PSL}_2(\mathbb{C})$**.** The group $\mathrm{PSL}_2(\mathbb{C})$ is algebraic, it is the quotient of $\mathrm{SL}_2(\mathbb{C})$ by the finite group $\{\pm \mathrm{Id}\}$.

It is useful to recall the isomorphism with $\mathrm{SO}_3(\mathbb{C})$, that we construct next. We denote by

$$\mathrm{Ad}\colon \mathrm{PSL}_2(\mathbb{C}) \to \mathrm{Aut}(\mathfrak{sl}_2(\mathbb{C}))$$

the adjoint action of $\mathrm{PSL}_2(\mathbb{C})$ on its Lie algebra $\mathfrak{sl}_2(\mathbb{C})$. The Killing form on $\mathfrak{sl}_2(\mathbb{C})$ is a non degenerate symmetric bilinear form over $\mathbb{C}$. For each $A \in \mathrm{PSL}_2(\mathbb{C})$, $\mathrm{Ad}(A)$ preserves the Killing form and $\det(\mathrm{Ad}(A)) = 1$, hence $\mathrm{Ad}(\mathrm{PSL}_2(\mathbb{C})) \subseteq \mathrm{SO}_3(\mathbb{C})$. The following lemma is well known from representation theory (see for instance [FH]):

Lemma (2.2.1). *The action of* $\mathrm{PSL}_2(\mathbb{C})$ *on the Lie algebra induces an isomorphism* $\mathrm{Ad}\colon \mathrm{PSL}_2(\mathbb{C}) \to \mathrm{SO}_3(\mathbb{C})$.

In this paper the trace will be abbreviated by tr, and $\mathrm{tr}^2(A)$ stands for $(\mathrm{tr}(A))^2$. By direct computation we obtain the equality

$$\mathrm{tr}(\mathrm{Ad}(A)) = \mathrm{tr}^2(A) - 1 = \mathrm{tr}(A^2) + 1 \quad \text{for all } A \in \mathrm{PSL}_2(\mathbb{C}) \tag{2.2.2}$$

that will be used later.

Given $\gamma \in \Gamma$, we have a well defined function

$$\begin{array}{rccc} \tau_\gamma\colon & R(\Gamma) & \to & \mathbb{C} \\ & \rho & \mapsto & \mathrm{tr}^2(\rho(\gamma)) \end{array}$$

Since it is invariant by conjugation, it induces a function

$$J_\gamma\colon X(\Gamma) \to \mathbb{C}.$$

(2.3) Proof of Theorem (1.3). Theorem (1.3) is a consequence of:

Proposition (2.3.1). *The ring of invariant functions* $\mathbb{C}[R(\Gamma)]^{\mathrm{PSL}_2(\mathbb{C})}$ *is generated by the functions* τ_γ*, with* $\gamma \in \Gamma$.

Proof. There is a surjection $\psi\colon F_n \to \Gamma$ where F_n is a free group of rank $n \in \mathbb{N}$. We obtain an inclusion $\psi^*\colon R(\Gamma) \subset R(F_n)$. This inclusion induces a surjection

$\psi_*\colon \mathbb{C}[R(F_n)] \to \mathbb{C}[R(\Gamma)]$. Now, $\mathrm{PSL}_2(\mathbb{C})$ is reductive and acts regularly by conjugation on the representation varieties. Hence we obtain a surjection

$$\psi_*\colon \mathbb{C}[R(F_n)]^{\mathrm{PSL}_2(\mathbb{C})} \to \mathbb{C}[R(\Gamma)]^{\mathrm{PSL}_2(\mathbb{C})}$$

and it is sufficient to prove the proposition for $\Gamma = F_n$ since $\psi_*(\tau_\gamma) = \tau_{\psi(\gamma)}$.

Using Lemma (2.2.1) and (2.2.2), we have to prove that $\mathbb{C}[R(F_n)]^{\mathrm{SO}_3(\mathbb{C})}$ is generated by the trace functions on elements of F_n. Equivalently, we claim that

$$\mathbb{C}[\mathrm{SO}_3(\mathbb{C}) \times \cdots \times \mathrm{SO}_3(\mathbb{C})]^{\mathrm{SO}_3(\mathbb{C})}$$

is generated by traces of products of matrices and their transposes.

Let $\mathrm{M}_3(\mathbb{C})$ denote the algebra of 3×3 matrices with complex coefficients. The group $\mathrm{PSL}_2(\mathbb{C}) \cong \mathrm{SO}_3(\mathbb{C})$ acts on the product $\mathrm{M}_3(\mathbb{C}) \times \cdots \times \mathrm{M}_3(\mathbb{C})$ diagonally by conjugation. A theorem of Aslaksen, Tan and Zhu (see [ATZ]) states that the algebra of invariant functions

$$\mathbb{C}[\mathrm{M}_3(\mathbb{C}) \times \cdots \times \mathrm{M}_3(\mathbb{C})]^{\mathrm{SO}_3(\mathbb{C})}$$

is generated by the traces of products of matrices and their transposes. Thus the proof of the proposition reduces to show that we have a natural surjection

$$\mathbb{C}[\mathrm{M}_3(\mathbb{C}) \times \cdots \times \mathrm{M}_3(\mathbb{C})]^{\mathrm{SO}_3(\mathbb{C})} \to \mathbb{C}[\mathrm{SO}_3(\mathbb{C}) \times \cdots \times \mathrm{SO}_3(\mathbb{C})]^{\mathrm{SO}_3(\mathbb{C})}\,.$$

Since $\mathrm{SO}_3(\mathbb{C}) \times \cdots \times \mathrm{SO}_3(\mathbb{C}) \subset \mathrm{M}_3(\mathbb{C}) \times \cdots \times \mathrm{M}_3(\mathbb{C})$ is a closed subvariety we obtain a natural surjection

$$\mathbb{C}[\mathrm{M}_3(\mathbb{C}) \times \cdots \times \mathrm{M}_3(\mathbb{C})] \to \mathbb{C}[\mathrm{SO}_3(\mathbb{C}) \times \cdots \times \mathrm{SO}_3(\mathbb{C})]$$

which is of course $\mathrm{SO}_3(\mathbb{C})$-invariant. Using the fact that $\mathrm{SO}_3(\mathbb{C})$ is reductive gives the surjection $\mathbb{C}[\mathrm{M}_3(\mathbb{C}) \times \cdots \times \mathrm{M}_3(\mathbb{C})]^{\mathrm{SO}_3(\mathbb{C})} \to \mathbb{C}[\mathrm{SO}_3(\mathbb{C}) \times \cdots \times \mathrm{SO}_3(\mathbb{C})]^{\mathrm{SO}_3(\mathbb{C})}$. □

Since $\mathbb{C}[X(\Gamma)] = \mathbb{C}[R(\Gamma)]^{\mathrm{SO}_3(\mathbb{C})}$ is finitely generated, we also obtain:

Corollary (2.3.2). *There are finitely many elements $\gamma_1, \dots, \gamma_N$ in Γ such that $J_{\gamma_1} \times \cdots \times J_{\gamma_N}\colon X(M) \to \mathbb{C}^N$ is an embedding and its image is a closed algebraic set.*

(2.4) Other invariant functions. There are other natural functions to consider. Let Γ^2 be the subgroup of Γ generated by the squares γ^2 of all elements γ of Γ. It is well known that we have an exact sequence:

$$1 \to \Gamma^2 \to \Gamma \to H_1(\Gamma, C_2) \to 1,$$

where $C_2 = \{\pm 1\}$ is the group with 2 elements. For instance, if Γ is a finite group of odd order, then $\Gamma^2 = \Gamma$. In general, if $\gamma, \mu \in \Gamma$ the commutator $[\gamma,\mu] = \gamma\mu\gamma^{-1}\mu^{-1} = (\gamma\mu)^2(\mu^{-1}\gamma^{-1}\mu)^2\mu^{-2}$ is in Γ^2 and hence Γ^2 contains the commutator group $\Gamma' = [\Gamma,\Gamma]$. Notice that

$$\Gamma^2 = \bigcap_{\epsilon \in H^1(\Gamma, C_2)} \mathrm{Ker}(\epsilon)$$

where $H^1(\Gamma, C_2) = \mathrm{Hom}(\Gamma, C_2)$. Let $R(\Gamma, \mathrm{SL}_2(\mathbb{C}))$ denote the variety of representations of Γ in $\mathrm{SL}_2(\mathbb{C})$. The cohomology group $H^1(\Gamma, C_2)$ acts on this variety of representations as follows: an homomorphism $\epsilon\colon \Gamma \to C_2 = \{\pm 1\}$ maps the representation $\rho \in R(\Gamma, \mathrm{SL}_2(\mathbb{C}))$ to the product of representations $\epsilon \cdot \rho$ (which maps $\gamma \in \Gamma$ to $\epsilon(\gamma) \cdot \rho(\gamma)$).

Invariant functions for the free group. Let F be a finitely generated free group. For $\gamma \in F^2$ and $\rho \in R(F)$, $\operatorname{tr}(\rho(\gamma))$ is well defined since the representation $\rho\colon F \to \mathrm{PSL}_2(\mathbb{C})$ lifts to $\tilde\rho\colon F \to \mathrm{SL}_2(\mathbb{C})$ and for $\gamma \in F^2$ the trace $\operatorname{tr}(\tilde\rho(\gamma))$ depends only on γ. Note that two lifts $\tilde\rho_1$ and $\tilde\rho_2$ of ρ differ by a homomorphism $\epsilon \in H^1(F, C_2)$ and that $F^2 \subset \operatorname{Ker}(\epsilon)$ for each $\epsilon \in H^1(F, C_2)$.

PROPOSITION (2.4.1). *Let F be a free group. For every k-tuple $\gamma_1, \dots, \gamma_k \in F$ such that the product $\gamma_1 \cdots \gamma_k \in F^2$, the function*

$$\begin{array}{rrcl} \sigma_{\gamma_1,\dots,\gamma_k}\colon & R(F) & \to & \mathbb{C} \\ & \rho & \mapsto & \operatorname{tr}(\tilde\rho(\gamma_1)) \cdots \operatorname{tr}(\tilde\rho(\gamma_k)) \end{array}$$

is regular (i.e. $\sigma_{\gamma_1,\dots,\gamma_k} \in \mathbb{C}[R(F)]$). Here, $\tilde\rho\colon F \to \mathrm{SL}_2(\mathbb{C})$ denotes a lift of ρ.

In order to prove this proposition we shall use the following:

LEMMA (2.4.2). *Let F_n be the free group of rank n. We have a natural isomorphism*

$$R(F_n, \mathrm{SL}_2(\mathbb{C}))//H^1(F_n, C_2) \cong R(F_n)\,.$$

Proof. Since $R(F_n, \mathrm{SL}_2(\mathbb{C})) \cong \mathrm{SL}_2(\mathbb{C})^n$, $R(F_n) \cong \mathrm{PSL}_2(\mathbb{C})^n$ and $\mathrm{SL}_2(\mathbb{C})/C_2 \cong \mathrm{PSL}_2(\mathbb{C})$, we have the lemma. $\square$

Proof of Proposition (2.4.1). For a free group F and $\gamma_1, \dots, \gamma_k \in F$, the function $\tilde\sigma\colon R(F, \mathrm{SL}_2(\mathbb{C})) \to \mathbb{C}$ given by $\tilde\sigma(\rho) = \operatorname{tr}(\rho(\gamma_1)) \cdots \operatorname{tr}(\rho(\gamma_k))$ is regular. Moreover, we have $\tilde\sigma(\epsilon \cdot \rho) = \epsilon(\gamma_1 \cdots \gamma_k)\tilde\sigma(\rho)$. Since the product $\gamma_1 \cdots \gamma_k \in F^2$ we get that $\tilde\sigma \in \mathbb{C}[R(F_n, \mathrm{SL}_2(\mathbb{C}))]^{H^1(F_n, C_2)}$ is an invariant regular function on the $\mathrm{SL}_2(\mathbb{C})$ representation variety. By Lemma (2.4.2), this function factors through $R(F)$ and gives the regular function $\sigma_{\gamma_1,\dots,\gamma_k} \in \mathbb{C}[R(F)]$. $\square$

Example (2.4.3). Given $\gamma, \eta \in F$, by Proposition (2.4.1), $\sigma_{\gamma,\eta,\gamma\eta} \in \mathbb{C}[R(F)]$, thus by Proposition (2.3.1), $\sigma_{\gamma,\eta,\gamma\eta}$ is a polynomial on the functions τ.

To compute explicitly the polynomial of Example (2.4.3), we recall some identities of traces in $\mathrm{SL}_2(\mathbb{C})$:

$$\operatorname{tr}(AB) = \operatorname{tr}(BA) \quad \text{and} \quad \operatorname{tr}(A) = \operatorname{tr}(A^{-1}) \qquad \forall\, A, B \in \mathrm{SL}_2(\mathbb{C})\,.$$

In addition, we have the fundamental identity:

$$\operatorname{tr}(AB) + \operatorname{tr}(A^{-1}B) = \operatorname{tr}(A)\ \operatorname{tr}(B) \qquad \forall\, A, B \in \mathrm{SL}_2(\mathbb{C})\,. \tag{2.4.4}$$

This identity can be deduced from $A^2 - (\operatorname{tr} A)A + \mathrm{Id} = 0$ multiplying by $A^{-1}B$ and taking traces. Taking the square of $\operatorname{tr}(AB^{-1}) = \operatorname{tr}(A)\ \operatorname{tr}(B) - \operatorname{tr}(AB)$ we deduce:

$$2\operatorname{tr}(A)\operatorname{tr}(B)\operatorname{tr}(AB) = \operatorname{tr}^2(A)\ \operatorname{tr}^2(B) + \operatorname{tr}^2(AB) - \operatorname{tr}^2(AB^{-1})\,.$$

Thus

$$\sigma_{\gamma,\eta,\gamma\eta} = \frac{1}{2}(\tau_\gamma\tau_\eta + \tau_{\gamma\eta} - \tau_{\gamma\eta^{-1}}). \tag{2.4.5}$$

Example (2.4.6). For every $\gamma, \mu \in F$, the commutator $[\gamma, \mu] = \gamma\mu\gamma^{-1}\mu^{-1}$ belongs to F^2 and therefore $\sigma_{[\gamma,\mu]} \in \mathbb{C}[R(F)]$. Using the the same method as for Equation (2.4.5) one can find:

$$\sigma_{[\gamma,\eta]} = \tau_\gamma + \tau_\eta + \frac{1}{2}\tau_{\gamma\eta} + \frac{1}{2}\tau_{\gamma\eta^{-1}} - \frac{1}{2}\tau_\gamma\tau_\eta - 2\,. \tag{2.4.7}$$

Invariant functions for other groups. Let Γ be a finitely generated group, F a free group and $\psi\colon F \to \Gamma$ a surjection. It induces another surjection $\psi_*\colon \mathbb{C}[R(F)] \to \mathbb{C}[R(\Gamma)]$, $\psi_* f(\rho) = f(\rho\circ\psi)$. Hence we obtain for all $\eta_1,\dots,\eta_k \in F$ such that the product $\eta_1\cdots\eta_k \in F^2$ a regular function $\psi_*\sigma_{\eta_1,\dots,\eta_k} \in \mathbb{C}[R(\Gamma)]$. Note that the functions $\psi_*\sigma_{\eta_1}$ and $\psi_*\sigma_{\eta_2}$ might be different even if $\psi(\eta_1)=\psi(\eta_2)$ in Γ. This reflects the fact that in general not every representation $\rho\colon \Gamma \to \mathrm{PSL}_2(\mathbb{C})$ lifts to $\mathrm{SL}_2(\mathbb{C})$.

Example (2.4.8). Let $\psi\colon F \to \Gamma$ be the canonical projection where $F = \langle x, y \mid -\rangle$ and $\Gamma = \langle x, y \mid [x,y]=1\rangle$. We consider the representation $\rho\colon \Gamma \to \mathrm{PSL}_2(\mathbb{C})$ given by $\rho(x) = \pm A_x$ and $\rho(y) = \pm A_y$ where

$$A_x = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \text{ and } A_y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

We obtain $\mathrm{tr}([A_x,A_y]) = -2$ and hence $\psi_*\sigma_{[x,y]}(\rho) = -2$. On the other hand we have $[x,y]=1$ in Γ and $\psi_*\sigma_1 = 2$ is a constant function.

If the representation $\rho \in R(\Gamma)$ admits a lift $\tilde\rho\colon \Gamma \to \mathrm{SL}_2(\mathbb{C})$ then

$$\psi_*\sigma_{\eta_1,\dots,\eta_k}(\rho) = \mathrm{tr}(\tilde\rho(\psi(\eta_1)))\cdots\mathrm{tr}(\tilde\rho(\psi(\eta_k))) \tag{2.4.9}$$

only depends on the elements $\psi(\eta_1),\dots,\psi(\eta_k)\in\Gamma$.

3. Irreducibility

To study the fiber of the map $t\colon R(\Gamma) \to X(\Gamma)$ we shall consider two different notions of irreducibility for $\rho \in R(\Gamma)$, the usual one as a representation in $\mathrm{PSL}_2(\mathbb{C})$ and the so called Ad-irreducibility, for the three dimensional representation $\mathrm{Ad}\circ\rho\colon \Gamma \to SO_3(\mathbb{C})$.

(3.1) Irreducible representations.

Definition (3.1.1). *A representation $\rho \in R(\Gamma)$ is called reducible if $\rho(\Gamma)$ preserves a point of $P^1(\mathbb{C})$. Otherwise it is called irreducible. A character $\chi\colon \Gamma \to \mathbb{C}$ is called reducible if it is the character of a reducible representation.*

Remark (3.1.2). Up to conjugation, the image of a reducible representation is contained in the set of upper-triangular matrices $\left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right)$.

We shall require the following well known lemma (see [Bea, § 4.3]).

LEMMA (3.1.3). *Two non-trivial elements $g, h \in \mathrm{PSL}_2(\mathbb{C})$ have a common fixed point in $P^1(\mathbb{C})$ if and only if $\mathrm{tr}([g,h]) = 2$. In addition, this fixed point is unique if $[g,h]$ is not the identity.*

Irreducibility is a property that can be detected from characters:

LEMMA (3.1.4). *A representation $\rho \in R(\Gamma)$ is reducible iff $\mathrm{tr}([\rho(\gamma),\rho(\eta)]) = 2$ for all elements γ,η in Γ.*

Proof. If ρ is reducible then all the $\rho(\gamma)$ have a common fixed point and Lemma (3.1.3) gives the result.

Assume now that $\mathrm{tr}([\rho(\gamma),\rho(\eta)]) = 2$ for all elements γ,η in Γ.

Case 1: There are two elements γ and η in Γ such that $[\rho(\gamma),\rho(\eta)]$ is not the identity. Then $A = [\rho(\gamma),\rho(\eta)]$ is a non-trivial parabolic element in the image of

Γ. For any $\mu \in \Gamma$, either $\rho(\mu)$ commutes with A or $[\rho(\mu), A]$ is non-trivial. The former possibility implies that $\rho(\mu)$ fixes the unique fixed point of A, the latter too by Lemma (3.1.3).

Case 2: The image of ρ is an abelian group. Abelian subgroups of $\mathrm{PSL}_2(\mathbb{C})$ are well-known: either they have a global fixed point in $P^1(\mathbb{C})$ or they are conjugated to the group with four elements generated by $\pm\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $\pm\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$. Since the commutator of these two generators is $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, this possibility does not occur. $\square$

Definition (3.1.5). *A non-cyclic abelian subgroup of* $\mathrm{PSL}_2(\mathbb{C})$ *with four elements is called Klein's 4-group. Such a group is realized by rotations about three orthogonal geodesics and it is conjugated to the one generated by* $\pm\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ *and* $\pm\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$.

Let $R^{red}(\Gamma)$ denote the set of reducible representations and $X^{red}(\Gamma) = t(R^{red}(\Gamma))$. Let F be a free group and let $\psi\colon F \to \Gamma$ be surjective. Lemma (3.1.4) implies that

$$R^{red}(\Gamma) = \{\rho \in R(\Gamma) \mid \psi_* \sigma_{[\gamma,\eta]}(\rho) = 2 \quad \forall\, \gamma, \eta \in F\}$$

is a Zariski closed subset invariant by conjugation. Thus, by invariant theory we have:

COROLLARY (3.1.6). *The set* $X^{red}(\Gamma)$ *is Zariski closed and* $R^{red}(\Gamma) = t^{-1}(X^{red}(\Gamma))$.

Remark (3.1.7). Every reducible character χ is the character of a diagonal representation, because if $\rho(\gamma) = \pm\begin{pmatrix} a_\gamma & b_\gamma \\ 0 & c_\gamma \end{pmatrix}$ is a representation, then $\rho'(\gamma) = \pm\begin{pmatrix} a_\gamma & 0 \\ 0 & c_\gamma \end{pmatrix}$ is also a representation with $\chi_\rho = \chi_{\rho'}$.

(3.2) Ad-irreducibility.

Definition (3.2.1). *A representation* $\rho \in R(\Gamma)$ *is* Ad*-reducible if* $\mathfrak{sl}_2(\mathbb{C})$ *has a proper invariant subspace by the action of* $\mathrm{Ad} \circ \rho$. *Otherwise it is* Ad*-irreducible.*

Let $\mathbb{H}^3$ denote the three-dimensional hyperbolic space and $\partial_\infty \mathbb{H}^3$ its ideal boundary. We use the isomorphism $\mathrm{Isom}^+(\mathbb{H}^3) \cong \mathrm{PSL}_2(\mathbb{C})$ and the natural identification $\partial_\infty \mathbb{H}^3 \cong P^1(\mathbb{C})$.

LEMMA (3.2.2). *A representation* $\rho\colon \Gamma \to \mathrm{PSL}_2(\mathbb{C})$ *is* Ad*-reducible if and only if* $\rho(\Gamma)$ *preserves either a point in* $\partial_\infty \mathbb{H}^3$ *or a geodesic in* $\mathbb{H}^3$.

Proof. Let V be a proper subspace of $\mathfrak{sl}_2(\mathbb{C})$ invariant by $\mathrm{Ad} \circ \rho(\Gamma)$. Up to taking $V^\perp$, we may assume $\dim V = 1$, because the Killing form is not degenerate. We have then two possibilities: either the Killing form restricted to V vanishes or not. In the first case V consists of parabolic Killing fields, in particular the 1-parameter group $\{\exp(v) \mid v \in V\} \cong \mathbb{C}$ is parabolic and fixes a unique point at infinity, that has to be fixed also by ρ. In the second case, when the Killing form restricted to V does not vanish, the 1-parameter group $\{\exp(v) \mid v \in V\} \cong \mathbb{C}^*$ is a subgroup of index two in the group of isometries which preserve a geodesic in $\mathbb{H}^3$. This geodesic has to be preserved by the representation. Conversely, if a representation preserves a point in $\partial_\infty \mathbb{H}^3$ or a geodesic, the previous argument shows how to construct an invariant subspace of $\mathfrak{sl}_2(\mathbb{C})$. $\square$

Corollary (3.2.3). *Reducible representations are also* Ad*-reducible.*

Remark (3.2.4). A representation Ad-reducible but not reducible is a C_2-extension of an abelian one that fixes an oriented geodesic. Thus it preserves an unoriented geodesic.

We call a representation $\rho \in R(\Gamma)$ *abelian* respectively *metabelian* if its image is an abelian respectively metabelian subgroup of $\mathrm{PSL}_2(\mathbb{C})$

Lemma (3.2.5). *A representation $\rho \in R(\Gamma)$ is* Ad*-reducible iff it is metabelian.*

Proof. If ρ is Ad-reducible then its image is contained in the stabilizer of either a point in $P^1(\mathbb{C})$ or a geodesic in $\mathbb{H}^3$. Those stabilizers are metabelian, since they are respectively the group of affine transformations of $\mathbb{C}$ and the semidirect product $\mathbb{C}^* \rtimes C_2$.

Now assume that $\rho(\Gamma) \subset \mathrm{PSL}_2(\mathbb{C})$ is a metabelian subgroup. We use the fact that an abelian subgroup of $\mathrm{PSL}_2(\mathbb{C})$ preserves a unique point of $P^1(\mathbb{C})$, a unique geodesic or it is Klein's 4-group (Definition (3.1.5)). If $\rho([\Gamma,\Gamma])$ is trivial then ρ is Ad-reducible by this fact. If $\rho([\Gamma,\Gamma])$is not trivial, then we look at those *unique* invariant objects: the *unique* point in $P^1(\mathbb{C})$, the *unique* geodesic, or the *unique* three geodesics if it is Klein's 4-group. Since $[\Gamma,\Gamma]$ is normal in Γ, uniqueness implies that $\rho(\Gamma)$ preserves the same objects, hence ρ is Ad-reducible. $\square$

Lemma (3.2.6). *The set of characters of* Ad*-reducible representations is Zariski closed.*

Proof. Lemma (3.2.5) gives that the set of Ad-reducible representations is

$$R^{\mathrm{Ad}-red} = \{\rho \in R(\Gamma) \mid \rho(c) = \pm\mathrm{Id} \quad \forall\, c \in \Gamma''\}$$

where Γ'' denotes the second commutator group of Γ. This is a closed subset of $R(\Gamma)$ invariant under conjugation. Hence we have $X^{\mathrm{Ad}-red}(\Gamma) = t(R^{\mathrm{Ad}-red})$ is a closed subset of $X(\Gamma)$. $\square$

Remark (3.2.7). The image of an Ad-reducible representation is elementary, but elementary groups also include groups that fix a point in $\mathbb{H}^3$.

(3.3) The fibers of $t\colon R(\Gamma) \to X(\Gamma)$.

Lemma (3.3.1). *The fiber of an irreducible character consists of a single closed orbit (i.e. two irreducible representations have the same character iff they are conjugate).*

Proof. Let $\rho_1, \rho_2 \in R(\Gamma)$ be two irreducible representations with $\chi_{\rho_1} = \chi_{\rho_2}$.

We assume first that each ρ_i is irreducible but Ad-reducible. Thus each ρ_i preserves a geodesic l, that we may assume to be the same after conjugation. The action of $\rho_i(\gamma)$ on l is determined by the value of $\chi_{\rho_i}(\gamma)$, except in the case $\chi_{\rho_i}(\gamma) = 0$, which means that $\rho_i(\gamma)$ is a rotation through angle π, but it can be either about γ or about an axis perpendicular to γ. Thus if there exists an element $\gamma_0 \in \Gamma$ with $\chi_{\rho_i}(\gamma_0) \neq 4, 0$ (i.e. $\rho_i(\gamma_0)$ is either a loxodromic element or a rotation of angle $\neq \pi$) then $\forall\, \gamma \in \Gamma$ the action of $\rho_i(\gamma)$ on the geodesic l is determined by $\chi_{\rho_i}(\gamma)$ and $\chi_{\rho_i}(\gamma\gamma_0)$. In particular ρ_i is unique up to conjugation.

The exceptional case occurs when $\chi_{\rho_i}(\gamma) = 0$ or 4 for every $\gamma \in \Gamma$. In this special case, ρ_i is necessarily a representation into Klein's 4-group. The lemma is also clear in this case.

When ρ_i are Ad-irreducible, we can assume that Γ is a free group. Thus we can lift ρ_i to $\tilde{\rho}_i\colon \Gamma \to \mathrm{SL}_2(\mathbb{C})$. By Example (2.4.3), for every pair $\gamma, \gamma' \in \Gamma$ we obtain a regular function $\sigma_{\gamma,\gamma',\gamma\gamma'}\colon X(\Gamma) \to \mathbb{C}$, given by

$$\sigma_{\gamma,\gamma',\gamma\gamma'}(\chi_\rho) = \operatorname{tr}\tilde{\rho}(\gamma\gamma')\ \operatorname{tr}\tilde{\rho}(\gamma)\ \operatorname{tr}\tilde{\rho}(\gamma')$$

where $\tilde{\rho}\colon \Gamma \to \mathrm{SL}_2(\mathbb{C})$ is any lift of ρ. Thus:

$$\operatorname{tr}\tilde{\rho}_1(\gamma\gamma')\ \operatorname{tr}\tilde{\rho}_1(\gamma)\ \operatorname{tr}\tilde{\rho}_1(\gamma') = \operatorname{tr}\tilde{\rho}_2(\gamma\gamma')\ \operatorname{tr}\tilde{\rho}_2(\gamma)\ \operatorname{tr}\tilde{\rho}_2(\gamma')\,. \tag{3.3.2}$$

We define $\epsilon\colon \Gamma \to C_2 = \{\pm 1\}$ by the formula:

$$\operatorname{tr}\tilde{\rho}_1(\gamma) = \epsilon(\gamma)\operatorname{tr}\tilde{\rho}_2(\gamma), \qquad \forall\, \gamma \in \Gamma \text{ such that } \chi_{\rho_1}(\gamma) \neq 0.$$

When $\chi_{\rho_1}(\gamma) = 0$, since we assume that ρ_i is Ad-irreducible, we can find $\gamma_0 \in \Gamma$ with $\chi_{\rho_i}(\gamma_0) \neq 0$ and $\chi_{\rho_i}(\gamma\gamma_0) \neq 0$. In this case we define $\epsilon(\gamma) = \epsilon(\gamma_0) \cdot \epsilon(\gamma\gamma_0)$.

By (3.3.2), ϵ is a morphism. Hence $\tilde{\rho}_1$ and $\epsilon \cdot \tilde{\rho}_2$ are irreducible representations in $\mathrm{SL}_2(\mathbb{C})$ with the same character. By [CS] they are conjugate. ☐

Proposition (3.3.3). (i) *A character χ is irreducible iff* $\mathrm{PSL}_2(\mathbb{C})$ *acts transitively on the fiber and with finite stabilizer.*

(ii) *A character is* Ad-*irreducible iff* $\mathrm{PSL}_2(\mathbb{C})$ *acts faithfully on the fiber.*

Proof. (i) By Lemma (3.3.1), if χ is irreducible then $\mathrm{PSL}_2(\mathbb{C})$ acts transitively on $t^{-1}(\chi)$. Assume now that the stabilizer is infinite: i.e. there exists nontrivial $A \in \mathrm{PSL}_2(\mathbb{C})$ of order ≥ 3 (possibly infinite) and ρ in the fiber such that A commutes with ρ. If A is parabolic, then it has a fixed point in $P^1(\mathbb{C})$ and therefore ρ fixes this point. Otherwise A has an invariant geodesic; since A has order ≥ 3, ρ preserves the oriented geodesic, and therefore ρ is also reducible.

Assume the character is reducible, then it has a diagonal representation ρ on the fiber (Remark (3.1.7)), and therefore the group of diagonal matrices stabilizes it. Thus the stabilizer is infinite.

(ii) Assume $\mathrm{PSL}_2(\mathbb{C})$ does not act faithfully on the fiber, i.e. there exists nontrivial $A \in \mathrm{PSL}_2(\mathbb{C})$ and ρ in the fiber such that A commutes with ρ. If A is parabolic, then ρ fixes a point in $P^1(\mathbb{C})$ by the previous argument. Otherwise A has an invariant geodesic, and by commutativity, ρ must preserve this geodesic. In both cases, ρ is Ad-reducible.

If the character is irreducible but Ad-reducible, then it preserves a geodesic, and the rotation through angle π about this geodesic commutes with ρ. Hence the stabilizer is nontrivial. ☐

Remark (3.3.4). The projection $t\colon\ R(\Gamma) \to X(\Gamma)$ induces a bijection between irreducible components.

A priori $R(\Gamma)$ could have more components than $X(\Gamma)$, but the number of components is the same, because $\mathrm{PSL}_2(\mathbb{C})$ is irreducible.

From Corollary (3.1.6) and Proposition (3.3.3) we deduce:

Corollary (3.3.5). *Let $\rho \in R(\Gamma)$ be an irreducible representation. Let R_0 denote an irreducible component of $R(\Gamma)$ that contains ρ and let X_0 denote the corresponding irreducible component of $X(\Gamma)$. Then*

$$\dim R_0 = \dim X_0 + 3.$$

4. Lifts of representations to $\mathrm{SL}_2(\mathbb{C})$

Let $\overline{R}(\Gamma) \subset R(\Gamma)$ denote the set of representations $\rho \in R(\Gamma)$ that lift to $\mathrm{SL}_2(\mathbb{C})$. According to [Cul, Thm. 4.1] $\overline{R}(\Gamma)$ is a union of connected components of $R(\Gamma)$. In particular $\overline{R}(\Gamma)$ is a Zariski-closed algebraic subset of $R(\Gamma)$, since irreducible complex varieties are connected in the $\mathbb{C}$-topology [Sha, VII, §2]. Moreover, $\overline{R}(\Gamma)$ is invariant under conjugation and hence the algebraic quotient

$$\overline{X}(\Gamma) = \overline{R}(\Gamma)//\mathrm{PSL}_2(\mathbb{C})$$

is a well defined closed subset of $X(\Gamma)$.

In many cases, $\overline{X}(\Gamma) = X(\Gamma)$. For instance this is clear when Γ is a free group. It is also true if $H^2(\Gamma, C_2) = 0$ by the following remark (see [GM] or [Cul]).

Remark (4.1). Let $\rho\colon \Gamma \to \mathrm{PSL}_2(\mathbb{C})$ be a representation. There is a second Stiefel-Whitney class $w_2(\rho) \in H^2(\Gamma, C_2)$ which is exactly the obstruction for the existence of a lift $\overline{\rho}\colon \Gamma \to \mathrm{SL}_2(\mathbb{C})$.

(4.2) Properties of $\overline{X}(\Gamma)$. Let $R(\Gamma, \mathrm{SL}_2(\mathbb{C}))$ and $X(\Gamma, \mathrm{SL}_2(\mathbb{C}))$ denote the variety of representations and characters in $\mathrm{SL}_2(\mathbb{C})$. The ring $\mathbb{C}[R(\Gamma, \mathrm{SL}_2(\mathbb{C}))]^{\mathrm{SL}_2(\mathbb{C})}$ is generated by the trace functions $\tilde{\tau}_\gamma\colon R(\Gamma, \mathrm{SL}_2(\mathbb{C})) \to \mathbb{C}$, $\tilde{\tau}_\gamma(\rho) = \mathrm{tr}(\rho(\gamma))$. The function induced by $\tilde{\tau}_\gamma$ is denoted by $I_\gamma\colon X(\Gamma) \to \mathbb{C}$, therefore $\mathbb{C}[X(\Gamma)]$ is finitely generated by the functions I_γ, $\gamma \in \Gamma$ [CS].

Elements of the cohomology group $H^1(\Gamma, C_2)$ are homomorphisms $\theta\colon \Gamma \to C_2 = \{\pm 1\}$ that act on representations by multiplication. The action of $\epsilon \in H^1(\Gamma, C_2)$ on I_γ is given by: $\epsilon \cdot I_\gamma = \epsilon(\gamma) I_\gamma$. Since $H^1(\Gamma, C_2)$ is finite, it is reductive and we may take the quotient of invariant theory.

Let F be a finitely generated free group and $\psi\colon F \to \Gamma$ be a surjection. We fix a k-tuple $\gamma_1, \ldots, \gamma_k \in \Gamma$ such that the product $\gamma_1 \cdots \gamma_k \in \Gamma^2$. Moreover, we choose $\eta_i \in F$ such that $\psi(\eta_i) = \gamma_i$ and such that the product $\eta_1 \cdots \eta_k \in F^2$. The function $\psi^* \sigma_{\eta_1, \ldots, \eta_k} \in \mathbb{C}[\overline{R}(\Gamma)]$ is invariant under conjugation and gives us a function $\psi^* \sigma_{\eta_1, \ldots, \eta_k} \in \mathbb{C}[\overline{X}(\Gamma)]$. By Equation (2.4.9) we have $\psi^* \sigma_{\eta_1, \ldots, \eta_k}(\chi) = \tilde{\chi}(\gamma_1) \cdots \tilde{\chi}(\gamma_k)$ where $\tilde{\chi} \in X(\Gamma, \mathrm{SL}_2(\mathbb{C}))$ is a character such that $\pi(\tilde{\chi}) = \chi$. Note that $\pi\colon X(\Gamma, \mathrm{SL}_2(\mathbb{C})) \to \overline{X}(\Gamma)$ is surjective. The function

$$\Sigma_{\gamma_1, \ldots, \gamma_k} := \phi^* \sigma_{\eta_1, \ldots, \eta_k} \in \mathbb{C}[\overline{X}(\Gamma)] \tag{4.2.1}$$

depends only on the elements $\gamma_i \in \Gamma$.

Proposition (4.2.2). *There is a natural isomorphism:*

$$X(\Gamma, \mathrm{SL}_2(\mathbb{C}))//H^1(\Gamma, C_2) \cong \overline{X}(\Gamma).$$

Proof. Composition with the projection $\mathrm{SL}_2(\mathbb{C}) \to \mathrm{PSL}_2(\mathbb{C})$ induces a surjection

$$\pi\colon X(\Gamma, \mathrm{SL}_2(\mathbb{C})) \to \overline{X}(\Gamma),$$

which is easily seen to be algebraic and is given by $\pi(\chi) = \chi^2$. At the level of function rings it induces an injection

$$\pi^*\colon\ \mathbb{C}[\overline{X}(\Gamma)] \hookrightarrow \mathbb{C}[X(\Gamma, \mathrm{SL}_2(\mathbb{C}))].$$

We have $\pi^* f(\chi) = f(\chi^2)$ for $f \in \mathbb{C}[\overline{X}(\Gamma)]$ and $\chi \in X(\Gamma, \mathrm{SL}_2(\mathbb{C}))$. The image of π^* is contained in the set of invariant functions:

$$\operatorname{Im} \pi^* \subseteq \mathbb{C}[X(\Gamma, \mathrm{SL}_2(\mathbb{C}))]^{H^1(\Gamma, C_2)}.$$

More precisely, we have $\pi^* f(\epsilon\chi) = f(\epsilon^2\chi^2) = \pi^* f(\chi)$ for all $\epsilon \in H^1(\Gamma, C_2)$. It remains to prove that this inclusion is an equality.

Since $\mathbb{C}[X(\Gamma, \mathrm{SL}_2(\mathbb{C}))]$ is generated as $\mathbb{C}$-algebra by the functions I_γ with $\gamma \in \Gamma$, the monomials

$$I_{\gamma_1} I_{\gamma_2} \cdots I_{\gamma_k}$$

generate $\mathbb{C}[X(\Gamma, \mathrm{SL}_2(\mathbb{C}))]$ as a $\mathbb{C}$-vector space. Taking the average of the action of $H^1(\Gamma, C_2)$, we deduce that the subspace of invariant functions $\mathbb{C}[X(\Gamma, \mathrm{SL}_2(\mathbb{C}))]^{H^1(\Gamma, C_2)}$ is generated by

$$\frac{1}{2^r} \sum_{\epsilon \in H^1(\Gamma, C_2)} \epsilon \cdot I_{\gamma_1} \cdots I_{\gamma_k} = \Big(\frac{1}{2^r} \sum_{\epsilon \in H^1(\Gamma, C_2)} \epsilon(\gamma_1 \cdots \gamma_k)\Big) I_{\gamma_1} \cdots I_{\gamma_k}$$

where r is the rank of $H^1(\Gamma, C_2)$ (see [Kra, II.3.6] for instance). Using the fact that

$$\frac{1}{2^r} \sum_{\epsilon \in H^1(\Gamma, C_2)} \epsilon(\gamma) = \begin{cases} 1 & \text{if } \gamma \in \Gamma^2 \\ 0 & \text{otherwise} \end{cases}$$

we deduce that $\mathbb{C}[X(\Gamma, \mathrm{SL}_2(\mathbb{C}))]^{H^1(\Gamma, C_2)}$ is generated by the monomials $I_{\gamma_1} \cdots I_{\gamma_k}$ such that the product $\gamma_1 \ldots \gamma_k \in \Gamma^2$.

On the other hand we have for $\chi \in X(\Gamma, \mathrm{SL}_2(\mathbb{C}))$:

$$\pi^* \Sigma_{\gamma_1, \ldots, \gamma_k}(\chi) = \Sigma_{\gamma_1, \ldots, \gamma_k}(\chi^2) = \chi(\gamma_1) \cdots \chi(\gamma_k) = I_{\gamma_1} \cdots I_{\gamma_k}(\chi)\,,$$

where $\Sigma_{\gamma_1, \ldots, \gamma_k}$ is the function defined in (4.2.1). This gives that the monomials $I_{\gamma_1} \cdots I_{\gamma_k}$ such that the product $\gamma_1 \ldots \gamma_k \in \Gamma^2$ is in the image of π^* and therefore $\mathbb{C}[X(\Gamma, \mathrm{SL}_2(\mathbb{C}))]^{H^1(\Gamma, C_2)} = \operatorname{Im} \pi^*$. □

Remark (4.2.3). Let $p\colon\ X(\Gamma, \mathrm{SL}_2(\mathbb{C})) \to \overline{X}(\Gamma)$ denote the projection. If $\chi \in \overline{X}(\Gamma)$ is Ad-irreducible, then $p^{-1}(\chi)$ has 2^r points where r is the rank of $H^1(\Gamma, C_2)$. If χ is Ad-reducible then the cardinality of $p^{-1}(\chi)$ is strictly less than 2^r. Thus p is a branched covering with branching locus the set of Ad-reducible characters.

Example (4.2.4). Let F_2 be the free group of rank 2, with generators α and β. There is an isomorphism:

$$(I_\alpha, I_\beta, I_{\alpha\beta})\colon\ X(F_2, \mathrm{SL}_2(\mathbb{C})) \to \mathbb{C}^3$$

where I_γ denotes the regular function induced by $\tilde{\tau}_\gamma$. In particular $X(F_2, \mathrm{SL}_2(\mathbb{C}))$ is smooth.

Since every representation in $R(F_2)$ lifts to $\mathrm{SL}_2(\mathbb{C})$, we deduce

$$X(F_2) = X(F_2, \mathrm{SL}_2(\mathbb{C}))//H^1(F_2, C_2).$$

The group $H^1(F_2, C_2) \cong (C_2)^2$ has four elements, and its action on $X(F_2, \mathrm{SL}_2(\mathbb{C}))$ is generated by the involutions

$$\begin{array}{rcl} (I_\alpha, I_\beta, I_{\alpha\beta}) & \mapsto & (-I_\alpha, I_\beta, -I_{\alpha\beta}) \\ (I_\alpha, I_\beta, I_{\alpha\beta}) & \mapsto & (I_\alpha, -I_\beta, -I_{\alpha\beta}). \end{array}$$

Thus $\mathbb{C}[X(F_2), \mathrm{SL}_2(\mathbb{C})]^{H^1(F_2,C_2)}$ is generated by $X = I_\alpha^2$, $Y = I_\beta^2$, $Z = I_{\alpha\beta}^2$ and $W = I_\alpha I_\beta I_{\alpha\beta}$. Hence

$$X(F_2) \cong \{(X, Y, Z, W) \in \mathbb{C}^4 \mid W^2 = XYZ\} \tag{4.2.5}$$

The relationship with Corollary (2.3.2) is given by the change of coordinates (cf. Equality (2.4.5))

$$\begin{cases} J_\alpha = X \\ J_\alpha = Y \\ J_{\alpha\beta} = Z \\ J_{\alpha\beta^{-1}} = XY + Z - 2W. \end{cases}$$

Remark (4.2.6). From Equality (4.2.5) we remark that the singular set of $X(F_2)$ consists of those points such that two of $\{X, Y, Z\}$ vanish. This is the same as the set of characters of representations generated by two rotations of angle π. This is also the set of Ad-reducible but non-reducible representations.

Example (4.2.7). If M is a knot exterior in S^3, then $H_2(\pi_1 M) \cong H_2(M) \cong 0$ and therefore $X(M) = \overline{X}(M)$. When in addition M is a 2-bridge knot exterior, explicit methods of how to compute $X(M)$ are given in [HLM1] and [HLM2], where $X(M)$ for this particular case was already defined as $X(M, \mathrm{SL}_2(\mathbb{C}))//C_2$. The explicit computation for the figure eight knot exterior is found in [GM], for instance.

(4.3) Representations that do not lift.

Proof of Theorem (1.4). The manifold M is a bundle over S^1 with fiber $\dot{T}^2$ a torus minus a disk. Up to homeomorphism, M is described by the action of the monodromy on $H_1(\dot{T}^2, \mathbb{Z})$, which is given by the matrix

$$\begin{pmatrix} 1 & m_2 \\ m_1 & 1 + m_1 m_2 \end{pmatrix}$$

with $m_i \in 2\mathbb{Z}$, $m_i > 0$. We shall show that $X(M) - \overline{X}(M)$ has arbitrarily many components by choosing m_i sufficiently large.

To have a presentation of $\pi_1 M$, we use an automorphism f of $\pi_1 \dot{T}^2$ induced by the monodromy. Since $\pi_1 \dot{T}^2$ is the free group of rank 2 generated by α and β,

$$\pi_1 M = \langle \alpha, \beta, \mu \mid \mu\alpha\mu^{-1} = f(\alpha), \mu\beta\mu^{-1} = f(\beta) \rangle$$

We choose f such that:

$$\begin{cases} \mu\alpha\mu^{-1} & = & \alpha\beta^{m_2} \\ \mu\beta\mu^{-1} & = & \beta(\alpha\beta^{m_2})^{m_1} \end{cases}$$

We choose odd numbers $p_1, p_2 \in 2\mathbb{Z} + 1$, with $1 \leq p_i \leq m_i/2$ and an arbitrary complex number $z \in \mathbb{C}$. By Example (4.2.4), there exist matrices $A_z, B_z \in$

$\mathrm{SL}_2(\mathbb{C})$ with

$$\operatorname{tr}(A_z) = 2\cos\frac{\pi p_1}{m_1},\ \operatorname{tr}(B_z) = 2\cos\frac{\pi p_2}{m_2} \text{ and } \operatorname{tr}(A_z B_z) = z\,.$$

Those trace equalities imply that $A_z^{m_1} = B_z^{m_2} = -\mathrm{Id}$. In particular

$$\begin{array}{rcl} A_z B_z^{m_2} & = & -A_z, \\ B_z(A_z B_z^{m_2})^{m_1} & = & -B_z. \end{array}$$

Let $\rho_z \in R(\Gamma)$ be the representation that $\rho_z(\alpha) = \pm A_z$, $\rho_z(\beta) = \pm B_z$ and $\rho_z(\mu) = \pm\mathrm{Id}$. Since m_1 and m_2 are even, this representation does not lift to $\mathrm{SL}_2(\mathbb{C})$. In addition, for each value of p_1 and p_2 we have defined a one parameter family of characters, with parameter $z = \operatorname{tr}(A_z B_z) \in \mathbb{C}$. By [CCGLS, Proposition 2.4] the dimension of each component of $X(M)$ is at most one, hence different values of p_1 and p_2 give different components. □

5. The singular set of $X(F_n)$

In this section we compute the singular set of $X(F_n)$, but before we need two preliminary subsections: in Subsection (5.1) we recall some basic facts about the Zariski tangent space and Luna's slice theorem, and in Subsection (5.2) we compute the cohomology of free groups with twisted coefficients.

(5.1) The Zariski tangent space. Given a representation $\rho \in R(\Gamma)$, we define the space of cocycles

$$Z^1(\Gamma, \mathrm{Ad}\circ\rho) = \left\{ \theta\colon\ \Gamma \to \mathfrak{sl}_2(\mathbb{C}) \ \middle|\ \begin{array}{c} \theta(\gamma_1\gamma_2) = \theta(\gamma_1) + \mathrm{Ad}_{\rho(\gamma_1)}(\theta(\gamma_2)), \\ \forall\, \gamma_1, \gamma_2 \in \Gamma \end{array} \right\}.$$

Given a smooth path of representations ρ_t, with t in a neighborhood of the origin, one can construct a cocycle as follows:

$$\begin{array}{ccl} \Gamma & \to & \mathfrak{sl}_2(\mathbb{C}) \\ \gamma & \mapsto & \frac{d}{dt}\rho_t(\gamma)\rho_0(\gamma)^{-1}|_{t=0} \end{array}.$$

This construction defines an isomorphism, due to Weil [Weil]:

Theorem (5.1.1) ([Weil]). *The previous construction defines an isomorphism*

$$T_\rho^{\mathrm{Zar}}(R(\Gamma)) \cong Z^1(\Gamma, \mathrm{Ad}\circ\rho).$$

Here $T_\rho^{\mathrm{Zar}}(R(\Gamma))$ denotes the Zariski tangent space in the scheme sense (i.e. the defining ideals are not necessary reduced).

We also consider the space of coboundaries

$$B^1(\Gamma, \mathrm{Ad}\circ\rho) = \left\{ \theta\colon\ \Gamma \to \mathbb{R}^2 \ \middle|\ \begin{array}{c} \text{there exists } a \in \mathfrak{sl}_2(\mathbb{C}) \text{ such that} \\ \theta(\gamma) = \mathrm{Ad}_{\rho(\gamma)}(a) - a,\ \forall\, \gamma \in \Gamma \end{array} \right\}.$$

The isomorphism of Theorem (5.1.1) identifies the subspace of the Zariski tangent space corresponding to the orbits by conjugation with $B^1(\Gamma, \mathrm{Ad}\circ\rho)$. So it seems natural that in some cases $T_\chi^{\mathrm{Zar}}(X(\Gamma))$ is isomorphic to the cohomology group

$$H^1(\Gamma, \mathrm{Ad}\circ\rho) = Z^1(\Gamma, \mathrm{Ad}\circ\rho)/B^1(\Gamma, \mathrm{Ad}\circ\rho)$$

as we will show next.

The stabilizer of a representation $\rho \in R(\Gamma)$ is denoted by

$$Stab_\rho = \{A \in \mathrm{PSL}_2(\mathbb{C}) \mid A\rho A^{-1} = \rho\}\,.$$

In particular, for and Ad-irreducible representation $Stab_\rho$ is trivial.

PROPOSITION (5.1.2). *If ρ is a smooth point of $R(\Gamma)$ with closed orbit, then*

$$T^{\mathrm{Zar}}_{\chi_\rho}(X(\Gamma)) \cong T^{\mathrm{Zar}}_0(H^1(\Gamma, \mathrm{Ad}\circ\rho)//Stab_\rho)\,.$$

Proof. We use the slice theorem of Luna: there exists an algebraic subvariety $S \subset R(\Gamma)$ that contains ρ and that is $Stab_\rho$-invariant, such that

$$Z^1(\Gamma, \mathrm{Ad}\circ\rho) = B^1(\Gamma, \mathrm{Ad}\circ\rho) \oplus T^{\mathrm{Zar}}_\rho(S) \tag{5.1.3}$$

and the map induced by the projection

$$S//Stab_\rho \to X(\Gamma)$$

is an étale isomorphism (in particular their tangent spaces are isomorphic). Since we assume that ρ is a smooth point, Luna's theorem shows that $S//Stab_\rho$ and $T^{\mathrm{Zar}}_\rho(S)//Stab_\rho$ are étale equivalent (see [KSS, p. 97]). Since $T^{\mathrm{Zar}}_\rho(S)$ and $H^1(\Gamma, \mathrm{Ad}\circ\rho)$ are isomorphic as $Stab_\rho$-modules (by Equation (5.1.3)), the proposition follows. □

(5.2) Cohomology of Free groups. We start with irreducible characters:

LEMMA (5.2.1). *Let $\chi_\rho \in X(F_n)$ be an irreducible character. Then*

$$\dim H^1(F_n, \mathrm{Ad}\circ\rho) = 3n - 3.$$

Proof. Notice first that $Z^1(F_n, \mathrm{Ad}\circ\rho) \cong \mathfrak{sl}_2(\mathbb{C})^n \cong \mathbb{C}^{3n}$. Irreducibility implies that $\dim B^1(F_n, \mathrm{Ad}\circ\rho) = 3$, which is maximal (even if Ad-reducible representations have invariant subspaces, irreducibility implies that the eigenvalues are different from 1). □

We are interested in computing $H^1(F_n, \mathrm{Ad}\circ\rho)$ as a $Stab_\rho$-module. If ρ is Ad-irreducible, then $Stab_\rho$ is trivial, and therefore $H^1(F_n, \mathrm{Ad}\circ\rho)$ is the trivial module $\mathbb{C}^{3n-3}$. In the reducible and Ad-reducible cases we need further computations.

Reducible characters. Let $\chi \in X(F_n)$ be a non trivial reducible character. There exists a representation $\rho \in R(F_n)$ with character χ such that ρ consists of diagonal matrices, constructed in Remark (3.1.7).

We decompose the Lie algebra $\mathfrak{sl}_2(\mathbb{C}) = h_0 \oplus h_- \oplus h_+$, where h_0, h_+ and h_- are the one dimensional $\mathbb{C}$-vector spaces generated respectively by $\left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}0 & 1\\ 0 & 0\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}0 & 0\\ 1 & 0\end{smallmatrix}\right)$.

LEMMA (5.2.2). *If ρ is diagonal then* $\mathrm{Ad}\circ\rho$ *preserves the splitting $\mathfrak{sl}_2(\mathbb{C}) = h_0 \oplus h_- \oplus h_+$. If in addition ρ is non-trivial, then $Stab_\rho$ preserves the splitting $\mathfrak{sl}_2(\mathbb{C}) = h_0 \oplus (h_- \oplus h_+)$ (some elements may permute h_+ and h_-).*

Proof. The first assertion is clear, because diagonal matrices preserve each factor h_0 and $h_\pm$.

When the image of ρ has order ≥ 3, the group $Stab_\rho$ is precisely the set of diagonal matrices. When the image has order precisely 2, then $Stab_\rho$ is the group of diagonal and anti-diagonal ones $\left(\begin{smallmatrix}0 & *\\ * & 0\end{smallmatrix}\right)$. Antidiagonal matrices preserve h_0 and permute h_- with h_+, hence the second assertion is proved. □

LEMMA (5.2.3). *Let $\rho \in R(F_n)$ be a non-trivial diagonal representation, then $H^1(F_n, \mathrm{Ad}\circ\rho) \cong h_0^n \oplus (h_+ \oplus h_-)^{n-1}$ as $Stab_\rho$-modules.*

Proof. By construction, $Z^1(F_n, \mathrm{Ad}\circ\rho) \cong \mathfrak{sl}_2(\mathbb{C})^n$. We have the splitting

$$H^1(F_n, \mathrm{Ad}\circ\rho) \cong H^1(F_n, h_0) \oplus H^1(F_n, h_+) \oplus H^1(F_n, h_-).$$

A diagonal matrix $\pm\left(\begin{smallmatrix} a & 0 \\ 0 & a^{-1}\end{smallmatrix}\right)$ acts trivially on h_0 and by multiplication by a factor $a^{\pm 2}$ on $h_\pm$. Therefore $B^1(F_n, h_0) \cong 0$ and $B^1(F_n, h_\pm) \cong h_\pm$, and the lemma follows. $\square$

Ad-reducible but irreducible characters. Let $\rho \in R(\Gamma)$ be irreducible but Ad-reducible. Up to conjugation the image of ρ is contained in the group of diagonal and anti-diagonal matrices. There are two possibilities for the stabilizer $Stab_\rho$. If the image of ρ has more than four elements, then $Stab_\rho$ has two elements: the identity and $\pm\left(\begin{smallmatrix} i & 0 \\ 0 & -i\end{smallmatrix}\right)$. Otherwise the image of ρ is Klein's 4-group (i.e. the group generated by $\pm\left(\begin{smallmatrix} i & 0 \\ 0 & -i\end{smallmatrix}\right)$ and $\pm\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0\end{smallmatrix}\right)$). In this case $Stab_\rho$ equals the image of ρ. With the same argument as in Lemma (5.2.2), one can prove:

LEMMA (5.2.4). *Let ρ be as above. Then both* $\mathrm{Ad}\circ\rho$ *and* $Stab_\rho$ *preserve the splitting* $\mathfrak{sl}_2(\mathbb{C}) = h_0 \oplus (h_+ \oplus h_-)$.

LEMMA (5.2.5). *Let $\rho \in R(F_n)$ be an irreducible but* Ad*-reducible representation, then* $H^1(F_n, \mathrm{Ad}\circ\rho) \cong \mathfrak{sl}_2(\mathbb{C})^{n-1}$ *as* $Stab_\rho$*-modules.*

Proof. Again $Z^1(F_n, \mathrm{Ad}\circ\rho) \cong \mathfrak{sl}_2(\mathbb{C})^n$, and we have the decomposition

$$H^1(F_n, \mathrm{Ad}\circ\rho) \cong H^1(F_n, h_0) \oplus H^1(F_n, h_+ \oplus h_-).$$

The group $B^1(F_n, h_0)$ has dimension one, because the antidiagonal matrices act on h_0 by change of sign. In addition, $\dim(B^1(F_n, h_+ \oplus h_-)) = 2$ is also maximal, because this is the case when we restrict it to diagonal representations (see the proof of Lemma (5.2.3)). $\square$

(5.3) Singular locus for free groups. We saw above that $X(F_2, \mathrm{SL}_2(\mathbb{C})) \cong \mathbb{C}^3$ is smooth. We also showed that the singular points of $X(F_2)$ are Ad-reducible but irreducible characters.

PROPOSITION (5.3.1). *For $n \geq 3$ the singular set of $X(F_n)$ is precisely the set of* Ad*-reducible characters.*

Proof. Since $R(F_n) \cong \mathrm{PSL}_2(\mathbb{C})^n$, $X(F_n)$ is irreducible and of dimension $3n-3$. Thus $\chi \in X(F_n)$ is singular if and only if

$$\dim T^{Zar}_\chi X(F_n) > 3n - 3.$$

This dimension is computed by means of Proposition (5.1.2): if the orbit of $\rho \in t^{-1}(\chi)$ is closed then

$$\dim T^{Zar}_\chi X(F_n) = \dim T^{Zar}_0 (H^1(F_n, \mathrm{Ad}\circ\rho)//Stab_\rho).$$

If $\rho \in R(F_n)$ is irreducible, by Lemma (5.2.1) $\dim H^1(F_n, \mathrm{Ad}\circ\rho) = 3n - 3$. If in addition ρ is Ad-irreducible, then $Stab_\rho$ is trivial and therefore χ_ρ is smooth.

If ρ is irreducible but Ad-reducible, then $H^1(F_n, \mathrm{Ad}\circ\rho) \cong \mathfrak{sl}_2(\mathbb{C})^{n-1}$ as $Stab_\rho$ modules, by Lemma (5.2.5). We may assume that the image of ρ has more than 4 elements, because the adherence set of such characters is the whole set of irreducible but Ad-reducible characters, and the singular set is closed. Hence $Stab_\rho$ is the group generated by the involution $\pm\left(\begin{smallmatrix} i & 0 \\ 0 & -i\end{smallmatrix}\right)$, that acts trivially on h_0 but as a change of sign on $h_+ \oplus h_-$. Thus the action of $Stab_\rho$ on $H^1(F_n, \mathrm{Ad}\circ\rho)$ is equivalent to the involution on $\mathbb{C}^{3n-3}$ that fixes $(n-1)$ coordinates and changes the sign of the remaining $(2n-2)$ coordinates. The quotient of $\mathbb{C}^{3n-3}$ by this involution is not smooth, hence $\dim T_0^{Zar}(H^1(F_n, \mathrm{Ad}\circ\rho)//Stab_\rho) > 3n-3$.

When χ_ρ is reducible but non trivial, we may assume that ρ is diagonal and its image has more that three elements (again the adherence set of those characters is the whole set of reducible ones). Thus $Stab_\rho$ is the group of diagonal matrices, and by Lemma (5.2.3), $H^1(F_n, \mathrm{Ad}\circ\rho) \cong h_0^n \oplus (h_+ \oplus h_-)^{n-1}$ as $Stab_\rho$-module. We have an isomorphism $Stab_\rho \cong \mathbb{C}^*$ and $t \in \mathbb{C}^*$ acts on h_0 trivially and on $h_\pm$ by multiplication by $t^{\pm 1}$. An elementary computation shows that $(h_+ \oplus h_-)^{n-1}//\mathbb{C}^*$ has dimension $2n-3$ and it is not smooth for $n > 2$. $\square$

A similar argument yields that for $n \geq 3$ the singular part of $X(F_n, \mathrm{SL}_2(\mathbb{C}))$ is precisely the set of reducible characters.

Acknowledgment

The second author was partially supported by MCYT through grant BFM2000-0007.

Received February 03, 2003

Final version received September 09, 2003

Michael Heusener
Laboratoire de Mathématiques
UMR 6620 du CNRS
Université Blaise Pascal
63177, Aubière Cedex
France
heusener@math.univ-bpclermont.fr

Joan Porti
Departament de Matemàtiques
Universitat Autònoma de Barcelona
08193 Bellaterra
Spain
porti@mat.uab.es

References

[ATZ] H. Aslaksen, E. Tan, C. Zhu, *Invariant theory of special orthogonal groups*, Pacific J. Math. **168** (1995), 207–215.

[Bea] A.F. Beardon, *The geometry of discrete groups*, Grand. Texts in Math. **91** Springer-Verlag, New York-Berlin-Heidelberg, 1983.

[BMP] M. Boileau, S. Maillot, J. Porti, *Three-dimensional orbifolds and their geometric structures*, Panoramas et Synthèses **14**, Soc. Math. France, Paris, 2003.

[BZ] S. Boyer, X. Zhang, *On Culler-Shalen seminorms and Dehn filling*, Ann. of Math. (2) **148** (1998), 737–801.

[Bur90] G. Burde, *$SU(2)$-representation spaces for two-bridge knot groups*, Math. Ann. **288** (1990), 103–119.

[Cul] M. Culler, *Lifting representations to covering groups*, Adv. in Math. **59** (1986), 64–70.

[CS] M. Culler, P. B. Shalen, *Varieties of group representations and splittings of 3-manifolds*, Ann. of Math. (2) **117** (1983), 109–146.

[CCGLS] D. Cooper, M. Culler, H. Gillet, D.D. Long, P.B. Shalen, *Plane curves associated to character varieties of 3-manifolds*, Invent. Math. **118** (1994), 47–84.

[FH] W. Fulton, J. Harris, *Representation Theory*, Grad. Texts in Math. **129**, Springer-Verlag, New York-Berlin-Heidelberg, 1991.

[GM] F. González-Acuña, J.M. Montesinos-Amilibia, *On the character variety of group representations in* $SL(2,\mathbf{C})$ *and* $PSL(2,\mathbf{C})$, Math. Z. **214** (1993), 627–652.

[HLM1] H.M. Hilden, M.T. Lozano, J.M. Montesinos-Amilibia, *On the character variety of group representations of a 2-bridge link* $p/3$ *into* $\mathrm{PSL}(2,C)$, Bol. Soc. Mat. Mexicana **37** (1992), 241–262.

[HLM2] H.M. Hilden, M.T. Lozano, J.M. Montesinos-Amilibia, *On the arithmetic 2-bridge knots and link orbifolds and a new knot invariant*, J. Knot Theory Ramifications **4** (1995), 81–114.

[Kra] H. Kraft, *Geometrische Methoden in der Invariantentheorie*, Aspects of Mathematics, Vieweg Verlag, Braunschweig, 1985.

[KSS] H. Kraft, P. Slodowy, T. A. Springer, *Algebraische Transformationsgruppen und Invariantentheorie*, DMV Seminar, **13** Birkhäuser Verlag, Basel, 1989.

[PV] V.L. Popov, È.B. Vinberg, *Invariant theory*, In Algebraic Geometry IV, EMS, **55**, Springer-Verlag, Berlin, 1994.

[Ril84] R. Riley, *Non abelian representations of 2-bridge knot groups*, Quart. J. Math. Oxford **2** (1984), 191–208.

[Sha] I.R. Shafarevich, Basic Algebraic Geometry, Springer-Verlag, Berlin, 1977.

[Weil] A. Weil, *Remarks on the cohomology of groups*, Ann. of Math. **80** (1964), 149–157.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

ON 2-UNIVERSAL KNOTS

HUGH M. HILDEN*, MARÍA TERESA LOZANO**, AND JOSÉ MARÍA MONTESINOS–AMILIBIA**

Dedicado a FICO con cariño

Abstract. A knot is said to be 2-universal if every closed orientable 3-manifold occurs as a branched covering of S^3 with branch set the knot and all branching of order one or two. In this paper we show that 2-universal knots exist and we comment on the possible significance of this results.

1. Introduction

In several papers ([17], [2], [3], [4], [5], [6], [16]) universal knots and links have been studied. In these papers it was shown that certain knots and links, such as for example the figure eight knot and the Borromean rings, are universal, and others, such as the trefoil knot, are not.

A knot or link K is said to be universal if every closed orientable 3-manifold occurs as a finite branched covering of S^3 with branch set equal to K.

There are good reasons for refining the definition of universal knot or link to the concept of universal orbifold (see [8], [10], [13]).

In this paper we refine the definitions of universal knot and link, and in particular we define the concept of 2-universal knot and link (corresponding to a π-orbifold), and we show that 2-universal knots exist.

The organisation of the paper is as follows: In section two we give definitions and examples and we introduce the notation that we use throughout the paper.

Our main result, Theorem (3.4), is proved in section three. Our basic idea is to begin with a certain branched covering $p: S^3 \longrightarrow S^3$, which is called a special branched covering, and to perform a series of modifications so as to obtain another branched covering $p: S^3 \longrightarrow S^3$ branched over a knot K. The natura of the modifications is such that a 2-universal link appears as a sublink of the preimage of the branch set. The knot K is then 2-universal. Thus 2-universal knots exist.

In the final section we comment on the importance of 2-universal knots and their possible geometric applications. In particular we raise the question as to whether 2-universal knots occur as the singular sets of hyperbolic orbifolds or cone-manifolds with cone angle 180 degrees. We speculate as to whether a 2-universal knot occurs in Rolfsen's Table ([15]) and, if so, which one it might be.

2000 *Mathematics Subject Classification*: Primary 57M.
Keywords and phrases: universal knot, 3-manifold.
*Partially supported by grant Sab2002-0004 of MECD Spain.
**Partially supported by grants PB98-0826 and BFM2002-4137 Spain.

2. Definitions and examples

In this section "manifold" will always mean closed orientable 3-dimensional piecewise linear manifold. Maps between manifolds will always be piecewise linear. A map between manifolds $p : M \longrightarrow N$ is a branched covering space map if there is a piecewise linearly embedded link L in N such that the map $p : M \setminus p^{-1}(L) \longrightarrow N \setminus L$ is a finite covering space map in the usual sense, and the following additional condition is fulfilled:

Let D be a meridian disc for L; that is to say a disc in N that intersects L in exactly one interior point. Then the preimage of D is a finite collection $D_1, ..., D_m$ of disjoint discs in M. Each disc D_j is mapped by p onto D in a manner equivalent to the map $z \to z^n$, for some n, of the unit disc in the complex plane to itself. In this situation the integer n is independent of the choice of D and depends only on the component of the link $p^{-1}(L)$ that intersects D_j. The link L is called the branch set.

Thus given a branched covering space map $p : M \longrightarrow N$, with branch set L, each component of the link $p^{-1}(L)$ in M is labelled with an integer which is called its ramification index. We also label each component of L in N with the *set* of ramification indices of its preimages. Thus we can speak of a branched covering $p : M \longrightarrow N$, branched over the knot K of type $\{1, 2, 3\}$ for example.

A knot K is said to be *universal of type* $\{a, b, c, ...\}$ if for every closed orientable 3-manifold M, there is a branch covering space map $p : M \longrightarrow S^3$ branched over K of type $\{a, b, c, ...\}$.

A knot K is said to be *2-universal* if it is universal of type $\{1, 2\}$. Later we will show that 2-universal knots exists.

The definition of universal link of a particular type is similar, but a little more complicated, in that there can be different sets of integers attached to different components. As an example we refer to [10, Theorem 1.1] where it is shown that the Borromean rings are universal of such a type that two components are labelled with $\{1, 2, 4\}$ and one of the components is labelled with $\{2, 4\}$. We can compose with the three fold symmetry of the Borromean rings to find that the link 6^2_2 (Rolfsen's notation) is universal with one component labelled $\{1, 2, 4\}$ and the other labelled $\{3\}$.

There is another way to refine the definition of universal knot or link. A branched covering space map $p : M \longrightarrow S^3$ branched over, say, the knot K is completely determined by the unbranched covering space map $p : M \setminus p^{-1}(K) \longrightarrow S^3 \setminus K$. This map induces an injective homomorphism

$$p_* : \pi_1(M \setminus p^{-1}(K)) \longrightarrow \pi_1(S^3 \setminus K)$$

in which the image is a subgroup of $\pi_1(S^3 \setminus K)$ of finite index; say index equal n. Labelling the left cosets of $p_*(\pi_1(M \setminus p^{-1}(K)))$ as

$$\{H_1 = p_*(\pi_1(M \setminus p^{-1}(K))),\ g_2H_1, ..., g_nH_1\},$$

a natural transitive representation of $\pi_1(S^3 \setminus K)$ in Σ_n is induced (left multiplication induces a permutation of left cosets). In this way there is a one to one correspondence between n fold branched covering space maps $p : M \longrightarrow S^3$ branched over K, and transitive representations $\omega : \pi_1(S^3 \setminus K) \longrightarrow \Sigma_n$. Given a transitive representation, ω, the subgroup of the covering is

$$\{g \in \pi_1(S^3 \setminus K) | \omega(g)(1) = 1\}.$$

However a transitive representation $\omega : \pi_1(S^3 \setminus K) \longrightarrow \Sigma_n$ need not be surjective. We say the branched covering space map $p : M \longrightarrow S^3$ is associated to the group $G \subset \Sigma_n$ if $\omega(\pi_1(S^3 \setminus K))$is contained in G.

Given a class of groups $\mathcal{C}$, for example $\mathcal{C}$ might be the class of dihedral groups, or $\mathcal{C}$ might be the class of groups $\{SL(2, finite\ ring)\}$, we say the knot K is universal for the class of groups $\mathcal{C}$ if given any closed orientable 3-manifold M, there is a branched covering space map $p : M \longrightarrow S^3$, branched over K such that the group $G \subset \Sigma_n$ associated to the branched covering space map belongs to the class $\mathcal{C}$.

At this point we can ask several questions (to which we do not know the answers).

Question (2.1). The character variety of representations in $SL(2, \mathbb{C})$ for the figure eight knot, (4_1 in Rolfsen's notation), has been computed in various places (see [18], [7], [9]) and contains the character of many representation of the form

$$\rho : \pi_1(S^3 \setminus 4_1) \longrightarrow SL(2, R)$$

where R is the ring of integers of an algebraic number field k. Let $\mathcal{C}$ be the class of finite groups $G = SL(2, R/I)$ where I is an ideal in R. Is the figure eight knot universal of type $\mathcal{C}$? How about the Borromean rings? In fact we can ask a weaker question.

Question (2.2). Given a 3-manifold M is there a knot K and branched covering map $p : M \longrightarrow S^3$, branched over K such that the group G associated to the covering space map is derived from a representation $\rho : \pi_1(S^3 \setminus 4_1) \longrightarrow SL(2, R)$ by factoring out by some ideal I contained in R?

A positive answer to questions (2.1) and (2.2) would lead to an interesting connection between character varieties and the classifications of 3-manifolds. In the next section we show, among others things, that 2-universal knots exist.

3. 2-Universal knots exist

In this section $p : M^3 \longrightarrow S^3$ will be a *special* branched covering space map branched over a knot K, unless otherwise indicated.

The word *special*, which we will use only in this section, will mean that $\omega : \pi_1(S^3 \setminus K) \longrightarrow \Sigma_n$ sends meridians to transpositions, whereas branched covering of type $\{1, 2\}$ implies that meridians are sent to disjoint products of transpositions.

LEMMA (3.1). *The map $\omega : \pi_1(S^3 \setminus K) \longrightarrow \Sigma_n$ is surjective.*

Proof. We know that ω is transitive by definition. Let $X = \{n_1, ..., n_l\}$ be a subset of $\{1, ..., n\}$ such that $\omega : \omega^{-1}(\Sigma(X)) \longrightarrow \Sigma(X)$ is surjective and X is maximal with respect to this property. Here $\Sigma(X)$ is the full group of permutations of X and it is understood that the elements of $\Sigma(X)$ fix the points not in X.

Suppose $X \neq \{1, ..., n\}$. Since ω is transitive and $\pi_1(S^3 \setminus K)$ is generated by meridians, there must be a meridian m such that $\omega(m) = (a, b)$ where a belong to X but b does not. Let $Y = X \cup \{b\}$, then $\Sigma(Y)$ is generated by (a, b) and $\Sigma(X)$ so that $\omega : \omega^{-1}(\Sigma(Y)) \longrightarrow \Sigma(Y)$ is surjective, contradicting the maximality of X. ☐

In the branched covering $p : M^3 \longrightarrow S^3$ the preimage of K is a link consisting of various components $\{K_1, ..., K_l\}$ of branching index one, and a single component B of branching index two. The component B is called the *branch cover* and the link $K_1 \cup ... \cup K_l$ is called the *pseudo-branch cover.*

LEMMA (3.2). *Let S be any non empty subset of the set of components of the pseudo-branch cover. Let H be the subgroup of $\pi_1(M^3 \setminus p^{-1}(K))$ generated by meridians in the components of S. Then the map*

$$\omega|_{p_*(H)} : p_*(H) \longrightarrow \Sigma_{n-1} = \{\sigma \in \Sigma_n | \sigma(1) = 1\}$$

is surjective.

Proof. The group H is a normal subgroup of $\pi_1(M^3 \setminus p^{-1}(K))$ because the conjugate of a meridian of S is again a meridian of S. Thus since p_* is injective, $p_*(H)$ is normal in $p_*(\pi_1(M^3 \setminus p^{-1}(K)))$ and as ω is surjective, $p_*(H)$ is a normal subgroup of Σ_{n-1}. However, p_* sends meridians of S to meridians of K, not powers of meridians of K, the image of a meridian of K under ω is a transposition. And no proper normal subgroup of Σ_{n-1} contains a transposition. Therefore the image of $p_*(H)$ under ω must be Σ_{n-1} itself and $\omega|_{p_*(H)} : p_*(H) \longrightarrow \Sigma_{n-1}$ is surjective. ☐

Now we would like to describe certain moves that change a special branch covering space map $p : M^3 \longrightarrow S^3$ branched over a link L to a different special branch covering space map $p : \widehat{M}^3 \longrightarrow \widehat{S}^3$ branched over a link $\widehat{L}$. The nature of the moves is such that $M^3 \simeq \widehat{M}^3$ and $S^3 \simeq \widehat{S}^3$ but the link $\widehat{L}$, and therefore its preimages, will be different from the link L and its preimages.

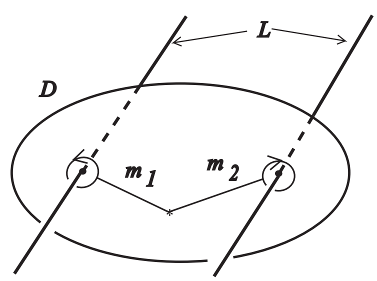


Figure 1. The disc D.

Consider a disc D that cuts the link L transversely in two points as in Figure 1. The branched covering space map when restricted to $p^{-1}(D)$ is a branched covering of D by a disconnected bounded surface. The meridians m_1 and m_2 are sent to transpositions (a, b) and (c, d) respectively by the representation $\omega : \pi_1(S^3 \setminus L) \longrightarrow \Sigma_n$. There are three possibilities; $\{a, b\}$ and $\{c, d\}$ are disjoint, $\{a, b\}$ and $\{c, d\}$ have one common point, and $\{a, b\} = \{c, d\}$. We shall arrange that the case $\{a, b\} = \{c, d\}$ never occurs so we don't need to consider it. The other two cases give rise to moves.

In each of the other two cases the preimage of a disc is a disjoint union of discs. To see this we compute the preimage of D by splitting D along arcs A and B as in Figure 2, following [14].

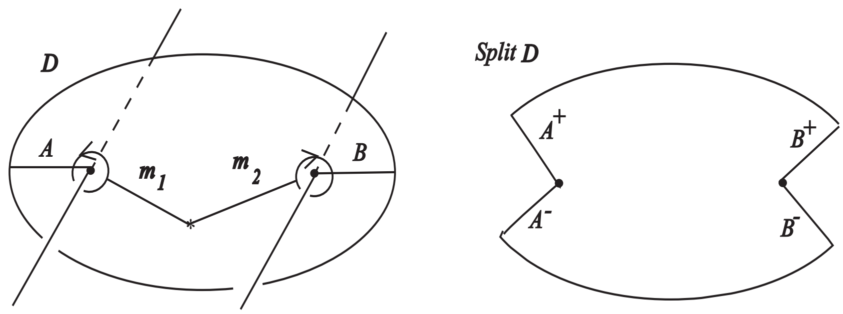


Figure 2. Splitting the disc D along A and B.

We take n copies of D; $D_1, \ldots , D_n$ and if $\omega(m_1) = (j, k)$ we glue A_j^+ to A_k^- and A_j^- to A_k^+ and similarly for B.

In the case where $\{a, b\}$ and $\{c, d\}$ are disjoint, the preimage of D consists of $n - 4$ discs that are mapped homeomorphically by p to D, and two discs that are mapped by p as double branched covers. In this case our move consists of splitting S^3 below along D and splitting along all the preimages of D above. We then do a double Dehn twist below and reglue. We do the same double Dehn disc twist above on all the homeomorphic preimages. In the preimage discs for which p is a double branch cover we do a single Dehn disc twist above and then reglue, as the double Dehn disc twist below lifts to a single Dehn disc twist above. The effect of this move, called a *type I move* is illustrated in Figures 3 and 4. We note that a type I move does not change L, except in a small neighbourhood of D and its preimages, and does not change the number of components of L below.

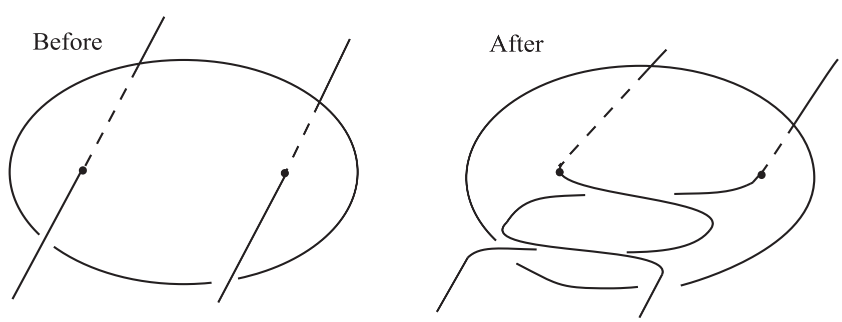


Figure 3. Effect on the disc below and on the $n - 4$ homeomorphic preimages above.

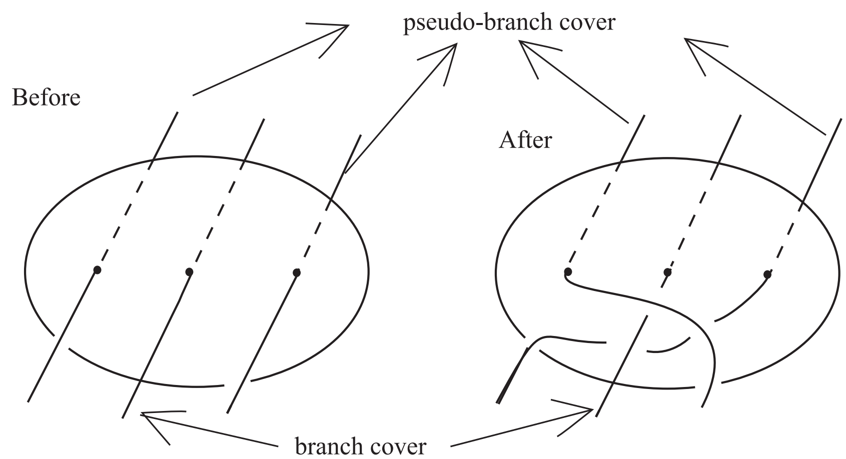


Figure 4. Effect on the disc above that is mapped by p as a double branched cover of D.

In the case where $\{a, b\}$ and $\{c, d\}$ intersect in one common point, the preimage of D consists of $n-3$ discs that are mapped homeomorphically by p to D, and one disc that is mapped as a three-to-one irregular branched cover of D. In this case again our move consists of first splitting S^3 along D below and splitting M^3 along all the disc preimages of D above. This time we do a triple Dehn disc twist below before regluing (the single and double Dehn disc twists do not lift). Above we do a triple Dehn disc twist on all the homeomorphic preimages and a single Dehn disc twist on the preimage that is mapped as a three-to-one irregular cover. Then we reglue, the effects of this move, called a *type II move*, are illustrated in Figures 5 and 6. Again we note that a type II move changes nothing outside a small neighbourhood of D and the preimages of this small neighbourhood of D.

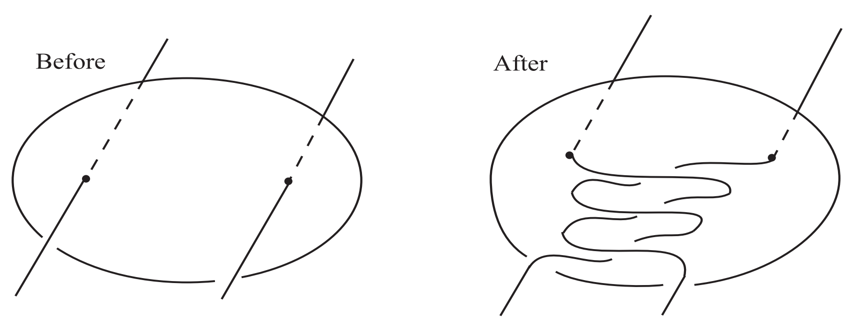


Figure 5. Effect on the disc below and on the $n-3$ homeomorphic preimages above.

We note that moves of type I and II have appeared before (see [10], [4]).

Our next task will be to produce a specific special branched covering space map $p : S^3 \longrightarrow S^3$ branched over a knot. We start with the trivial link of $n-1$

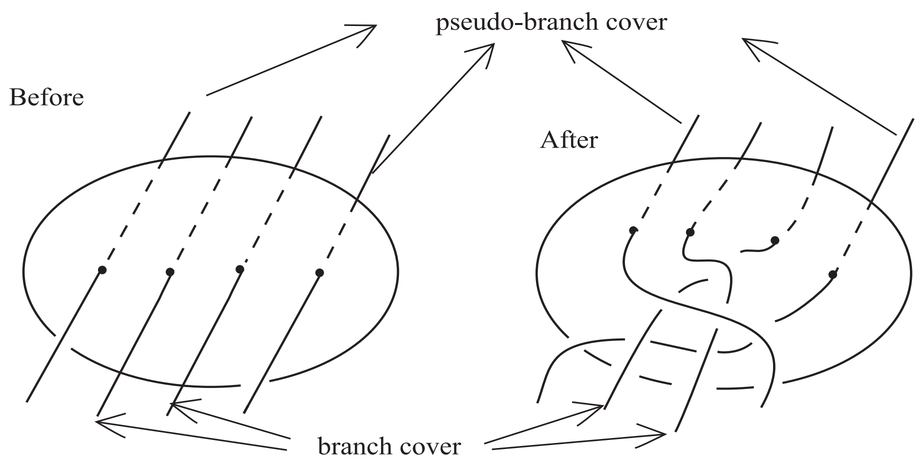


Figure 6. Effect on the disc above that is mapped by p as a 3 to 1 irregular cover of D.

components, and assign to the meridian generators of its fundamental group, which is a free group, the transpositions indicated in Figure 7.

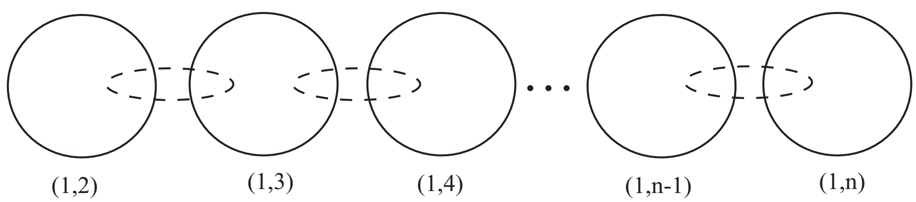


Figure 7. The trivial $n-1$ component link and assigned transpositions.

This gives a special n to 1 branched covering of S^3 to S^3. We then do type II moves between adjacent circles in the discs indicated by the dotted lines to obtain the connected sum of trefoils where we twist in opposite directions in adjacent discs, such that the last one is in the right handed direction. In this way we obtain the branched covering $p : S^3 \longrightarrow S^3$ branched over a knot K with images of meridians as indicated in Figure 8.

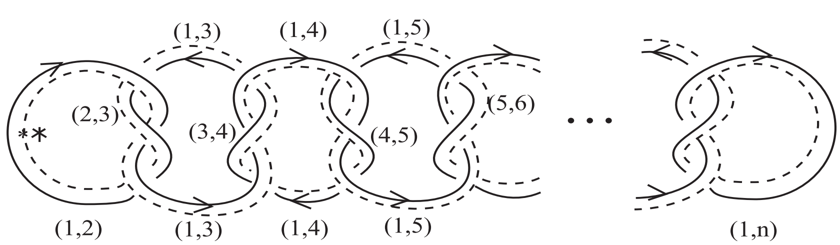


Figure 8. The knot K.

The knot K is the connected sum of $n-2$ trefoils, some of which are right handed and some of which are left handed. The dotted line is just the knot pushed off to its right.

This curve, together with a meridian, forms a basis for the fundamental group of the torus boundary of a regular neighbourhood of the knot K. The knot K is obtained by pasting together several segments of the type pictured in the right hand side of Figure 9, plus two trivial arcs in the case of an even connected sum of trefoils, or plus a trivial arc and the segment in the left hand side of Figure 9 in the case of an odd connected sum.

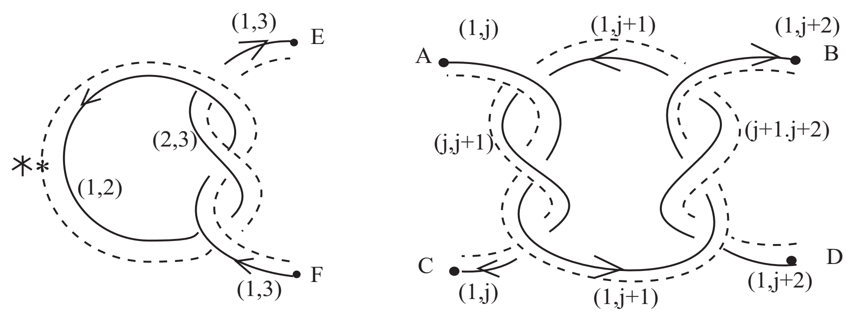


Figure 9. Computing the image of the dotted curve.

Next we compute the image of the dotted curve under $\omega : \pi_1(S^3 \setminus K) \longrightarrow \Sigma_n$. As we move from A to B we obtain $(j, j+1)(j+1, j+2) = (j, j+2, j+1)$. As we move from C to D we obtain $(1, j+1)(1, j)(1, j+2)(1, j+1) = (j, j+2, j+1)$, which is the same permutation.

Starting at the extreme right of the knot in Figure 8 and working toward the left, an inductive argument shows that in the case where we have the connected sum of an even number of trefoils, that the image of the dotted curve under ω is the identity. And in the case where we have the connected sum of an odd number of trefoils, multiplying transpositions assigned to over crossings as we move to the right from E in Figure 9 though the rest of the knot returning to F from the right also gives the identity. The rest of the computation, the part pictured in Figure 9, shows that, the element of Σ_n assigned to the dotted curve is $(1, 3)(1, 2)(1, 3)$ which is equal to $(1, 2)$ using $\star$ as a basepoint.

In both the odd and the even cases we see that the image by ω of the fundamental group of a torus boundary T of a regular neighbourhood of K in Σ_n is a two element group. If follows that the preimage of this torus neighbourhood consists of $n-2$ tori that are mapped homeomorphically to T and one torus that is mapped to T as a double cover. And from this it follows that the pseudo-branch cover of K has $n-2$ components and the branch cover one. For convenience we summarise the above in a lemma.

LEMMA (3.3). *For $n \geq 3$, there exists a knot K, and a special branched covering $p : S^3 \longrightarrow S^3$ branched over K, such that the pseudo-branch cover has $n-2$ components and the branch cover has only one.* $\square$

We now state and prove our main theorem. We remark that our proof is constructive.

Theorem (3.4). *Let L be any m component link with $m \geq 1$. There is a special $(2m+3)$ to 1 branched covering space map $p : S^3 \longrightarrow S^3$ branched over a knot K such that L is a sublink of the pseudo-branch cover.*

An immediate corollary of this theorem, the case $m = 1$, is the following

Corollary (3.5). *Let $\widehat{K}$ be any knot. There is a special five to one branched covering $p : S^3 \longrightarrow S^3$ branched over a knot K such that $\widehat{K}$ is contained in the pseudo-branch cover.*

The above corollary also makes sense for special three to one covers, but is it true? Or, stated differently, which knots occur as the pseudo-branch covers of special three to one branched coverings $p : S^3 \longrightarrow S^3$ branched over knots?

Proof of Theorem (3.4). Let $p : S^3 \longrightarrow S^3$ be an $n = 2m+3$ to 1 special branched covering space map branched over the knot K_0, such as is described in Lemma (3.3), so that the pseudo-branch cover has $2m+1$ components. Let L_0 be an m component sublink of the pseudo-branch cover, let M_0 be the link consisting of the other $m+1$ components and let B_0 be the branch cover.

Consider a regular projection of any link. This can be changed to a regular projection of the trivial link simply by changing some of the crossings. It follows that any link can be changed to any other link of the same number of components by a finite number of operations that we shall describe and that are analogous to crossing changes.

Let A be an arc connecting two points P and Q of the link in the same or different components. Let D be a disc intersecting the link in points P and Q as in Figure 10.

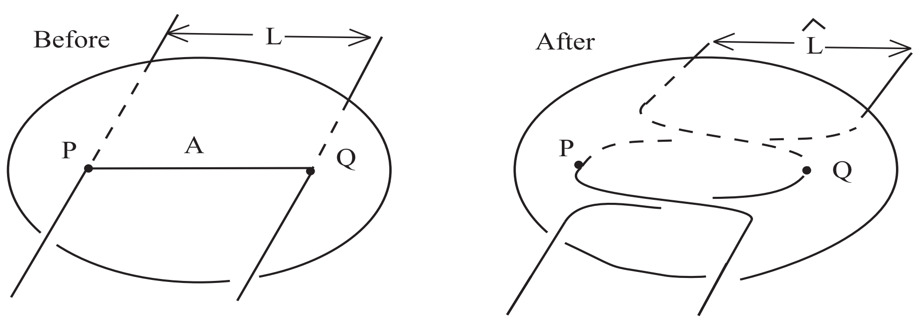


Figure 10.

Replace the link L by the link $\widehat{L}$ where $\widehat{L}$ differs from L only in a small neighbourhood of the disc D as indicated in Figure 10.

Thus there is a sequence of links L_0, L_1, ... , L_t, all with m components, such that L_{j+1} is obtained from L_j by this operation and L_t is the desired link L.

In order to prove Theorem (3.4) it suffices to show there are sequences of links B_0, B_1, ... , B_t and M_0, M_1, ... , M_t and a sequence of special branched covering space maps $\{p_j : S^3 \longrightarrow S^3\,; 0 \le j \le t\}$ branched over knots K_j, $0 \le j \le t$, such that the pseudo-branch cover for p_j is $L_j \cup M_j$ and the branch cover is B_j.

We shall prove this theorem by induction. The initial step of the induction is done as p_0, B_0, L_0, M_0 and K_0 have already been defined.

Suppose, as inductive hypothesis, that $p_j : S^3 \longrightarrow S^3$ is a special branched covering space map branched over the knot K_j with pseudo-branch cover $L_j \cup M_j$ where L_j has m components.

Also suppose that D^j is the disc whereby L_{j+1} is obtained from L_j via the operation illustrated in Figure 10. The operation illustrated in Figure 10 is analogous to what happens in Figure 3 in the "above"discs when we do a type I move. The problem is that we must also do the type I move on all the other "above"discs and on the disc below in order to get another branched covering $p_{j+1} : S^3 \longrightarrow S^3$ branched over a new knot K_{j+1}. We next show how to deal with this problem.

Suppose that the following set of conditions, which we call Hypotheses set A, are satisfied.

Hypotheses set A.

A.1 *The disc D^j is mapped by p_j homeomorphically onto its image.*

A.2 *The preimage of the image of the disc D^j consists of $n-4$ discs, $D^j = D_1^j$ and D_2^j, ... , D_{n-4}^j, each of which is mapped homeomorphically onto its image, together with two other discs E_1^j and E_2^j, each of which is mapped as a double branched covering onto the disc $p_j(D^j)$.*

A.3 *The disc D^j intersects L_j exactly in the two points P and Q.*

A.4 *Each disc D_i^j; $2 \le i \le n-4$; intersects L_j in one point and M_j in one point.*

A.5 *The discs E_1^j and E_2^j each intersects B_j in one point and M_j in two points.*

Under these hypotheses the type I moves on the discs D_i^j; $2 \le i \le n-4$ and the discs E_1^j and E_2^j have no effect on the link L_j although they radically alter the pseudo-branch cover $L_j \cup M_j$ as a whole. The move in the disc D_1^jchanges L_j to L_{j+1} and the type I move below changes the knot K_j to a new knot K_{j+1}. Thus to complete the proof it suffices to show that Hypotheses set A can be satisfied.

By general position we can isotope the arc A^j so that it is embedded by p_j and we can shrink the disc D^j so that it too is embedded thus satisfying A.1. We choose a point x_1 on the arc A^j to serve as base point for the group $\pi_1(S^3 \setminus p_j^{-1}(K_j))$. Let $x_0 = p_j(x_1)$ so that x_0 is the base point for $\pi_1(S^3 \setminus K_j)$ below and label the other preimages of x_0 as x_2, ... , x_n. Then $\omega : \pi_1(S^3 \setminus K_j, x_0) \longrightarrow \Sigma_n$ is defined by the condition $\omega(g) : i \to k$ if the lift of a closed curve representing g that begins at x_i ends at x_k.

The subgroup of the covering space is then $\{g \in \pi_1(S^3 \setminus K_j, x_0) | \omega(1) = 1\}$.

We denote the points in which arc $A_1^j = A^j$ intersects L_j by P_1^j and Q_1^j, and we choose meridian discs for these points which we denote O_{P1}^j and O_{Q1}^j, respectively. We denote the homeomorphic images of the points, arc, and discs below by using the subscript 0 instead of 1. All this is illustrated in Figure 11.

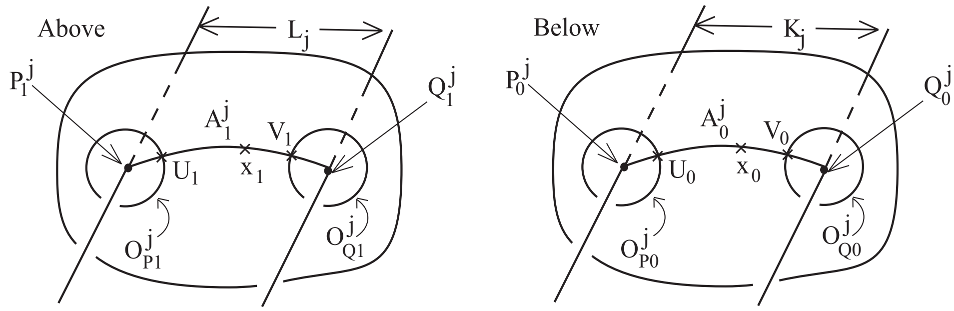


Figure 11. Notation on the discs D_1^j and $p(D_1^j)$.

We label the point $A_1^j \cap \partial O_{P1}^j$ by U_1, $A_0^j \cap \partial O_{P0}^j$ by U_0, $A_1^j \cap \partial O_{Q1}^j$ by V_1, $A_0^j \cap \partial O_{Q0}^j$ by V_0, where ∂O is the boundary of the disc O.

The preimage of O_{P0}^j consists of $n-2$ discs that are mapped homeomorphically onto O_{P0}^j by p_j, including O_{P1}^j, and one that is mapped onto O_{P0}^j as a double branched covering. Since $n = 2m + 3$ and each component of L_j is mapped homeomorphically onto K_j by p_j, we see that exactly m of the discs in the preimage of O_{P0}^j that are mapped homeomorphically onto O_{P0}^j intersect L_j, and $m + 1$ intersect M_j. We label the discs intersecting L_j as O_{P1}^j, ... , O_{Pm}^j and those intersecting M_j as $\widehat{O}_{Pm+1}^j$, ... , $\widehat{O}_{P2m+1}^j$. We label the disc that double covers O_{P0}^j as $\widetilde{O}_{P2m+2}^j$.

The above remarks apply also to the discs O_{Q1}^j and O_{Q0}^j as well and we label the preimages of O_{Q0}^j by O_{Q1}^j, ... , O_{Qm}^j, $\widehat{O}_{Qm+1}^j$, ... , $\widehat{O}_{Q2m+1}^j$ and $\widetilde{O}_{Q2m+2}^j$ in corresponding fashion.

The arc that goes from U_0 to V_0 lifts to $n = 2m + 3$ arcs, one of which goes from U_1 on O_{P1}^j to V_1 on O_{Q1}^j.

Suppose we could arrange that the other $2m + 2$ lifts satisfy the following conditions which we call Hypotheses set B.

Hypotheses set B.

B.1 *Those lifts that begin at a point on* O_{Pi}^j *end at a point on* $\widehat{O}_{Qm+i}^j$ *for* $2 \le i \le m$.

B.2 *Those lifts that begin at a point on* $\widehat{O}_{Pm+i}^j$ *end at a point on* O_{Qi}^j *for* $2 \le i \le m$.

B.3 *Those lifts that begin at a point on* $\widehat{O}_{Pm+1}^j$ *and* $\widehat{O}_{P2m+1}^j$ *end at points on* $\widetilde{O}_{Q2m+2}^j$.

B.4 *The two lifts that begin at points on* $\widetilde{O}_{P2m+2}^j$ *end at the points* $\widehat{O}_{Qm+1}^j$ *and* $\widehat{O}_{Q2m+1}^j$.

If we could do this then the intersection properties A.2 through A.5 would be satisfied and we would be done. The rest of the proof consists of showing that we can in fact do this.

We can arrange that B.1 though B.4 are satisfied in the following way: Let σ be a fixed permutation of the numbers $\{2, ..., n\}$ to be chosen shortly. By Lemma (3.2), with M_j playing the role of S, there is an element $[\alpha]$ in $\pi_1(S^3\backslash p^{-1}(K_j); x_1)$ that is a product of meridians of S such that $\omega(p_*([\alpha])) = \sigma \in \Sigma_{n-1}$.

Consider the lifts of the curve γ that we now define. The curve γ consists of the arc $[U_0, x_0]$ of Figure 11, followed by α, a curve representing $[\alpha]$, followed by the arc $[x_0, V_0]$.

The lift of the arc $[U_0, x_0]$ that begins at U_k^j ends at x_i for some i. The lift of α that begins at x_i ends at $x_{\sigma(i)}$. The lift of the arc $[x_0, V_0]$ that begins at $x_{\sigma(i)}$ ends at V_l^j for some l. Since we are free to choose any $\sigma \in \Sigma_{n-1}$ we can always arrange that the lifts of γ satisfy conditions B.1 though B.4.

Next we show that there is an isotopy of S^3 above that fixes the branch cover and L^j but moves M^j, and such that, after performing the isotopy, the arc $[U_0, V_0]$ satisfies all the conditions of Hypotheses set B.

First let $\alpha = \prod_{i=1}^r \beta_i$ where β_i is a meridian of M^j. We homotope α slightly so that the meridian discs are all disjoint, and the arcs leading from x_1 to the meridian discs are all disjoint, and then we modify the new α in the manner indicated in Figure 12 defining discs $G_1, ... , G_r$, one for each meridian curve, in the manner indicated in Figure 12 where we illustrate the case of three meridian curves.

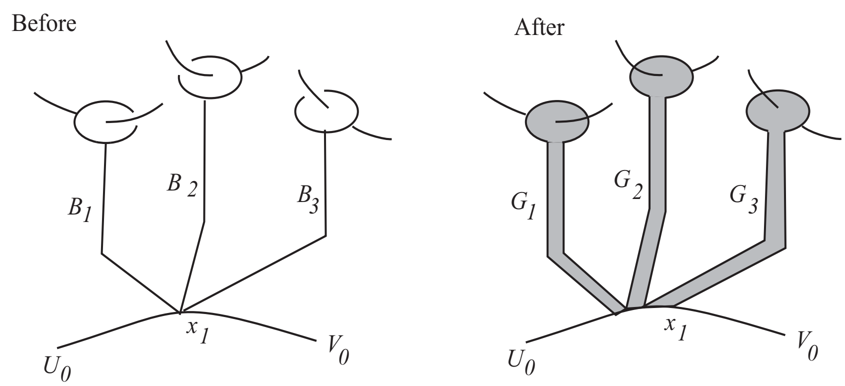


Figure 12. The discs G_i.

The boundary of the discs G_i consists of a small arc on $[U_0, V_0]$, two curves close to and parallel to the arc leading to the meridian disc and most of the boundary of the meridian disc.

We can assume that $\widehat{\gamma}$ is an embedded curve in S^3 above and that the set of discs $\{G_j\}$ are pairwise disjoint.

Now we simply perform ambient isotopies in a small neighbourhood of each G_j so that the arc of the boundary of the disc G_j consisting of the two arcs from the arc $[U_0, V_0]$ to the meridian disc and an arc from the boundary of the meridian disc is pushed onto the arc of the boundary of each G_j that is the intersection of G_j with $[U_0, V_0]$. During the course of this ambient isotopy properties B.1 though B.4 are preserved.

This concludes the proof of the theorem. □

For examples of 2-universal links we refer the reader to [13] where some seventeen different 2-universal links are exhibited. One of them, with only three components, is reproduced below in Figure 13.

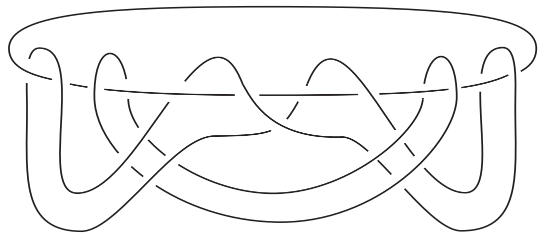


Figure 13. A 2-universal link.

Corollary (3.6). *Knots that are 2-universal exist.*

Proof. Let L be a 2-universal link and apply Theorem (3.4) to find a special branched covering space $p : S^3 \longrightarrow S^3$ branched over a knot K such that L occurs as a sublink of the pseudo-branch cover.

Given any closed orientable 3-manifold M^3 there is a branched covering $q : M^3 \longrightarrow S^3$ of type $\{1, 2\}$ branched over L. Then $p \circ q : M^3 \longrightarrow S^3$ is a branched covering branched over the knot K of type $\{1, 2\}$. Therefore K is a 2-universal knot. □

In the final section we discuss the significance of Corollary (3.6)

4. Discussion of the significance of 2-universal knots

Suppose that K is a 2-universal hyperbolic knot. Then there is a one parameter family of hyperbolic cone-manifolds, topologically equal to S^3, with singular set the knot K. The parameter can be chosen to be the cone angle, in which case the parameter varies from zero, corresponding to the complete hyperbolic structure on the knot complement, to the angle θ_h called the limit of hyperbolicity. For values θ, with $0 < \theta < \theta_h$, S^3 has the structure of hyperbolic cone-manifold with singular set the knot K and cone angle θ, but no such structure exists for $\theta = \theta_h$. (As background for the above, see [1]).

Suppose, for a knot K, that (S^3, K) has spherical cone-manifold structure at 180^o. Let $p : M^3 \longrightarrow S^3$ be a branched covering of type $\{1, 2\}$. The spherical cone manifold structure of S^3 lifts, via p, to a spherical cone manifold structure on M^3. The induced cone manifold structure on M^3 can be altered slightly in a neighbourhood of the singular set to an actual Riemannian structure, (no singularities), that has non negative curvature at all points. Thus M^3 cannot be hyperbolic, (compare [11], [12]), and therefore K cannot be 2-universal as no hyperbolic 3-manifolds occur as branched covering of K of type $\{1, 2\}$.

Since two bridge knots have such spherical cone-manifold structure at 180^o, they cannot be 2-universal.

We have computed the character variety for almost all knots of nine crossings and many ten crossing knots in Rolfsen's table. A short list of some of the (necessarily three bridge) hyperbolic knots for which θ_h exceeds 180^0 is $\{8_{18}, 9_{40}, 9_{41}, 9_{47}, 9_{49}, 10_{123}, 10_{161}\}$. We do not have enough nerve to conjecture that one of these is 2-universal but we do point out that it is possible and here is why that would be interesting.

A hyperbolic 2-universal knot would give rise to a hyperbolic orbifold structure on S^3 with angle 180^o. This would, in turn, give rise to a discrete universal group of hyperbolic isometries, call it G_0.

A discrete group G of hyperbolic isometries is said to be universal if, given any closed orientable 3-manifold M^3, there is a finite index subgroup H of G such that M is homeomorphic to the orbit space H^3/G. Universal groups necessarily contain rotations as not every 3-manifold is hyperbolic. The universal group U, defined in [10] is generated by three 90^o rotations and contains only rotations of 90^o and 180^o. The group G_0, if it exists, would contain only 180^o rotations and this would make it an interesting group, indeed.

Received June 18, 2003

Final version received May 03, 2004

Hugh M. Hilden
Department of Mathematics
University of Hawaii
Honolulu, HI 96822
USA
mike@math.hawaii.edu

María Teresa Lozano
Departamento de Matemáticas
Universidad de Zaragoza
50009 Zaragoza
Spain
tlozano@unizar.es

José María Montesinos–Amilibia
Departamento de Geometría y Topología
Facultad de Matemáticas
Universidad Complutense
28040 Madrid
Spain
montesin@mat.ucm.es

References

[1] M. Boileau, B. Leeb, J. Porti, *Uniformization of small 3-orbifolds*, C. R. Acad. Sci. Paris Sér. I Math. **332** (1), (2001), 57–62.

[2] H.M. Hilden, M.T. Lozano, J.M. Montesinos, *Universal knots*, Bull. Amer. Math. Soc. (N.S.) **8** (3) (1983), 449–450.

[3] H.M. Hilden, M.T. Lozano, J.M. Montesinos, *The Whitehead link, the Borromean rings and the knot* 9_{46} *are universal*, Collect. Math. **34** (1), (1983), 19–28.

[4] H.M. Hilden, M.T. Lozano, J.M. Montesinos, *Universal knots*, Knot theory and manifolds (Vancouver, B.C., 1983), 25–59, Lecture Notes in Math. **1144**, Springer, Berlin, 1985.

[5] H.M. Hilden, M.T. Lozano, J.M. Montesinos, *On knots that are universal*, Topology **24** (4), (1985), 499–504.

[6] H.M. Hilden, M.T. Lozano, J.M. Montesinos, *Nonsimple universal knots*, Math. Proc. Cambridge Philos. Soc. **102** (1987), 87–95

[7] H.M. Hilden, M.T. Lozano and J.M. Montesinos-Amilibia, *The Arithmeticity of the Figure Eight Knot Orbifolds*, in TOPOLOGY'90. Ohio State University, Math. Research Inst. Pub. 1 (eds. B Apanasov, W. Neumann, A. Reid and L. Siebebenmann) De Gruyter, (1992), 169–183.

[8] H.M. Hilden, M.T. Lozano and J.M. Montesinos-Amilibia, *Universal 2-bridge knot and link orbifolds*, J. Knot Theory Ramifications **2** (1993), 141–148.

[9] H.M. Hilden, M.T. Lozano and J.M. Montesinos-Amilibia, *Character varieties and perhipheral polynomials of a class of knots*, J. Knot Theory Ramifications **12** (2003), 1093–1130.

[10] H.M. Hilden, M.T. Lozano, J.M. Montesinos and W.C. Whitten, *On universal groups and three-manifolds*, Invent. math. **87** (1987), 441–456.

[11] K. N. Jones, *Geometric structures on branched covers over universal links*, Geometric topology (Haifa,1992), 47–58, Contemp. Math. **164** Amer. Math. Soc., Providence, RI, (1994),

[12] K. N. Jones, *The structure of closed non-positively curved Euclidean cone* 3*-manifolds*, Pacific J. Math. **163** (1994), 297–313.

[13] M.T. Lozano,and J.M. Montesinos-Amilibia, *Geodesic flows on hyperbolic orbifolds, and universal orbifolds*, Pacific J. Math. **177** (1997), 109–147.

[14] L. Neuwwirth, Knot groups, Ann. of Math. Stud. **56** Princeton Univ. Press, Princeton, NJ, 1965.

[15] D. Rolfsen, Knots and Links, Mathematics Lecture Series 7, Publish or Perish, Inc., Berkeley, CA 1976.

[16] Y. Uchida, *Universal chains.* Kobe J. Math. **8** (1991), 55–65.

[17] W. Thurston, *Universal links*, Preprint (1982)

[18] A. Whittemore, *On representations of the group of Listing's knot by subgroups of* $SL(2,\mathbb{C})$, Proc. Amer. Math. Soc. **40** (1973), 378–382.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

SOME RESULTS ON ONE-RELATOR SURFACE GROUPS

JAMES HOWIE

To Fico González Acuña on his 60'th birthday

Abstract. A one-relator surface group is the quotient of the fundamental group of an orientable surface by the normal closure of a single element. Inspired by a question from González Acuña and by a paper of Hempel, we extend a number of the classical theorems of one-relator group theory to one-relator surface groups.

1. Introduction

This short note was inspired by a question from Fico González Acuña:

Question (1.1). *If α and β are two closed curves (nonsimple, in general) on an orientable surface S, such that the normal closures of α and β in $\pi_1(S)$ coincide, is β freely homotopic to $\alpha^{\pm 1}$?*

If S is noncompact, or has nonempty boundary, then $\pi_1(S)$ is free, and the answer to Question (1.1) is yes, by an old result of Magnus [7] on one-relator groups. (Essentially, the defining relator in a one-relator group on a given generating set is unique up to conjugacy and inversion.)

We will show (see Theorem (3.4) below) that Question (1.1) also has an affirmative answer in the case of a closed surface S. In this case Question (1.1) can be interpreted in terms of *one-relator surface groups*, as introduced by Hempel [3]. Among other results, Hempel proved analogues for one-relator surface groups of two theorems from one-relator group theory: (i) a one-relator surface group is locally indicable if and only if the relator is not a proper power in $\pi_1(S)$; (ii) a closed curve α in S lifts (up to homotopy) to a simple closed curve in the covering space corresponding to the normal closure of α in $\pi_1(S)$. These are analogues of results of Brodskiĭ[1] and Weinbaum [15] respectively. (In the latter case, the original form states that proper subwords of the defining relator represent nontrivial elements in a one-relator group.) Hempel [3] also proved (iii) that a power β^n of a simple closed curve β can belong to the normal closure in $\pi_1(S)$ of a curve α only in the obvious cases: either α is isotopic in S to β^m with $m|n$; or α is a nonseparating curve in a punctured torus in S bounded by β.

The purpose of this note is to show that many other results from one-relator group theory have natural analogues for one-relator surface groups. In most (but not all) cases, the proofs can be obtained by using a trick from [3] to reduce us to the classical one-relator case.

2000 *Mathematics Subject Classification*: Primary 20F06; Secondary 20F10, 20F32, 57M05, 57M07.

Keywords and phrases: surface group, one-relator.

Interest in one-relator surface groups first appeared in the work of Papakyriakopoulos [12], who reduced the Poincaré conjecture to two conjectures which can be expressed in terms of certain one-relator surface groups. (See [10, 13, 14] for more on these conjectures).

2. One-relator surface groups

By a *one-relator surface group* we will mean, following Hempel [3], the quotient of the fundamental group $\pi_1(S)$ of a connected, orientable surface S by the normal closure of a single element α. We will denote this group by $\pi_1(S)/\alpha$. In particular, any (countable) one-relator group can be regarded as a one-relator surface group by choosing S to be noncompact (or ∂S to be nonempty). We will consistently abuse notation to regard α as an immersed closed curve in S.

Since one-relator quotients of the torus group are well understood, we may in practice restrict attention to the case where S is a closed orientable surface of genus at least 2. The following basic trick is employed by Hempel in [3] to reduce his analogue of Brodskiĭ's Theorem to the classical case. We follow [3] in using $\langle -, - \rangle$ to denote the integer-valued algebraic intersection pairing on $H_1(S)$ or $\pi_1(S)$ as appropriate.

Proposition (2.1). *Let S be a closed, connected, oriented surface of genus at least 2, and let α be a closed curve in S. Then*

1. *There is a non-separating simple closed curve β in S such that $\langle \alpha, \beta \rangle = 0$.*
2. *For any such β, there are connected surfaces F, F_0, F_1 and a closed curve α' in F, such that*

(a) *$F_0 \cong F_1$, $F_0 \subset F$ and $F_1 \subset F$;*

(b) *$\pi_1(F_0) \to \pi_1(F)/\alpha'$ and $\pi_1(F_1) \to \pi_1(F)/\alpha'$ are injective;*

(c) *$\pi_1(S)$ (resp. $\pi_1(S)/\alpha$) is an HNN-extension of $\pi_1(F)$ (resp. $\pi_1(F)/\alpha'$) with associated subgroups $\pi_1(F_0)$ and $\pi_1(F_1)$;*

(d) *Each of ∂F, ∂F_0 and ∂F_1 consists of two circles, each of which represents (a conjugate of) $\beta \in \pi_1(S)$.*

Proof. The first part is Lemma 2.1 of [3]. The second is implicit in the proof of Theorem 2.2 of [3]. For completeness we repeat the argument here. Let S_0 denote the surface obtained from S by cutting along β, let S_n be an isomorphic copy of S_0 for each integer n, and form a covering $\tilde{S}$ of S from $\bigcup_{n\in\mathbb{Z}} S_n$ by joining one of the two boundary components of S_n to the other boundary component of S_{n+1}, for all n. Note that $\tilde{S}$ is the infinite cyclic covering of S corresponding to the kernel of $\langle -, \beta \rangle : \pi_1(S) \to \mathbb{Z}$.

There is a minimum $n \geq 0$ such that $S_0 \cup S_1 \cup \cdots \cup S_n$ contains a closed curve α' homotopic to a lift of α. Define $F = S_0 \cup S_1 \cup \cdots \cup S_n$, $F_0 = S_0 \cup \cdots \cup S_{n-1}$ and $F_1 = S_1 \cup \cdots \cup S_n$ (provided $n > 0$). Then property (b) follows from the classical Freiheitssatz of Magnus for one-relator groups [6], using the fact that α' cannot be homotoped into F_0 or F_1. The remaining properties are clear from the construction.

For the case $n = 0$ we adapt the construction slightly as follows: F_0 and F_1 are annuli which are regular neighbourhoods in $\tilde{S}$ of the two boundary components of S_0 (with $F_0 \cong F_1$ via a covering transformation), and $F = F_0 \cup S_0 \cup F_1$. $\square$

3. Results using Hempel's trick

In this section we list some results which follow easily from Hempel's trick. The first two were proved by Hempel in [3].

Theorem (3.1). [3, Theorem 2.2] *Let S be an oriented surface and α an essential closed curve in S. Then the following are equivalent:*
1. *α is not homotopic to β^m for any curve β and any integer $m > 1$;*
2. *$\pi_1(S)/\alpha$ is locally indicable;*
3. *$\pi_1(S)/\alpha$ is torsion-free.*

Theorem (3.2). [3, Theorem 2.3] *Let S be an oriented surface and α a closed curve in S. Then each lift of α to the regular covering corresponding to the normal closure of α in $\pi_1(S)$ is (homotopic to) a simple closed curve.*

Corollary (3.3). *If α is homotopic to β^m in $\pi_1(S)$ for some curve β and integer $m \geq 1$, then β has order m in $\pi_1(S)/\alpha$.*

Proof. Clearly $\beta^m = 1$ in $\pi_1(S)/\alpha$. On the other hand, β^m lifts to a simple closed curve (up to homotopy) in the covering corresponding to the normal closure N of α in $\pi_1(S)$, so for $0 < k < m$, β^k does not lift to a closed curve. In other words, $\beta^k \notin N$. □

The next result answers Question (1.1), and generalises the result of Magnus that was mentioned in the introduction.

Theorem (3.4). *Let S be an oriented surface and α, β two closed curves in S whose normal closures in $\pi_1(S)$ coincide. Then α is freely homotopic to $\beta^{\pm 1}$.*

Proof. For this we use the proof, as well as the statement, of Proposition (2.1). If either of α, β is nullhomotopic, then clearly so is the other, so we may assume that both α and β are essential. Let γ be a simple closed curve in S such that $\langle \alpha, \gamma \rangle = 0$ (and hence also $\langle \beta, \gamma \rangle = 0$, since the normal closures of α and β in $\pi_1(S)$ coincide, and so α and β are homologous). In the notation of Proposition (2.1), suppose that $F = S_0 \cup \cdots \cup S_n$ contains a closed curve α' homotopic to a lift of α, and that n is minimal with this property.

Similarly, suppose that $F' = S_0 \cup \cdots S_{n'}$ contains a closed curve β' homotopic to a lift of β, and that n' is minimal with respect to this property. Suppose that $n' < n$. Then $F' \subset F_0$, and $\pi_1(F_0)$ embeds into $\pi_1(S)/\alpha$. Hence β', and hence also β, must be nullhomotopic, contrary to assumption. Hence $n' \geq n$. By a symmetric argument $n \geq n'$, so $n = n'$ and $F' = F$.

Moreover, $\alpha' = 1$ in $\pi_1(F)/\beta'$, since $\alpha = 1$ in $\pi_1(S)/\beta$ which is an HNN extension of $\pi_1(F)/\beta'$. Similarly, $\beta' = 1$ in $\pi_1(F)/\alpha'$. Using Magnus' original theorem for one-relator groups [7], we see that α' is conjugate in the free group $\pi_1(F)$ to β' or its inverse. Hence α is conjugate in $\pi_1(S)$ to β or its inverse, as claimed. □

The next generalises a result of Dyer and Vasquez [2] for ordinary one-relator groups, and of Papakyriakopoulos [13] for certain one-relator surface groups.

Theorem (3.5). *Let S be an oriented surface and α an essential closed curve in S. Suppose that $\alpha = \beta^m$ in $\pi_1(S)$, with m maximal. Then the space formed by attaching a $K(\mathbb{Z}_m, 1)$-space X_m to S by identifying β with a curve in X_m that generates $\pi_1(X_m)$ is a $K(\pi_1(S)/\alpha, 1)$-space.*

(Note that, in the case $m = 1$ of the theorem, we may take X_m to be a disc, whose boundary is the curve to be identified with $\beta = \alpha$. In other words, the theorem says that the space formed by attaching a 2-cell to S along a non-power essential curve α is aspherical.)

Proof. By Proposition (2.1), there is a surface F with homeomorphic subsurfaces F_0 and F_1, such that S is homotopy equivalent to the double mapping cylinder Y formed from F and $F_0 \times [0,1]$ by identifying $F_0 \times \{0\}$ with $F_0 \subset F$ and $F_0 \times \{1\}$ with $F_1 \subset F$. By the theorem of Dyer and Vasquez [2], $Z := F \cup_\beta X_m$ is aspherical. Since F_0, F_1 and $F_0 \times [0,1]$ are aspherical, and the inclusion-induced maps $F_0 \to Z$, $F_1 \to Z$ are π_1-injective, it follows from a theorem of Whitehead [16] that $Y \cup_\beta X_m$ is aspherical, as claimed. $\square$

Arguing as in [2], we deduce from this an analogue of Lyndon's Identity Theorem [5], and the resulting structure of the (co-) homology of $\pi_1(S)/\alpha$.

Corollary (3.6). *Let S, α, β and m be as in Theorem (3.5). Let $G = \pi_1(S)/\alpha$, let N be the normal closure of α in $\pi_1(S)$, and C the cyclic subgroup of G generated by β (which has order precisely m, by Corollary (3.3)). Then $N/[N,N] \cong \mathbb{Z}G \otimes_{\mathbb{Z}C} \mathbb{Z} \cong \mathbb{Z}(G/C)$ as a (left) $\mathbb{Z}G$-module.*

Proof. Let K denote the $K(G,1)$-space constructed in Theorem (3.5). Then its universal cover $\tilde{K}$ is constructed from the regular cover S_N of S corresponding to N by attaching copies of the universal cover of X_m, one for each left coset of C in G. There is a long exact sequence

$$\cdots \to H_k(S_N) \to H_k(\tilde{K}) \to H_k(\tilde{K}, S_N) \to H_{k-1}(S_N) \to \cdots$$

in which $H_k(\tilde{K}) = 0$ for $k \geq 1$ by Theorem (3.5), and

$$H_k(\tilde{K}, S_N) \cong \mathbb{Z}(G/C) \otimes_{\mathbb{Z}} H_k(\tilde{X}_m, S^1) \cong \mathbb{Z}(G/C) \otimes_{\mathbb{Z}} H_{k-1}(S^1)$$

for $k \geq 2$ since X_m is aspherical. Hence

$$N/[N,N] \cong H_1(S_N) \cong H_2(\tilde{K}, S_N) \cong \mathbb{Z}(G/C) \otimes H_1(S^1) \cong \mathbb{Z}(G/C)$$

as claimed. $\square$

Corollary (3.7). *Let G and C be as in Corollary (3.6), and M a left $\mathbb{Z}G$-module. Then for each $q > 2$ there are isomorphisms $H_q(G,M) \cong H_q(C,M)$ and $H^q(G,M) \cong H^q(C,M)$.*

Combining the above corollary with a theorem of Serre [4] yields further consequences:

Corollary (3.8). *Let G and C be as in Corollary (3.6), and let H be a finite subgroup of G. Then there is a unique double coset HgC such that $H \subseteq gCg^{-1}$.*

Corollary (3.9). *Let G and C be as in Corollary (3.6). Then every element of finite order in G belongs to a conjugate of C.*

This last result generalises a theorem of Magnus, Karrass and Solitar [9] for one-relator groups.

We have not yet addressed the oldest results of one-relator group theory, Magnus' Freiheitssatz [6] and his solution of the word problem [8]. The Freiheitssatz for a one-relator group says that any proper subset of the generators, omitting a letter which essentially occurs in the relator, freely generates a free subgroup. Such subgroups are now known as *Magnus subgroups*.

The word problem is the algorithmic problem of deciding whether any given word in the generators represents the identity element of the group. For one-relator groups a stronger property is true: one can algorithmically decide whether any given word represents an element of the Magnus subgroup generated by any given recursive subset of the generators. This is called the *generalized word problem* for Magnus subgroups. (In the case of a finite presentation, all subsets of the generators are recursive.)

We will prove the analogues of both these results for one-relator surface groups. In general, this will require some more effort than just applying Proposition (2.1). However, there are special cases of both results which can be immediately deduced from Proposition (2.1).

PROPOSITION (3.10). *Let S be a closed oriented surface, α a closed curve in S, and β a simple closed curve in S such that α is not homotopic to a curve disjoint from β, and that $\langle\alpha,\beta\rangle = 0$. Then $\pi_1(S\setminus\beta)\to\pi_1(S)/\alpha$ is injective.*

Proof. This is immediate from the proof of Proposition (2.1), since $S\setminus\beta\cong \mathrm{Int}(S_0)\subseteq F$ (in the notation of (2.1)), and the natural maps $\pi_1(S_0)\to\pi_1(F_0)\to \pi_1(S)/\alpha$ are injective. $\square$

PROPOSITION (3.11). *Let $G = \langle u_1,\dots,u_{2g} \mid [u_1,u_2]\cdots[u_{2g-1},u_{2g}]\rangle$, let W be a word in the generators of G such that u_1 appears in W with exponent-sum zero, let N be the normal closure of W in G and let H be the subgroup of G generated by $\{u_2,\dots,u_{2g}\}$. Then there exists an algorithm which, given a word U in the generators of G, will determine whether or not $U\in NH$; and if so will find the (unique) word V in $\{u_2,\dots,u_{2g}\}$ such that $UV^{-1}\in N$.*

Again this follows more or less immediately from Proposition (2.1), where β is the closed curve representing u_2. We omit the details, since a stronger result will be proved in the next section.

4. Further results

In this section we complete the proofs of the Freiheitssatz and the solution of the generalized word problem for one-relator surface groups. First we prove the Freiheitssatz.

THEOREM (4.1). *Let S be a closed oriented surface, α a closed curve in S, and β a simple closed curve in S such that α is not homotopic to a curve disjoint from β. Then $\pi_1(S\setminus\beta)\to\pi_1(S)/\alpha$ is injective.*

Proof. The result is trivial if S is a torus, so we may assume that S has genus $g\geq 2$. We may also assume that α is not a proper power in $\pi_1(S)$, since if $\pi_1(S\setminus\beta)\to\pi_1(S)/\alpha$ is injective then so is $\pi_1(S\setminus\beta)\to\pi_1(S)/\alpha^m$ for all $m\geq 1$.

We first strengthen the first part of Proposition (2.1) to obtain a simple closed curve γ, disjoint from β, with $\langle\alpha,\gamma\rangle = 0$. Choose a simple closed curve β' that meets β transversely in a single point. Then a regular neighbourhood N of $\beta\cup\beta'$ is a punctured torus, so $S\setminus N$ is a punctured surface of genus $g-1\geq 1$. Take γ to be a simple closed curve in the kernel of the restriction of $\langle\alpha,-\rangle$ to $\pi_1(S\setminus N)$.

Now consider the cover S_K of S corresponding to the kernel K of $\langle -,\gamma\rangle : \pi_1(S)\to\mathbb{Z}$. Let A be a small regular neighbourhood of β in S, such that each component of $A\cap\alpha$ is an embedded arc joining the two components of ∂A. Then A is an annulus. Moreover, $\langle\beta,\gamma\rangle = 0$, so the preimage of A in S_K is the (disjoint) union of an infinite collection of annuli A_n ($n\in\mathbb{Z}$), such that $A_{n+1}=\tau(A_n)$, where τ is a generator of the covering transformation group. Let T denote the preimage in S_K of $S\setminus A$, so that $T = S_K\setminus(\bigcup_{n\in Z} A_n)$.

Since $\langle\alpha,\gamma\rangle = 0$, the preimage of α in S_K is the union of an infinite collection $\{\alpha_n, n\in\mathbb{Z}\}$ of closed curves, where $\alpha_{n+1}=\tau(\alpha_n)$. Now α_0 intersects a nonzero finite number of the A_n. Let λ, μ denote the least and greatest indices n such that $\alpha_0\cap A_n\neq\emptyset$, and assume that α has been isotoped to minimise $\mu-\lambda$.

Define $S_0 = T\cup A_\lambda\cup\cdots\cup A_\mu$, $S_1 = T\cup A_\lambda\cup\cdots\cup A_{\mu-1}$, and $S_2 = T\cup A_{\lambda+1}\cup\cdots\cup A_\mu$. Then $S_1\cong S_2$ via τ, the inclusion-induced maps $\pi_1(S_i)\to\pi_1(S_0)/\alpha_0$ ($i=1,2$) are injective (by Magnus' Freiheitssatz [6]), and $\pi_1(S)/\alpha$ is an HNN-extension of $\pi_1(S_0)/\alpha_0$ with associated subgroups $\pi_1(S_1)$, $\pi_1(S_2)$ and isomorphism $\tau_* : \pi_1(S_1)\to\pi_1(S_2)$.

It follows that $\pi_1(T)\to\pi_1(S)/\alpha$ is injective. Since $\pi_1(T)$ is the kernel of $\langle -,\gamma\rangle : \pi_1(S\setminus\beta)\to\mathbb{Z}$ and since $\langle\alpha,\gamma\rangle = 0$, it also follows that the inclusion-induced map $\pi_1(S\setminus\beta)\to\pi_1(S)/\alpha$ is injective. □

In a similar manner, we can obtain the solution of the generalized word problem for Magnus subgroups in one-relator surface groups.

THEOREM (4.2). *Let $G=\langle u_1,\dots,u_{2g}\mid [u_1,u_2]\cdots[u_{2g-1},u_{2g}]\rangle$, let W be a word in the generators of G, let N be the normal closure of W in G and let H be the subgroup of G generated by $\{u_2,\dots,u_{2g}\}$. Then there exists an algorithm which, given a word U in the generators of G, will determine whether or not $U\in NH$; and if so will find the (unique) word V in $\{u_2,\dots,u_{2g}\}$ such that $UV^{-1}\in N$.*

Proof. We follow the proof of Theorem (4.1), letting α be the closed curve represented by W (up to isotopy), and β the simple closed curve represented by u_2.

In order to find γ we replace the pair of generators $\{u_3,u_4\}$ by another basis $\{u_3',u_4'\}$ of $\langle u_3,u_4\rangle$, such that, on rewriting W in terms of the new generators

$$\{u_1,u_2,u_3',u_4',u_5,\dots,u_{2g}\},$$

the generator u_3' appears with exponent-sum zero. Then we take γ to be the simple closed curve representing u_4'. Note that this process can be carried out algorithmically as follows. The exponent sums of u_3 and u_4 in W give a vector (a,b) in $\mathbb{Z}^2$ that can be transformed to one of the form $(0,k)$ by a matrix in $SL(2,\mathbb{Z})$, where $k=\mathrm{lcm}(a,b)$. The Euclidean algorithm expresses this matrix as a product of elementary matrices. Realising each elementary matrix by a Nielsen transformation produces an automorphism σ of $\langle u_3,u_4\rangle$ such that W, written as a word in $u_1, u_2, \sigma(u_3), \sigma(u_4), u_5, \ldots, u_{2g}$, has exponent sum 0 in $\sigma(u_3)$. Note also that $[\sigma(u_3), \sigma(u_4)]$ is conjugate to $[u_3, u_4]^{\pm 1}$ [11]. Indeed, using the solution of the conjugacy problem in the free group $\langle u_3, u_4\rangle$, we can find a word $w \in \langle u_3, u_4\rangle$ and $\varepsilon = \pm 1$ such that $[\sigma(u_3), \sigma(u_4)] = w[u_3, u_4^{\varepsilon}]w^{-1}$. This gives an (algorithmically obtained) automorphism $\phi : u_3 \mapsto u_3' := w^{-1}\sigma(u_3)w$, $u_4 \mapsto u_4' := w^{-1}\sigma(u_4)^{\varepsilon}w$ of $\langle u_3, u_4\rangle$ such that $\phi([u_3, u_4]) = [u_3, u_4]$. Finally, we extend ϕ to an automorphism of G by setting $\phi(u_i) = u_i$ for $i \neq 3, 4$.

Let us assume that the above algorithmic automorphism has been carried out, so that u_3 appears in W with exponent-sum zero, and we can choose γ to be a curve representing u_4. The homomorphism $\langle -, \gamma\rangle : G \to \mathbb{Z}$ can then be interpreted as the exponent-sum of u_3, and its kernel K is generated by conjugates of u_i ($i \neq 3$) by powers of u_3. Since $W \in K$, we may rewrite W as a word $\tilde{W}$ in these generators. Let λ, μ be the least and greatest indices n respectively such that $u_3^{-n}u_1u_3^n$ occurs in $\tilde{W}$.

Let G_0 be the one-relator group with generators

$$\{u_3^{-n}u_1u_3^n; \lambda \leq n \leq \mu\} \cup \{u_3^{-n}u_ju_3^n; n \in \mathbb{Z}, j = 2, 4, 5, \ldots, 2g\}$$

and relator $\tilde{W}$. The proof of Theorem (4.1) then expresses G/N as an HNN-extension of G_0, in which the associated subgroups are the Magnus subgroups obtained by omitting $u_3^{-\mu}u_1u_3^{\mu}$ and $u_3^{-\lambda}u_1u_3^{\lambda}$ respectively from the generating set. By the solutions of the generalized word problems for one-relator groups and for HNN-extensions, it is decidable whether or not the generator u_1 may be eliminated from U in G/N, as required.

[Note that, while the proof of Theorem (4.1) makes use of the assumption that W is not a proper power in G, the HNN-construction of G/N described there does not depend on that assumption. We may therefore use it in full generality for the purposes of the present proof.] □

The statement of Theorem (4.2) asserts the solubility of the generalized word problem only for one particular Magnus subgroup - that obtained by omitting u_1 from the generating set. A similar argument applies to the Magnus subgroup obtained by omitting any other generator. For an arbitrary Magnus subgroup, one can combine the algorithm of Theorem (4.2) with the solution to the generalized word problem for a free factor of a finitely generated free group. One particular case is the absolute word problem, which is the generalized word problem for the trivial group.

COROLLARY (4.3). *Let $G = \langle u_1, \ldots, u_{2g} \mid [u_1, u_2] \cdots [u_{2g-1}, u_{2g}]\rangle$, let W be a word in the generators of G, and let N be the normal closure of W in G. Then the word problem for G/N is soluble. That is, there exists an algorithm which, given a word U in the generators of G, will determine whether or not $U \in N$.*

Acknowledgements

The author is grateful to the organisers of the FICOFEST conference for the opportunity to participate, and to the Royal Society of Edinburgh for financial support.

Received December 27, 2002

Final version received July 29, 2003

Department of Mathematics
Heriot-Watt University
Edinburgh EH14 4AS
Scotland
J.Howie@hw.ac.uk

References

[1] S. D. Brodskiĭ, *Equations over groups, and groups with one defining relation*, Siberian Math. J. **25** (1984), 235–251.

[2] E. Dyer and A. T. Vasquez, *Some small aspherical spaces*, J. Austral. Math. Soc. **16** (1973), 332–352.

[3] J. Hempel, *One-relator surface groups*, Math. Proc. Cambridge Philos. Soc. **108** (1990), 467–474.

[4] J. Huebschmann, *Cohomology theory of aspherical groups and of small cancellation groups*, J. Pure Appl. Algebra **14** (1979), 137–143.

[5] R. C. Lyndon, *Cohomology theory of groups with a single defining relation*, Ann. Math. (2) **52** (1950), 650–665.

[6] W. Magnus, *Über diskontinuierliche Gruppen mit einer definierenden Relation* (Der Freiheitssatz), J. reine angew. Math. **163** (1930), 141–165.

[7] W. Magnus, *Untersuchungen über einige unendliche diskontinuierliche Gruppen*, Math. Ann. **105** (1931), 52–74.

[8] W. Magnus, *Das Identitätsproblem für Gruppen mit einer definierenden Relation*, Math. Ann. **106** (1932), 295–307.

[9] A. Karrass, W. Magnus and D. Solitar, *Elements of finite order in groups with a single defining relation*, Comm. Pure. Appl. Math. **13** (1960), 57–66.

[10] B. Maskit, *A theorem on planar covering surfaces with applications to* 3*-manifolds*, Ann. Math. (2) **81** (1965), 341–355.

[11] J. Nielsen, *Die Isomorphismen der allgemeinen unendlichen Gruppe mit zwei Erzeugenden*, Math. Ann. **78** (1918), 385–397.

[12] C. D. Papakyriakopoulos, *A reduction of the Poincaré conjecture to group theoretic conjectures*, Ann. Math. (2) **77** (1963), 250–305.

[13] C. D. Papakyriakopoulos, *Attaching* 2*-dimensional cells to a complex*, Ann. Math. (2) **78** (1963), 205–222.

[14] E. S. Rapaport, *Proof of a conjecture of Papakyriakopoulos*, Ann. Math. (2) **79** (1964), 506–513.

[15] C. M. Weinbaum, *On relators and diagrams for groups with one defining relation*, Illinois J. Math. **16** (1972), 308–322.

[16] J. H. C. Whitehead, *On asphericity of regions in a* 3*-sphere*, Fund. Math. **32** (1939), 149–166.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

C_n-MOVES AND THE HOMFLY POLYNOMIALS OF LINKS

TAIZO KANENOBU

Abstract. We consider the difference between the HOMFLY polynomials of two links that are related by a C_n-move. This gives the difference between the first HOMFLY coefficient polynomials of such two knots, and further implies the differences of some finite type invariants between such two knots.

1. Introduction

It is known ([6], [7]) that two knots have the same finite type invariants of order less than n ([28]) if and only if they are related by a finite sequence of C_n-moves. Here a *C_n-move* is a local move for oriented links involving $n+1$ arcs as shown in Figure 1, $n \geq 2$, with a C_1-move a crossing change; in particular, a C_2-move is equivalent to a *delta move* ([16], [19]).

The *HOMFLY polynomial* $P(L; t, z) \in \mathbb{Z}[t^{\pm 1}, z^{\pm 1}]$ is an invariant of the isotopy type of an oriented link L, which is defined, as in [9], by the following formulas:

$$P(U; t, z) = 1; \tag{1.1}$$

$$t^{-1}P(L_+; t, z) - tP(L_-; t, z) = zP(L_0; t, z), \tag{1.2}$$

where U is the unknot and L_+, L_-, L_0 are three links that are identical except inside the depicted regions as shown in Figure 2; see [5], [24]. We call (L_+, L_-, L_0) a *skein triple*; also, we say that L_- (resp. L_+) is obtained from L_+ (resp. L_-) by *changing* the crossing, and that L_0 is obtained from L_+ (or L_-) by *smoothing* the crossing.

A *delta skein quadruple* consists of four links (L, M, L_0, M_0) which are identical except inside the depicted regions as shown in Figure 3; two links L and M are related by a delta move. Nikkuni and the author ([11, Theorem 3.1]) have shown that it holds that:

$$P(L; t, z) - P(M; t, z) = t^2 z^2 \left(P(L_0; t, z) - P(M_0; t, z)\right). \tag{1.3}$$

In this paper, we prove a formula giving the difference of the HOMFLY polynomials of two links that are related by a C_n-move (Theorem (2.7)), which generalizes (1.3) above.

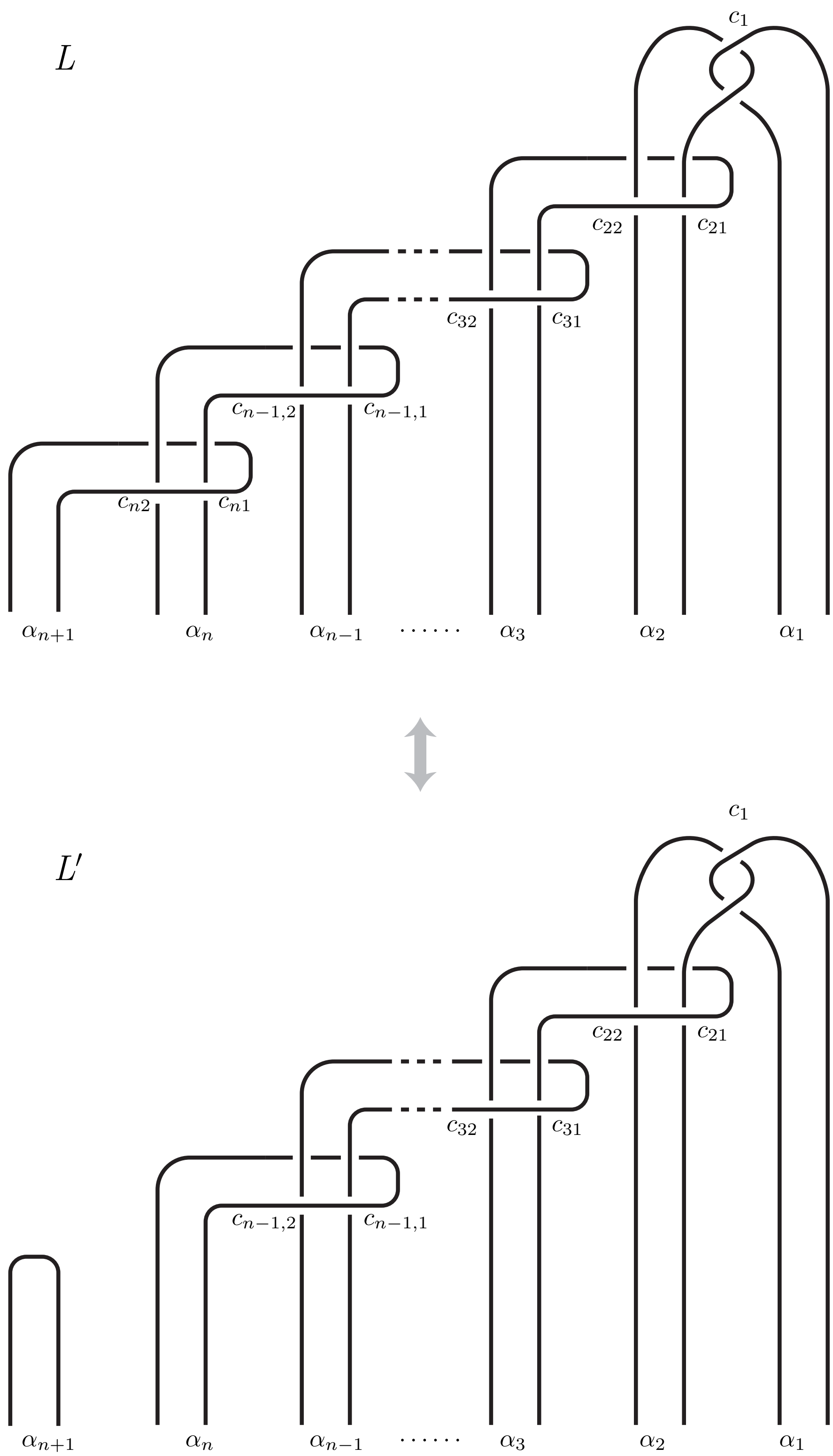
If K is a knot, then its HOMFLY polynomial is of the form:

$$P(K; t, z) = \sum_{i \geq 0} P_{2i}(K; t) z^{2i}, \tag{1.4}$$

2000 *Mathematics Subject Classification*: 57M25.

Keywords and phrases: knot, link, C_n-move, delta move, HOMFLY polynomial, coefficient polynomial, Jones polynomial, Conway polynomial, finite type invariant, Q polynomial.

The author was partially supported by Grant-in-Aid for Scientific Research (B) (No. 14340027), Japan Society for the Promotion of Science.



L
c1
c22
c21
c32
c31
cn−1,2
cn−1,1
cn2
cn1
αn+1
αn
αn−1
······
α3
α2
α1
L′
c1
c22
c21
c32
c31
cn−1,2
cn−1,1
αn+1
αn
αn−1
······
α3
α2
α1

Figure 1

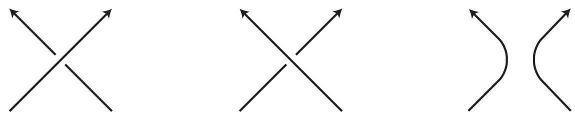


Figure 2

Figure 3

where each $P_{2i}(K;t) \in \mathbb{Z}[t^{\pm 1}]$ is called a *coefficient polynomial*; see [14, Proposition 22]. For a skein triple (L_+, L_-, L_0) with L_+ and L_- knots, and for a delta skein quadruple (L, M, L_0, M_0) with L, M knots, each of (1.2) and (1.3) yields a formula for the first HOMFLY coefficient polynomials (the P_0 polynomials); see (3.3), (3.4). Similarly, for two knots which are related by a C_n-move, from Theorem (2.7), we may obtain a formula for their P_0 polynomials ((3.9), (3.10)). These formulas are much simpler than that in Theorem (2.7). As mentioned above, two knots that are related by a C_n-move have the same finite type invariants of order less than n. On the other hand, there are several results on the difference of the finite type invariants of order n between such two knots (also, such links) ([15], [17], [20], [21], [22], [23], [27]). From (3.9) and (3.10), we obtain that if two knots are related by a C_n-move, then the difference of the nth derivatives of their P_0 polynomials at $t = 1$, which is a finite type invariant of order n, is either 0 or $\pm n! \cdot 2^n$ (Theorem (3.11)). Also, we obtain some results on the Conway polynomials and the constant terms of the Q polynomials (Sect. 4).

2. C_n-moves and the HOMFLY polynomials

In this section, we will prove Theorem (2.7), which gives a formula for the HOMFLY polynomials involving two links that are related by a C_n-move. The proof is essentially analogous to that of [22, Theorem 1.2].

We shall use the following notation: Let L be a link and $C = \{c_1, \ldots, c_k\}$ a subset of the crossings of L. If a link L' is obtained from L by changing the crossings in C_1 and smoothing the crossings in C_2, where C_1, $C_2 \subset C$ and $C_1 \cap C_2 = \emptyset$, then we denote L by $L_{c_1,\ldots,c_k}$ and L' by $L_{c'_1,\ldots,c'_k}$, where

$$c'_i = \begin{cases} \bar{c}_i & \text{if } c_i \in C_1; \\ \dot{c}_i & \text{if } c_i \in C_2; \\ c_i & \text{if } c_i \notin C_1 \cup C_2. \end{cases} \tag{2.1}$$

We will use the following lemma; cf. [22, Proposition 3.1].

Lemma (2.2). *Let L_{c_1,c_2} be a link having two crossings c_1, c_2 with $\epsilon(c_1)\epsilon(c_2) = -1$. Then*

$$P(L_{c_1,c_2}) - P(L_{\bar{c}_1,\bar{c}_2}) = \epsilon(c_1)t^{\epsilon(c_1)}z\left(P(L_{\dot{c}_1,c_2}) - P(L_{\bar{c}_1,\dot{c}_2})\right). \tag{2.3}$$

Proof. Suppose $\epsilon(c_1) = 1$ and $\epsilon(c_2) = -1$. By (1.2), we have

$$t^{-1}P(L_{c_1,c_2}) - tP(L_{\bar{c}_1,c_2}) = zP(L_{\dot{c}_1,c_2}); \tag{2.4}$$

$$t^{-1}P(L_{\bar{c}_1,\bar{c}_2}) - tP(L_{\bar{c}_1,c_2}) = zP(L_{\bar{c}_1,\dot{c}_2}), \tag{2.5}$$

which imply

$$P(L_{c_1,c_2}) - P(L_{\bar{c}_1,\bar{c}_2}) = tz\left(P(L_{\dot{c}_1,c_2}) - P(L_{\bar{c}_1,\dot{c}_2})\right). \tag{2.6}$$

The case $\epsilon(c_1) = -1$ and $\epsilon(c_2) = 1$ is similar, and the proof is complete. □

Theorem (2.7). *Suppose $n \geq 2$. Let L and L' be links such that L' is obtained from L by a C_n-move as shown in Figure 1. Then*

$$P(L) - P(L') = \epsilon_1\epsilon_2\cdots\epsilon_n t^{\epsilon_1+\epsilon_2+\cdots+\epsilon_n} z^n \sum_{\delta_2,\ldots,\delta_n=\pm 1} \delta_2\cdots\delta_n P(L[\delta_2,\ldots,\delta_n]), \tag{2.8}$$

where:

- $\epsilon_1 = \epsilon(c_1)$ *and* $\epsilon_j = \epsilon(c_{j1})$, $j = 2,\ldots,n$;
- $L[\delta_2,\ldots,\delta_n]$ *is the link obtained from* $L = L_{c_1,c_{21},c_{22},\ldots,c_{n1},c_{n2}}$ *by replacing* c_1 *to* $\dot{c}_1$ *and* (c_{j1},c_{j2}) *to either* $(\dot{c}_{j1},c_{j2})$ *or* $(\bar{c}_{j1},\dot{c}_{j2})$ *according as* $\delta_j = 1$ *or* -1, $j = 2,\ldots,n$.

Proof. First notice that $L' = L_{c_1,c_{21},c_{22},\ldots,c_{n-1,1},c_{n-1,2},\bar{c}_{n1},\bar{c}_{n2}}$. We proceed by induction on n. Let us consider the case $n = 2$. Suppose $\epsilon_1 = 1$. Then by (1.2), we have

$$t^{-1}P(L_{c_1}) - tP(L_{\bar{c}_1}) = zP(L_{\dot{c}_1}); \tag{2.9}$$

$$t^{-1}P(L'_{c_1}) - tP(L'_{\bar{c}_1}) = zP(L'_{\dot{c}_1}), \tag{2.10}$$

Since $L_{\bar{c}_1}$ and $L'_{\bar{c}_1}$ are isotopic, we have

$$P(L_{c_1}) - P(L'_{c_1}) = tz\left(P(L_{\dot{c}_1}) - P(L'_{\dot{c}_1})\right). \tag{2.11}$$

Similarly, if $\epsilon_1 = -1$, we have

$$P(L_{c_1}) - P(L'_{c_1}) = -t^{-1}z\left(P(L_{\dot{c}_1}) - P(L'_{\dot{c}_1})\right), \tag{2.12}$$

and so we obtain

$$P(L_{c_1}) - P(L'_{c_1}) = \epsilon_1 t^{\epsilon_1}z\left(P(L_{\dot{c}_1}) - P(L'_{\dot{c}_1})\right). \tag{2.13}$$

As noticed above, $L_{\dot{c}_1} = L_{\dot{c}_1,c_{21},c_{22}}$ and $L'_{\dot{c}_1} = L_{\dot{c}_1,\bar{c}_{21},\bar{c}_{22}}$, and thus by Lemma (2.2), we have

$$\begin{aligned} P(L_{\dot{c}_1}) - P(L'_{\dot{c}_1}) &= \epsilon_2 t^{\epsilon_2}z\left(P(L_{\dot{c}_1,\dot{c}_{21},c_{22}}) - P(L_{\dot{c}_1,\bar{c}_{21},\dot{c}_{22}})\right) \\ &= \epsilon_2 t^{\epsilon_2}z\left(P(L[1]) - P(L[-1])\right). \end{aligned} \tag{2.14}$$

Substituting (2.14) into (2.13), we obtain

$$P(L_{c_1}) - P(L'_{c_1}) = \epsilon_1\epsilon_2 t^{\epsilon_1+\epsilon_2}z^2\left(P(L[1]) - P(L[-1])\right), \tag{2.15}$$

which gives (2.8) with $n = 2$.

Assume that the result holds for $n-1$. By the inductive hypothesis, we have

$$\begin{aligned}P(L) &= P(L_{c_1,c_{21},c_{22},\ldots,c_{n-1,1},c_{n-1,2},c_{n1},c_{n2}}) \\ &= P(L_{c_1,c_{21},c_{22},\ldots,\bar{c}_{n-1,1},\bar{c}_{n-1,2},c_{n1},c_{n2}}) \\ &\quad + at^b z^{n-1} \sum_{\delta_2,\ldots,\delta_{n-1}=\pm 1} \delta_2 \cdots \delta_{n-1} P(L[\delta_2,\ldots,\delta_{n-1}]_{c_{n1},c_{n2}});\end{aligned} \tag{2.16}$$

$$\begin{aligned}P(L') &= P(L_{c_1,c_{21},c_{22},\ldots,c_{n-1,1},c_{n-1,2},\bar{c}_{n1},\bar{c}_{n2}}) \\ &= P(L_{c_1,c_{21},c_{22},\ldots,\bar{c}_{n-1,1},\bar{c}_{n-1,2},\bar{c}_{n1},\bar{c}_{n2}}) \\ &\quad + at^b z^{n-1} \sum_{\delta_2,\ldots,\delta_{n-1}=\pm 1} \delta_2 \cdots \delta_{n-1} P(L[\delta_2,\ldots,\delta_{n-1}]_{\bar{c}_{n1},\bar{c}_{n2}}),\end{aligned} \tag{2.17}$$

where $a = \epsilon_1\epsilon_2\cdots\epsilon_{n-1}$, $b = \epsilon_1+\epsilon_2+\cdots+\epsilon_{n-1}$, and $L[\delta_2,\ldots,\delta_{n-1}]_{c_{n1},c_{n2}}$, $L[\delta_2,\ldots,\delta_{n-1}]_{\bar{c}_{n1},\bar{c}_{n2}}$ are the links obtained from $L(= L_{c_1,c_{21},c_{22},\ldots,c_{n1},c_{n2}})$, $L'(= L_{c_1,c_{21},c_{22},\ldots,\bar{c}_{n1},\bar{c}_{n2}})$, respectively, by replacing c_1 to $\dot{c}_1$ and (c_{j1},c_{j2}) to either $(\dot{c}_{j1},c_{j2})$ or $(\bar{c}_{j1},\dot{c}_{j2})$ according as $\delta_j = 1$ or -1, $j = 2,\ldots,n-1$. Since $L_{c_1,c_{21},c_{22},\ldots,\bar{c}_{n-1,1},\bar{c}_{n-1,2},c_{n1},c_{n2}}$ and $L_{c_1,c_{21},c_{22},\ldots,\bar{c}_{n-1,1},\bar{c}_{n-1,2},\bar{c}_{n1},\bar{c}_{n2}}$ are isotopic, we have

$$\begin{aligned}P(L)-P(L') = at^b z^{n-1} &\sum_{\delta_2,\ldots,\delta_{n-1}=\pm 1} \delta_2\cdots\delta_{n-1} \\ &(P(L[\delta_2,\ldots,\delta_{n-1}]_{c_{n1},c_{n2}}) - P(L[\delta_2,\ldots,\delta_{n-1}]_{\bar{c}_{n1},\bar{c}_{n2}}))\end{aligned} \tag{2.18}$$

By Lemma (2.2), we have

$$\begin{aligned}&P(L[\delta_2,\ldots,\delta_{n-1}]_{c_{n1},c_{n2}}) - P(L[\delta_2,\ldots,\delta_{n-1}]_{\bar{c}_{n1},\bar{c}_{n2}}) \\ &\quad = \epsilon_n t^{\epsilon_n} z \left(P(L[\delta_2,\ldots,\delta_{n-1}]_{\dot{c}_{n1},c_{n2}}) - P(L[\delta_2,\ldots,\delta_{n-1}]_{\bar{c}_{n1},\dot{c}_{n2}})\right) \\ &\quad = \epsilon_n t^{\epsilon_n} z \left(P(L[\delta_2,\ldots,\delta_{n-1},1]) - P(L[\delta_2,\ldots,\delta_{n-1},-1])\right).\end{aligned} \tag{2.19}$$

Substituting (2.19) into (2.18), we obtain (2.8), completing the proof. □

Example (2.20). Let us consider the knot J_n as shown in Figure (4), $n \geq 2$, with J_1 the trivial knot. In particular, J_2 is the right-hand trefoil, J_3 is the mirror image of the 7_6 knot, and J_4 is the mirror image of the 10_{73} knot; see Rolfsen's table ([25]). Performing a C$_n$-move on J_n, we obtain J_{n-1}.

We calculate $P(J_3)-P(J_2)$ using (2.8). We shall use a similar notation to that in Theorem (2.7). Since $\epsilon_1 = \epsilon_2 = \epsilon_3 = 1$, and $J_3[1,1]$ is the positive Hopf link, $J_3[1,-1]$ and $J_3[-1,1]$ are both the trivial links of two components, and $J_3[-1,-1]$ is the trivial link of four components, we have:

$$\begin{aligned}&P(J_3)-P(J_2) = \\ &= t^3z^3\left(P(J_3[1,1]) - P(J_3[1,-1]) - P(J_3[-1,1]) + P(J_3[-1,-1])\right) \\ &= t^3z^3\left(\left(-t^{-1}+t^{-2}\right)z^{-1} - t^{-1}z\right) - 2\left((t^{-1}-t)z^{-1}\right) + (t^{-1}-t)^3z^{-3}\right) \\ &= (1-t^2)^3 + (1-3t^2+2t^4)z^2 - t^2z^4.\end{aligned} \tag{2.21}$$

We may obtain this from the table of [13, p. 282].

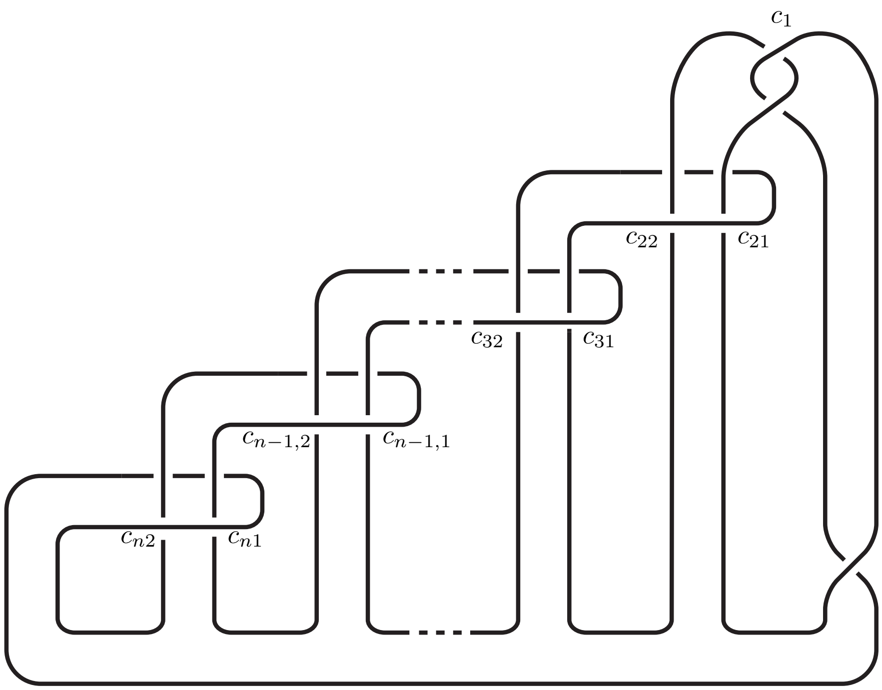


Figure 4

3. C$_n$-moves and the P_0 polynomials

In this section, we will give a formula for the P_0 polynomials of two knots which are related by a C$_n$-move. Before this, we give some properties of the coefficient polynomials of the HOMFLY polynomial of a link. Let $L = K_1 \cup K_2 \cup \cdots \cup K_r$ be an oriented r-component link and $\mathrm{Lk}(L)$ be the total linking number of L; $\mathrm{Lk}(L) = \sum_{i<j} \mathrm{lk}(K_i, K_j)$ with $\mathrm{lk}(K_i, K_j)$ the linking number of K_i and K_j. By [14, Proposition 22], the HOMFLY polynomial of L is of the form:

$$P(L;t,z) = \sum_{n\geq 0} P_{2n-r+1}(L;t) z^{2n-r+1}, \tag{3.1}$$

where each $P_{2n-r+1}(L;t) \in \mathbb{Z}[t^{\pm 1}]$ is called the *coefficient polynomial*; the powers of t which appear in it are all even or odd, depending on whether $2n-r+1$ is even or odd. In particular, the first coefficient polynomial satisfies the following relation:

$$P_{1-r}(L;t) = t^{2\,\mathrm{Lk}(L)} (t^{-1} - t)^{r-1} \prod_{i=1}^{r} P_0(K_i;t). \tag{3.2}$$

For C$_1$- and C$_2$-moves, the following are known: Let (L_+, L_-, L_0) be a skein triple with L_+ and L_- knots and L_0 a 2-component link $K_1 \cup K_2$. Then from (1.2), we obtain:

$$t^{-1} P_0(L_+;t) - t P_0(L_-;t) = t^{2\,\mathrm{Lk}(L_0)} (t^{-1} - t) P_0(K_1;t) P_0(K_2;t), \tag{3.3}$$

where $\mathrm{Lk}(L_0) = \mathrm{lk}(K_1, K_2)$; see [12]. Further, for a delta skein quadruple (L, M, L_0, M_0) with L, M, L_0 knots and M_0 a 3-component link $K_1 \cup K_2 \cup K_3$, we obtain from (1.3):

$$P_0(L;t) - P_0(M;t) = -t^{2\,\mathrm{Lk}(M_0)}(t^2-1)^2 P_0(K_1;t)P_0(K_2;t)P_0(K_3;t); \qquad (3.4)$$

([11, Theorem 4.1]).

In order to give a similar formula for a C_n-move, we consider the number of the components of the link $L[\delta_2,\dots,\delta_n]$, $n \geq 2$, given in Sect. 2, which we denote by $\#L[\delta_2,\dots,\delta_n]$. It is easy to see the following:

$$\#L[\delta_2,\dots,\delta_n] \equiv n + \#L \pmod 2; \qquad (3.5)$$

$$\#L[\delta_2,\dots,\delta_n] \leq n + \#L. \qquad (3.6)$$

Furthermore, we have:

Lemma (3.7). *Suppose that L is a knot. Then in the set of links $\{L[\delta_2,\dots,\delta_n] \mid \delta_i = \pm 1\}$, the number of links with $\#L[\delta_2,\dots,\delta_n] = n+1$ is at most one. In particular, for $n = 2$, one of the links $L[\pm 1]$ is a knot and the other is a 3-component link.*

Proof. We shall use a *chord diagram* of order n, which is an oriented circle with n chords; cf. [1], [2]. For each link $L[\delta_2,\dots,\delta_n]$, $\delta_i = \pm 1$, we construct the chord diagram of order n, $D[\delta_2,\dots,\delta_n]$, as follows: Consider the link L as an image of an embedding $h : S^1 \to S^3$; $L = h(S^1)$. Let $\alpha'_i = h^{-1}(\alpha_i)$, $i = 1, 2, \dots, n+1$. The preimage of each crossing point of L consists of two distinct points. Let $h^{-1}(c_1) = \{c'_1, c''_1\}$, $h^{-1}(c_{ij}) = \{c'_{ij}, c''_{ij}\}$, $i = 2,\dots,n$, $j = 1, 2$, where $c'_1 \in \alpha'_1$, $c''_1, c'_{21}, c'_{22} \in \alpha'_2$, $c''_{i-1,1}, c''_{i-1,2}, c'_{i1}, c'_{i2} \in \alpha'_i$, $i = 3,\dots,n+1$, and let τ_1, τ_{ij} be the chords joining the two points $\{c'_1, c''_1\}$, $\{c'_{ij}, c''_{ij}\}$, respectively. We define $D[\delta_2,\dots,\delta_n]$ to be a chord diagram consisting of the circle S^1 and n chords τ_1, τ_{ij}, $i = 2,\dots,n$, $j = (3-\delta_j)/2$.

First, consider the chords τ_1, τ_{21}, τ_{22}. On the arc α'_2 the point c''_1 lies between c'_{21} and c'_{22}; the other endpoint c'_1 lies on a'_1, and c''_{21}, c''_{22} lie on α'_3. Thus the chord τ_1 intersects either τ_{21} or τ_{22}. Next, consider the chords $\tau_{i-1,1}$, $\tau_{i-1,2}$, $\tau_{i,1}$, $\tau_{i,2}$, $i = 2,\dots,n$. Similarly, on the arc α'_i the four points lie in the order of c'_{i1}, $c''_{i-1,2}$, $c''_{i-1,1}$, c'_{i2}; the other endpoints $c'_{i-1,2}$, $c'_{i-1,1}$ lie on α'_{i-1}, and c''_{i1}, c''_{i2} lie on α'_{i+1}. Thus the chord $\tau_{i,j}$, $j = 1, 2$, intersects either $\tau_{i+1,1}$ or $\tau_{i+1,2}$. Therefore, in the set of all chord diagrams $D[\delta_2,\dots,\delta_n]$, there exists at most one that has no intersection among the chords; see [22, Proof of Lemma 3.2].

Add 1-handles along all the chords of $D[\delta_2,\dots,\delta_n]$, that is, change each chord as in Figure 5. Then we obtain a set of circles whose number is just $\#L[\delta_2,\dots,\delta_n]$, the number of the components of the link $L[\delta_2,\dots,\delta_n]$. Hence, if all n chords of $D[\delta_2,\dots,\delta_n]$ are separated from one another, then the corresponding link $L[\delta_2,\dots,\delta_n]$ has $n+1$ components. Conversely, if there exist a pair of chords that intersect, then the corresponding link $L[\delta_2,\dots,\delta_n]$ has less than $n+1$ components, and so the result follows.

Note that for $n = 2$, one of $D[\pm 1]$ is separated, and so one of the links $L[\pm 1]$ is of 3 components. □

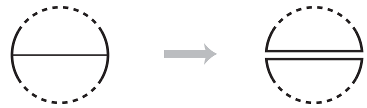


Figure 5

Let L and L' be as in Theorem (2.7). Then from (2.8), we obtain:

$$P_0(L) - P_0(L') = \epsilon_1\epsilon_2\cdots\epsilon_n t^{\epsilon_1+\epsilon_2+\cdots+\epsilon_n} \sum_{\delta_2,\ldots,\delta_n=\pm1} \delta_2\cdots\delta_n P_{-n}(L[\delta_2,\ldots,\delta_n]), \tag{3.8}$$

where $n \geq 2$. Suppose L and L' are knots. By (3.6) and Lemma (3.7), for the links $L[\delta_2,\ldots,\delta_n]$, $\delta_2,\ldots,\delta_n = \pm1$, we have either;

(i) all the links have less than $n+1$ components, or

(ii) only one link, say $L[\delta'_2,\ldots,\delta'_n]$, has $n+1$ components and the other have less than $n+1$ components.

In case (i), by (3.1) we have:

$$P_0(L) - P_0(L') = 0. \tag{3.9}$$

In case (ii), let $L[\delta'_2,\ldots,\delta'_n] = K_1 \cup K_2 \cup \cdots \cup K_{n+1}$ and λ be its total linking number, $\lambda = \sum_{i<j} \mathrm{lk}(K_i, K_j)$. Then using (3.1) and (3.2), we have:

$$\begin{aligned} P_0(L) - P_0(L') &= \epsilon_1\epsilon_2\cdots\epsilon_n\delta'_2\cdots\delta'_n t^{\epsilon_1+\epsilon_2+\cdots+\epsilon_n} P_{-n}(L[\delta'_2,\ldots,\delta'_n]) \qquad (3.10) \\ &= \epsilon_1\epsilon_2\cdots\epsilon_n\delta'_2\cdots\delta'_n t^{\epsilon_1+\epsilon_2+\cdots+\epsilon_n+2\lambda}(t^{-1}-t)^n \prod_{i=1}^{n+1} P_0(K_i). \end{aligned}$$

Moreover, we obtain:

Theorem (3.11). *Let L and L' be knots such that L' is obtained from L by a C_n-move as shown in Figure 1. Then*

$$P_0^{(n)}(L;1) - P_0^{(n)}(L';1) = \begin{cases} \pm 8 & \text{if } n = 2; \\ 0,\ \pm n!\cdot 2^n & \text{if } n \geq 3. \end{cases} \tag{3.12}$$

Conversely, for each value there exist knots L, L' satisfying this formula.

Proof. In case (i), from (3.9) we have $P_0^{(n)}(L;1) - P_0^{(n)}(L';1) = 0$. Note that if $n = 2$, then from Lemma (3.7) this case does not occur; see Remark (3.16) below. In case (ii), (3.10) is written as

$$P_0(L) - P_0(L') = (t-1)^n f(t) \tag{3.13}$$

with

$$f(t) = \epsilon_1\epsilon_2\cdots\epsilon_n\delta'_2\cdots\delta'_n(-1)^n(t+1)^n t^{\epsilon_1+\epsilon_2+\cdots+\epsilon_n+2\lambda-n} \prod_{i=1}^{n+1} P_0(K_i;t). \tag{3.14}$$

Then since $P_0(K_i;1) = 1$, we obtain

$$\begin{aligned} |P_0^{(n)}(L;1) - P_0^{(n)}(L';1)| &= n!\cdot|f(1)| \qquad (3.15) \\ &= n!\cdot 2^n. \end{aligned}$$

Conversely, for any knot there exists a C_n-move that does not change its knot type; see [22, Remark]. Also, in Examples (3.20) and (3.24) below, for any integer $n \geq 2$, we shall give an example of knots L and L' satisfying (3.12). □

Remark (3.16). (i) The case $n = 2$ of Theorem (3.11) is essentially Okada's result ([23]): She has proved that if two knots K and K' are related by a delta move, then

$$a_2(K) - a_2(K') = \pm 1, \tag{3.17}$$

where $a_2(K)$ is the coefficient of z^2 of the Conway polynomial of K. In fact, a delta move is equivalent to a C_2-move (see Remark (3.31) below) and $P_0^{(2)}(K;1) = -8a_2(K)$ ([10, (5.6)]); cf. [11, Remark 4.8].

(ii) The case $n = 3$ of Theorem (3.11) is essentially Tsukamoto's result ([27, Corollary 3.2]): He has proved that if two knots K and K' are related by a clasp-pass move, then $V^{(3)}(K;1) - V^{(3)}(K';1) = 0$ or ± 36, where $V^{(3)}(K;1)$ is the third derivative of the Jones polynomial of K at $t = 1$. In fact, a clasp-pass move is equivalent to a C_3-move and $P_0^{(3)}(K;1) = (4/3)V^{(3)}(K;1)$ ([18]); cf. [10, (5.9)].

(iii) For two knots K, K' that are related by a C_4-move, Matsuzaka [15] has studied the difference $v(K) - v(K')$, where v is a finite type invariant of order 4. More explicitly, he has given this difference in terms of the v-values of certain chord diagrams of order 4 ([15, Theorem 5.1.1]) using the result of Ohyama and Tsukamoto ([22]). From this he has shown:

$$a_4(K) - a_4(K') = 0,\ \pm 2; \tag{3.18}$$

$$V^{(4)}(K;1) - V^{(4)}(K';1) = 0,\ \pm 6 \cdot 4!,\ \pm 12 \cdot 4!. \tag{3.19}$$

Conversely, he has given examples of knots K and K' satisfying these equations. Similarly, from Matsuzaka's theorem, we can easily deduce Theorem (3.11) for $n = 4$.

Example (3.20). We have considered the HOMFLY polynomials of the knots J_n in Example (2.20). Here, we apply (3.10) to J_n. We see $\epsilon_i = 1$ for all i, and $J_n[-1, \ldots, -1]$ is the trivial link of $(n+1)$ components, which is the only $(n+1)$-component link in $J_n[\delta_2, \ldots, \delta_n]$, $\delta_i = \pm 1$. Then from (3.10), we have:

$$\begin{aligned} P_0(J_n) - P_0(J_{n-1}) &= (-1)^{n-1} t^n P_{-n}(J_n[-1, \ldots, -1]) \\ &= (-1)^{n-1} t^n (t^{-1} - t)^n \\ &= -(t^2 - 1)^n, \end{aligned} \tag{3.21}$$

from which, we obtain:

$$P_0^{(n)}(J_n;1) - P_0^{(n)}(J_{n-1};1) = -n! \cdot 2^n; \tag{3.22}$$

$$P_0^{(n+1)}(J_n;1) - P_0^{(n+1)}(J_{n-1};1) = -(n+1)! \cdot 2^{n-1} n. \tag{3.23}$$

Example (3.24). Consider the knot K_n as shown in Figure 6, $n \geq 2$, with K_1 the trivial knot, which is given by H. A. Miyazawa ([17, p. 107]). In particular, K_2 is the figure-eight knot, K_3 is the 7_7 knot, and K_4 is the mirror image of the 10_{60} knot; see [25]. Performing a C_n-move on K_n, we obtain K_{n-1}.

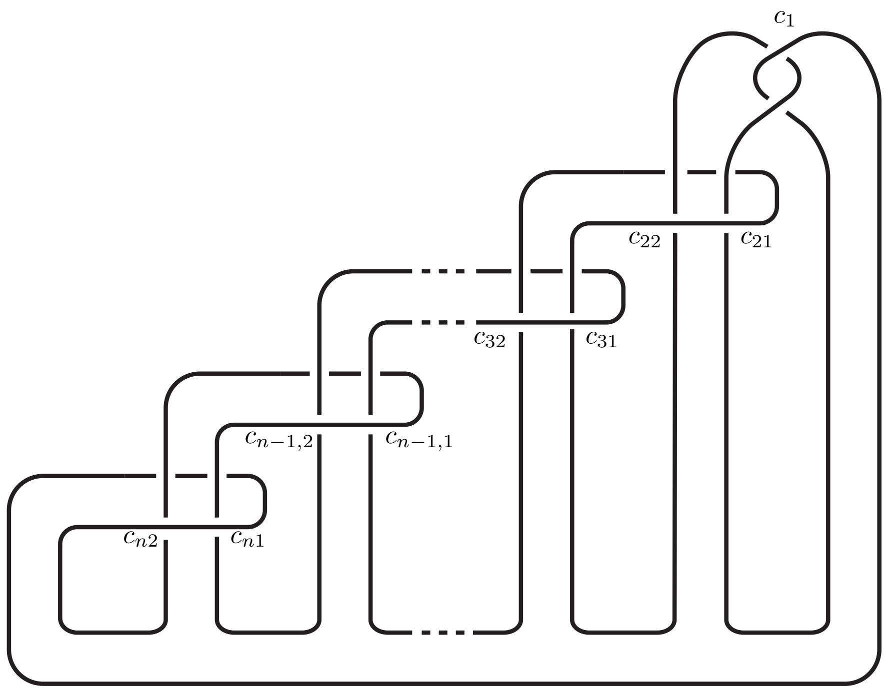


Figure 6

We see $\epsilon_1 = -1$ and $\epsilon_i = 1$ for $2 \leq i \leq n$. By changing the crossing c_1, we obtain the trivial knot U, and by smoothing c_1, we obtain $K_{n-1}\#H_+$, the connected sum of K_{n-1} and a positive Hopf link H_+. From (1.2) we have:

$$t^{-1}P(U) - tP(K_n) = zP(K_{n-1}\#H_+), \tag{3.25}$$

for $n \geq 2$, and so,

$$\begin{aligned} P(K_n) &= t^{-2}P(U) - t^{-1}zP(H_+)P(K_{n-1}) \\ &= t^{-2} - t^{-1}z\left((t-t^3)z^{-1} + tz\right)P(K_{n-1}) \\ &= t^{-2} + \varphi P(K_{n-1}), \end{aligned} \tag{3.26}$$

where $\varphi = (-1+t^2) - z^2$. Then

$$\begin{aligned} P(K_n) - P(K_{n-1}) &= \varphi\left(P(K_{n-1}) - P(K_{n-2})\right) \\ &= \varphi^{n-2}\left(P(K_2) - P(K_1)\right) \\ &= \varphi^{n-2}\psi, \end{aligned} \tag{3.27}$$

where $\psi = (t^{-2} - 2 + t^2) - z^2$; cf. [12, p. 282]. This implies:

$$P_0(K_n) - P_0(K_{n-1}) = t^{-2}(t^2-1)^n = (t-1)^n \cdot t^{-2}(t+1)^n, \tag{3.28}$$

and thus

$$P_0^{(n)}(K_n;1) - P_0^{(n)}(K_{n-1};1) = n! \cdot 2^n; \tag{3.29}$$

$$P_0^{(n+1)}(K_n;1) - P_0^{(n+1)}(K_{n-1};1) = (n+1)! \cdot 2^{n-1}(n-4). \tag{3.30}$$

We may obtain (3.28) using (3.10).

Remark (3.31). Since the delta move and the C$_2$-move are equivalent (cf. [26, Sect. 2]), we can obtain (1.3) from Theorem (2.7): A *tangle* is a disjoint union of finitely many properly embedded arcs in a 3-ball. We may think of the four parts of the link diagrams in Figure 3 as four 3-string tangles, which we denote by τL, τM, τL_0, τM_0 from left to right. If we put these tangles into the 3-ball D in Figure 7, we obtain the four tangles $\tau L'$, $\tau M'$, $\tau L_0'$, $\tau M_0'$ as shown in Figure 8, which correspond to four links L, L', $L[1]$, $L[-1]$, respectively, in Theorem (2.7) with $n = 2$, $\epsilon_1 = \epsilon_2 = 1$. Then from (2.8), we have:

$$P(L) - P(L') = t^2 z^2 \left(P(L[1]) - P(L[-1])\right), \tag{3.32}$$

giving (1.3).

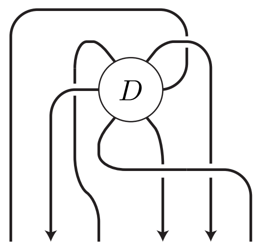
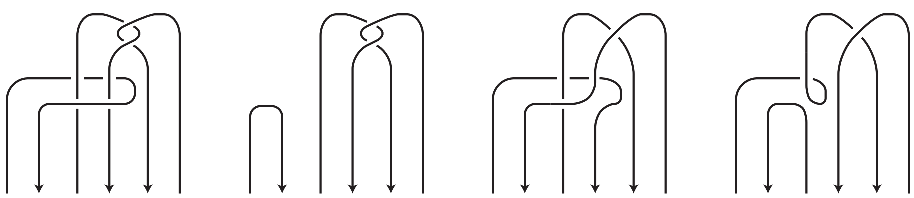


Figure 7



$\tau L'$ $\tau M'$ $\tau L_0'$ $\tau M_0'$

Figure 8

Let us consider the $(n+1)$st derivative of P_0. From (3.14), we have:

$$\begin{aligned} f'(t) =& \epsilon \left(n(t+1)^{n-1} t^m + m(t+1)^n t^{m-1}\right) \prod_{i=1}^{n+1} P_0(K_i; t) \\ &+ \epsilon (t+1)^n t^m \sum_{j=1}^{n+1} P_0'(K_j; t) \prod_{i \neq j} P_0(K_i; t), \end{aligned} \tag{3.33}$$

where $\epsilon = \epsilon_1 \epsilon_2 \cdots \epsilon_n \delta_2' \cdots \delta_n' (-1)^n$ and $m = \epsilon_1 + \epsilon_2 + \cdots + \epsilon_n + 2\lambda - n$. Then since $P_0(K_i; 1) = 1$ and $P_0'(K_i; 1) = 0$, we have:

$$\begin{aligned} f'(1) &= \epsilon (n \cdot 2^{n-1} + m \cdot 2^n) \\ &= \epsilon\ 2^{n-1} (2\epsilon_1 + 2\epsilon_2 + \cdots + 2\epsilon_n + 4\lambda - n). \end{aligned} \tag{3.34}$$

Therefore, we obtain:

$$\begin{aligned}|P_0^{(n+1)}(L;1) - P_0^{(n+1)}(L';1)| &= (n+1)! \cdot |f'(1)| \\ &= (n+1)! \cdot 2^{n-1} \left|2\epsilon_1 + 2\epsilon_2 + \cdots + 2\epsilon_n + 4\lambda - n\right| . \end{aligned} \tag{3.35}$$

which is a generalization of [11, Theorem 4.7, (4.18)].

4. C_n-moves and other polynomials

In this section, we will consider the Conway and Jones polynomials, which we may be obtained from the HOMFLY polynomial, and the constant terms of the Q polynomials.

The *Conway polynomial* $\nabla_L(z) \in \mathbb{Z}[z]$ and the *Jones polynomial* $V(L;t) \in \mathbb{Z}[t^{\pm 1/2}]$ of an oriented link L are given from the HOMFLY polynomial $P(L;t,z)$ by the following formulas; see [4], [9]:

$$\nabla_L(z) = P(L;1,z); \tag{4.1}$$

$$V(L;t) = P(L;t,t^{1/2} - t^{-1/2}). \tag{4.2}$$

The Conway polynomial of an r-component link L is of the form

$$\nabla_L(z) = \sum_{n\geq 0} a_{2n+r-1}(L) z^{2n+r-1}, \tag{4.3}$$

where $a_{2n+r-1}(L) = P_{2n+r-1}(L;1) \in \mathbb{Z}$.

Suppose $n \geq 2$ and let L and L' be links as in Theorem (2.7). Then using (4.1), we obtain immediately the following identity from (2.8):

$$\nabla_L(z) - \nabla_{L'}(z) = \epsilon_1\epsilon_2\cdots\epsilon_n z^n \sum_{\delta_2,\ldots,\delta_n=\pm 1} \delta_2\cdots\delta_n \, \nabla_{L[\delta_2,\ldots,\delta_n]}(z). \tag{4.4}$$

This implies the following identity, which has been given by H. A. Miyazawa [17, p. 102].

$$a_n(L) - a_n(L') = \epsilon_1\epsilon_2\cdots\epsilon_n \sum_{\delta_2,\ldots,\delta_n=\pm 1} \delta_2\cdots\delta_n \, a_0(L[\delta_2,\ldots,\delta_n]). \tag{4.5}$$

Notice that for a link M, $a_0(M) = 1$ or 0 according as if $\#M = 1$ or ≥ 2. For $n \geq 3$, using this formula, she has proved the following identity ([17, Theorem 1.3]):

$$a_n(L) - a_n(L') \equiv 0 \pmod 2. \tag{4.6}$$

For the case where L and L' are knots, this had been proved by Ohyama and Ogushi [21]. Note that if $n \equiv \#L \pmod 2$, then $a_n(L) = a_n(L') = 0$. Also remember Okada's equation (3.17).

Further, combining (4.5) and (4.6), we have the following:

PROPOSITION (4.7). *Suppose $n \geq 3$ and $n \equiv \#L - 1 \pmod 2$. Then*

$$a_n(L) - a_n(L') \in \left\{\pm 2k \mid k = 0, 1, \ldots, 2^{n-3}\right\}. \tag{4.8}$$

Proof. Put $b(L) = \sum_{\delta_2,\ldots,\delta_n=\pm 1} \delta_2\cdots\delta_n \, a_0(L[\delta_2,\ldots,\delta_n])$. If each of the links $L[\delta_2,\ldots,\delta_n]$ with $\delta_2\cdots\delta_n = 1$ (resp. -1) has one component and each of the links $L[\delta_2,\ldots,\delta_n]$ with $\delta_2\cdots\delta_n = -1$ (resp. 1) has more than one component, then $|b(L)| = 2^{n-2}$, otherwise $|b(L)| < 2^{n-2}$; cf. (3.5). Therefore, using (4.6), we obtain the result. □

Remark (4.9). (i) According to Matsuzaka's result (3.18), Proposition (4.7) with $n = 4$ and $\#L = 1$ is not best possible.

(ii) H. A. Miyazawa [17] has given examples of links L and L' such that L' is obtained from L by a C$_n$-move, $n \geq 3$, which satisfies $a_n(L) - a_n(L') = 0, \pm 2$ for each value.

Suppose $n \geq 2$ and let L and L' be links as in Theorem (2.7). Then by using (4.2), we obtain immediately the following identity from (2.8):

$$V(L;t) - V(L';t) = \epsilon_1\epsilon_2\cdots\epsilon_n t^{\epsilon_1+\epsilon_2+\cdots+\epsilon_n - n/2}(t-1)^n \cdot \sum_{\delta_2,\ldots,\delta_n=\pm 1} \delta_2\cdots\delta_n V(L[\delta_2,\ldots,\delta_n];t). \tag{4.10}$$

This implies that:

$$V^{(n)}(L;1) - V^{(n)}(L';1) = \epsilon_1\epsilon_2\cdots\epsilon_n \cdot n! \sum_{\delta_2,\ldots,\delta_n=\pm 1} \delta_2\cdots\delta_n\, V(L[\delta_2,\ldots,\delta_n];1). \tag{4.11}$$

Notice that for a link M, $V(M;1) = (-2)^{\#M-1}$. For $n \geq 3$, H. A. Miyazawa has shown the following identity ([17, Theorem 1.5]):

$$V^{(n)}(L;1) - V^{(n)}(L';1) \equiv 0 \pmod{6 \cdot n!}. \tag{4.12}$$

The Q *polynomial* $Q(M;x) \in \mathbb{Z}[x^{\pm 1}]$ is an invariant of an unoriented link M defined by the following formulas:

$$Q(U;x) = 1; \tag{4.13}$$

$$Q(M_+;x) + Q(M_-;x) = x\left(Q(M_0;x) + Q(M_\infty;x)\right), \tag{4.14}$$

where U is the unknot and $(M_+, M_-, M_0, M_\infty)$ are four links which are identical except inside the depicted regions as illustrated in Figure 9; see [3], [8].

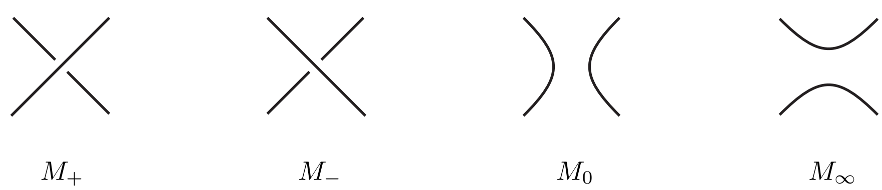


Figure 9

It is known that

$$Q(K;0) = P_0(K;\sqrt{-1}) \equiv 1 \pmod 4 \tag{4.15}$$

for a knot K; see [3, Property 7], [12, Theorem 4.12 (i)].

Suppose $n \geq 2$. Let L and L' be knots such that L' is obtained from L by a C$_n$-move. Then as noticed before Theorem (3.11), there are two cases (i) and (ii) to consider for the links $L[\delta_2,\ldots,\delta_n]$. In case (i), by (3.9) we have:

$$Q(L;0) - Q(L';0) = 0, \tag{4.16}$$

and in case (ii), from (3.10), we obtain:

$$|Q(L;0) - Q(L';0)| = 2^n \prod_{i=1}^{n+1} |Q(K_i;0)|, \tag{4.17}$$

where $K_1 \cup K_2 \cup \cdots \cup K_{n+1}$ is the only $(n+1)$-component link in $L[\delta_2, \ldots, \delta_n]$. This is a generalization of the result for a delta move (C_2-move) in [11, Sect. 4.5].

Example (4.18). For the knots J_n (Example (2.20)) and K_n (Example (2.24)), from (3.21) and (3.28) we obtain:

$$Q(J_n;0) - Q(J_{n-1};0) = P_0(J_n;\sqrt{-1}) - P_0(J_{n-1};\sqrt{-1}) = -(-2)^n; \tag{4.19}$$

$$Q(K_n;0) - Q(K_{n-1};0) = P_0(K_n;\sqrt{-1}) - P_0(K_{n-1};\sqrt{-1}) = -(-2)^n. \tag{4.20}$$

Received April 03, 2003

Final version received March 05, 2004

Department of Mathematics
Osaka City University
Sugimoto, Sumiyoshi-ku
Osaka 558-8585
Japan
kanenobu@sci.osaka-cu.ac.jp

References

[1] D. Bar-Natan, *On the Vassiliev knot invariant*, Topology **34** (1995), 423–472.

[2] J. S. Birman and X.-S. Lin, *Knot polynomials and Vassiliev's invariants*, Invent. Math. **111** (1993), 225–270.

[3] R. D. Brandt, W. B. R. Lickorish and K. C. Millett, *A polynomial invariant for unoriented knots and links*, Invent. Math. **84** (1986), 563–573.

[4] J. H. Conway, *An enumeration of knots and links*, in Computational Problems in Abstract Algebra, (J. Leech, ed.), Pergamon Press, New York, 1969, 329–358.

[5] P. Freyd, D. Yetter, J. Hoste, W. B. R. Lickorish, K. C. Millett and A. Ocneanu, *A new polynomial invariant of knots and links*, Bull. Amer. Math. Soc. **12** (1985), 239–246.

[6] M. Gusarov, *On n-equivalence of knots and invariants of finite degree*, in Topology of Manifolds and Varieties, (O. Viro, ed.), Adv. Soviet Math. **18**, Amer. Math. Soc., Providence, RI (1994), 173–192.

[7] K. Habiro, *Claspers and finite type invariants of links*, Geom. Topol. **4** (2000), 1–83.

[8] C. F. Ho, *A new polynomial for knots and links—preliminary report*, Abstracts Amer. Math. Soc. **6** (1985), 300.

[9] V. F. R. Jones, *Hecke algebra representations of braid groups and link polynomials*, Ann. Math. **126** (1987), 335–388.

[10] T. Kanenobu and Y. Miyazawa, *HOMFLY polynomials as Vassiliev link invariants*, in Knot Theory, Banach Center Publications **42**, (V. F. R. Jones, J. Kania-Bartoszyńska, J. H. Przytycki, P. Traczyk and V. G. Turaev, eds.) Institute of Mathematics, Polish Academy of Science, Warsawa (1998), 165–185.

[11] T. Kanenobu and R. Nikkuni, *Delta move and polynomial invariants of links*, Proceedings of the International Conference on Topology in Matsue 2002 (TIM 2002). Top. Appl. (to appear).

[12] A. Kawauchi, *On coefficient polynomials of the skein polynomial of an oriented link*, Kobe J. Math. **11** (1994), 49–68.

[13] A. Kawauchi, A Survey of Knot Theory, Birkhäuser, Basel, 1996.

[14] W. B. R. Lickorish and K. C. Millett, *A polynomial invariant of oriented links*, Topology **26** (1987), 107–141.

[15] B. Matsuzaka, *C_4-move and Vassiliev invariants of order four for knots*, Master Thesis, International Christian Univ. (1998).

[16] S. Matveev, *Generalized surgeries of three-dimensional manifolds and representations of homology sphere*, (Russian) Mat. Zametki **42** (1987), 268–278, 345 (English translation: Math. Notes **42** (1987), 651–656.)

[17] H. A. Miyazawa, *C_n-moves and polynomial invariants of links*, Kobe. J. Math. **17** (2000), 99–117.

[18] Y. Miyazawa, *The third derivative of the Jones polynomial*, J. Knot Theory Ramifications **6** (1997), 359–372.

[19] H. Murakami and Y. Nakanishi, *On a certain move generating link-homology*, Math. Ann. **284** (1989), 75–89.

[20] Y. Ohyama, *Remarks on C_n-moves for links and Vassiliev invariants of order n*, J. Knot Theory Ramifications **11** (2002), 507–514.

[21] Y. Ohyama and Y. Ogushi, *On the triviality index of knots*, Tokyo J. Math. **13** (1990), 249–255.

[22] Y. Ohyama and T. Tsukamoto, *On Habiro's C_n-moves and Vassiliev invariants of order n*, J. Knot Theory Ramifications **8** (1999), 15–26.

[23] M. Okada, *Delta-unknotting operation and the second coefficient of the Conway polynomial*, J. Math. Soc. Japan **42** (1990), 713–717.

[24] J. H. Przytycki and P. Traczyk, *Invariants of links of Conway type*, Kobe J. Math. **4** (1987), 115–139.

[25] D. Rolfsen, Knots and Links, Publish or Perish, 1976.

[26] K. Taniyama and A. Yasuhara, *Clasp-pass moves on knots, links and spatial graphs*, Top. Appl. **122** (2002), 501–529.

[27] T. Tsukamoto, *Clasp-pass move and Vassiliev invariants of type three for knots*, Proc. Amer. Math. Soc. **128** (1999), 1859–1867.

[28] V. A. Vassiliev, *Cohomology of knot spaces*, in *Theory of Singularities and its Applications*, (V. I. Arnold, ed.) Adv. Soviet Math. **1**, Amer. Math. Soc., Providence, RI (1990), pp. 23–69.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

A TABULATION OF 3-MANIFOLDS VIA DEHN SURGERY

AKIO KAWAUCHI

Abstract. We show that every well-order of the set of lattice points induces an embedding from the set of closed connected orientable 3-manifolds into the set of links which is a right inverse of the 0-surgery map, and this embedding further induces two embeddings from the set of closed connected orientable 3-manifolds into the well-ordered set of lattice points and into the set of link groups. In particular, the set of closed connected orientable 3-manifolds is a well-ordered set by a well-order inherited from the well-ordered set of lattice points, and the homeomorphism problem on the 3-manifolds can in principle be replaced by the isomorphism problem on the link groups. To determine the embedded images of every 3-manifold, we propose a tabulation program on the well-ordered set of 3-manifolds which can be carried out inductively until a concrete pair of indistinguishable 3-manifolds occurs (if there is such a pair). As a demonstration, we tabulate 3-manifolds corresponding to the lattice points of lengths up to 7.

1. Introduction

There are two fundamental problems in the theory of 3-manifolds, that is, the homeomorphism problem and the classification problem (see J. Hempel [11, p.169]). The homeomorphism problem is the problem of giving an effective procedure for determining whether two given 3-manifolds are homeomorphic, and the classification problem is the problem of effectively generating a list containing exactly one 3-manifold from every (unoriented) type of 3-manifolds. In this paper, we consider the classification problem on closed connected orientable 3-manifolds by establishing an embedding from the set of closed connected orientable 3-manifolds into the set of links in the 3-sphere S^3 which is a right inverse of the 0-surgery map. For this purpose, let $\mathbb{Z}$ be the set of integers, and $\mathbb{Z}^n$ the product of n copies of $\mathbb{Z}$ whose elements $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{Z}^n$ we will call *lattice points* of *length* $\ell(\mathbf{x}) = n$. The *set* $\mathbb{X}$ *of lattice points* is the disjoint union of $\mathbb{Z}^n$ for all $n = 1, 2, 3, \ldots$. Let Ω be any well-order in $\mathbb{X}$, although we define in §2 the canonical order[1] Ω_c, a particular well-order in $\mathbb{X}$ such that we have $\mathbf{x} < \mathbf{y}$ for any $\mathbf{x}, \mathbf{y} \in \mathbb{X}$ with $\ell(\mathbf{x}) < \ell(\mathbf{y})$. We are particularly interested in the delta set Δ, a special subset of $\mathbb{X}$ defined in §3 such that the lattice points of Δ smaller than any given $\mathbf{x} \in \mathbb{X}$ in Ω_c form a finite set. The class of oriented links L' in S^3 such that there is a homeomorphism $h : S^3 \to S^3$ sending L to L' is called the *unoriented link type* $[L]$ of an oriented link L in S^3, and the *oriented link type* $\langle L \rangle$ of L if moreover h preserves the orientation of S^3 and the orientations of L

2000 *Mathematics Subject Classification*: 57M25, 57M27.
Keywords and phrases: braid, lattice point, link, 3-manifold, link group, tabulation.

[1]The present definition is modified from the definition made in earlier research announcements to make an enumeration of lattice points easier.

and L'. Let $\mathbb{L}$ and $\vec{\mathbb{L}}$ be the sets of unoriented link types and oriented link types in S^3, respectively. *A link type will be identified with a link belonging to the link type unless confusion might occur.* Thus, $\mathbb{L}$ and $\vec{\mathbb{L}}$ are understood as the sets of unoriented links and oriented links in S^3, respectively. We have a canonical surjection

$$\mathrm{cl}\beta_\iota : \mathbb{X} \xrightarrow{\mathrm{cl}\beta} \vec{\mathbb{L}} \xrightarrow{\iota} \mathbb{L}$$

sending a lattice point to the closure of the associated braid (see §2 for details), where $\iota : \vec{\mathbb{L}} \to \mathbb{L}$ denotes the forgetful surjection, which simply ignores the orientations of S^3 and links. On the other hand, every well-order Ω in $\mathbb{X}$ induces an injection

$$\sigma : \mathbb{L} \longrightarrow \mathbb{X}$$

which is a right inverse of $\mathrm{cl}\beta_\iota$, so that Ω defines a well-order in $\mathbb{L}$, also denoted by Ω. This construction of σ is done in §2. In §3, we show that in the case of $\Omega = \Omega_c$ the image $\sigma(L)$ of a prime link L belongs to Δ. In §4, we define the concept of a π-minimal link (depending on a choice of a well-order Ω in $\mathbb{X}$). Let $\mathbb{L}^\pi$ be the subset of $\mathbb{L}$ consisting of π-minimal links. Then we see that the restriction

$$\sigma|_{\mathbb{L}^\pi} : \mathbb{L}^\pi \longrightarrow \mathbb{X}$$

is an embedding (see Lemma 4.4). Since a π-minimal link is a prime link by definition, we see in the case of $\Omega = \Omega_c$ that $\sigma(\mathbb{L}^\pi) \subset \Delta$ and every initial segment of $\mathbb{L}^\pi$ is a finite set. The *link group* of a link L in S^3 is the fundamental group $\pi_1 E(L)$ of the exterior $E(L) = \mathrm{cl}(S^3 - N(L))$ of L with $N(L)$ a tubular neighborhood of L in S^3. Let $\mathbb{G}$ be the set of the isomorphism types of the link groups for links in $\mathbb{L}$. *The isomorphism type of a group will be identified with a group belonging to the isomorphism type unless confusion might occur.* An *Artin presentation* is a finite group presentation

$$(x_1, x_2, \dots, x_n \,|\, x_i = w_i x_{p(i)} w_i^{-1}, i = 1, 2, \dots, n)$$

where $p(1), p(2), \dots, p(n)$ are a permutation of $1, 2, \dots, n$ and w_i $(i = 1, 2, \dots, n)$ are words in $x_1, x_2, \dots, x_n$ which satisfy the identity

$$\prod_{i=1}^{n} x_i = \prod_{i=1}^{n} w_i x_{p(i)} w_i^{-1}$$

in the free group F on the letters $x_1, x_2, \dots, x_n$. Then we have a braid $b \in B_n$ corresponding to the automorphism φ of F defined by

$$\varphi(x_i) = w_i x_{p(i)} w_i^{-1} \quad (i = 1, 2, \dots, n),$$

from which we see that the set $\mathbb{G}$ is characterized as the set of groups with Artin presentation (see for example [15; p.83] as well as J. S. Birman [2;p.46]). If the closure $\mathrm{cl}(b)$ is prime or π-minimal, then we say that the Artin presentation is *prime* or *π-minimal*, respectively. For the map

$$\pi : \mathbb{L} \longrightarrow \mathbb{G}$$

sending every link L to the link group $\pi_1 E(L)$, we also see that the restriction

$$\pi|_{\mathbb{L}^\pi} : \mathbb{L}^\pi \longrightarrow \mathbb{G}$$

is an embedding (see Lemma (4.4)). Let $\mathbb{M}$ and $\overrightarrow{\mathbb{M}}$ be the sets of unoriented types and oriented types of closed connected oriented 3-manifolds, respectively. *The type of a closed connected oriented 3-manifold will be identified with a 3-manifold belonging to the type unless confusion might occur.* We define the map $\chi_0 : \mathbb{L} \to \mathbb{M}$ by $\chi_0(L) = \chi(L, 0)$, where $\chi(L, 0)$ denotes the 0-surgery manifold of L. The following result is our main theorem, which is proved in §5:

THEOREM (1.1). *Every well-order Ω of $\mathbb{X}$ induces an embedding*

$$\alpha : \mathbb{M} \longrightarrow \mathbb{L}^{\pi} \subset \mathbb{L}$$

and hence two embeddings

$$\begin{aligned} \sigma_\alpha &= \sigma\alpha : \mathbb{M} \longrightarrow \mathbb{X}, \\ \pi_\alpha &= \pi\alpha : \mathbb{M} \longrightarrow \mathbb{G} \end{aligned}$$

which satisfy properties (1) and (2) below:

(1) $\chi_0\alpha = 1$.

(2) *If a lattice point $\sigma_\alpha(M) \in \mathbb{X}$ is given, then the π-minimal link $\alpha(M) \in \mathbb{L}$ with a braid presentation, the 3-manifold $M \in \mathbb{M}$ with a 0-surgery description along a π-minimal link and the link group $\pi_\alpha(M) \in \mathbb{G}$ with a π-minimal Artin presentation are determined.*

Furthermore, when $\Omega = \Omega_c$, we have $\sigma_\alpha(\mathbb{M}) \subset \Delta$ and the properties (3) and (4) below are obtained:

(3) *If a group $\pi_\alpha(M)$ with a prime Artin presentation is given, then the lattice point $\sigma_\alpha(M)$ is determined assuming a solution of the following problem:*

Problem. Let $\mathbf{x} \in \mathbb{X}$ be a lattice point induced from the prime Artin presentation of $\pi_\alpha(M)$, and $\mathbf{x}_1 < \mathbf{x}_2 < \cdots < \mathbf{x}_n$ the lattice points in Δ smaller than or equal to $\mathbf{x}$. Find the smallest index i such that the link $\mathrm{cl}\beta(\mathbf{x}_i)$ is prime and there is an isomorphism $\pi_1 E(\mathrm{cl}\beta(\mathbf{x}_i)) \to \pi_\alpha(M)$.

(4) *If a 3-manifold M with the 0-surgery description along a π-minimal link L is given, then the lattice point $\sigma_\alpha(M)$ is determined assuming a solution of the following problem:*

Problem. Let $\mathbf{x} \in \mathbb{X}$ be a lattice point induced from a π-minimal link L, and $\mathbf{x}_1 < \mathbf{x}_2 < \cdots < \mathbf{x}_n$ the lattice points in Δ smaller than or equal to $\mathbf{x}$. Find the smallest index i such that the link $\mathrm{cl}\beta(\mathbf{x}_i)$ is π-minimal and the 0-surgery manifold $\chi(\mathrm{cl}\beta(\mathbf{x}_i), 0)$ is $\chi(L, 0)$.

The embedding σ_α makes the set $\mathbb{M}$ a well-ordered set by a well-order inherited from the well-order Ω of $\mathbb{L}$ and denoted also by Ω. The *length* of a 3-manifold $M \in \mathbb{M}$ is the length of the lattice point $\sigma_\alpha(M) \in \mathbb{X}$. In §6, to determine the images $\alpha(M)$, $\sigma_\alpha(M)$ and $\pi_\alpha(M)$ of every $M \in \mathbb{M}$, we take the canonical order Ω_c and propose a classification program on $\mathbb{M}$ based on Theorem (1.1), which we can carry out inductively until a concrete pair of indistinguishable 3-manifolds occurs (if there is such a pair). As a demonstration, we carry out this classification for 3-manifolds with lengths up to 7. The embedding π_α implies that *two* 3-manifolds $M_i \in \mathbb{M}$ $(i = 1, 2)$ are homeomorphic if and only if the groups $\pi_\alpha(M_i)$ $(i = 1, 2)$ are isomorphic, and thus the homeomorphism problem on $\mathbb{M}$ can be in principle replaced by the isomorphism problem on $\mathbb{G}$ (see Remark (5.5)), although it appears difficult to calculate the group $\pi_\alpha(M)$ of any given 3-manifold $M \in \mathbb{M}$ apart from the classification program. A lifting of the embedding α to the oriented version is discussed in §7 together with an observation on a relationship between oriented 3-manifold invariants and oriented link invariants.

This paper is a grown up version of a part of the research announcement "*Link corresponding to closed* 3*-manifold* ". A version of the remaining part will appear in [16] (see http://www.sci.osaka-cu.ac.jp/˜kawauchi/index.htm). The author is grateful to Dr. Ikuo Tayama for finding errors from an earlier version of this paper and to the referees for finding further errors and for helpful comments.

2. Representing links in the set of lattice points

For a lattice point $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ of length n, we denote the lattice points $(x_n, \ldots, x_2, x_1)$ and $(|x_1|, |x_2|, \ldots, |x_n|)$ by $\mathbf{x}^T$ and $|\mathbf{x}|$, respectively. Let $|\mathbf{x}|_N$ be a permutation $(|x_{j_1}|, |x_{j_2}|, \ldots, |x_{j_n}|)$ of the coordinates $|x_j|$ $(j = 1, 2, \ldots, n)$ of $|\mathbf{x}|$ such that $|x_{j_1}| \leqq |x_{j_2}| \leqq \cdots \leqq |x_{j_n}|$. For convenience, we use k^n for the lattice point of length n with k for every coordinate and $-k^n$ for $(-k)^n$. The integers $\min_{1\leqq i\leqq n} |x_i|$ and $\max_{1\leqq i\leqq n} |x_i|$ are also denoted by $\min |\mathbf{x}|$ and $\max |\mathbf{x}|$, respectively. Furthermore, we define the *dual* lattice point $\delta(\mathbf{x}) = (x'_1, x'_2, \ldots, x'_n)$ of $\mathbf{x}$ by

$$x'_i = \begin{cases} \operatorname{sign}(x_i)(\max|\mathbf{x}| + 1 - |x_i|) & x_i \neq 0 \\ 0 & x_i = 0\,. \end{cases}$$

Defining $\delta^0(\mathbf{x}) = \mathbf{x}$ and $\delta^n(\mathbf{x}) = \delta(\delta^{n-1}(\mathbf{x}))$ inductively, we note that $\delta^2(\mathbf{x}) \neq \mathbf{x}$ in general, but $\delta^{n+2}(\mathbf{x}) = \delta^n(\mathbf{x})$ for all $n \geqq 1$. For example, taking $\mathbf{x} = (2^3, 3, -2, 3)$, we have $\delta^{2m-1}(\mathbf{x}) = (2^3, 1, -2, 1)$ and $\delta^{2m}(\mathbf{x}) = (1^3, 2, -1, 2)$ for all $m \geqq 1$. For a lattice point $\mathbf{y} = (y_1, y_2, \ldots, y_m)$ of length m, we denote by $(\mathbf{x}, \mathbf{y})$ the lattice point $(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m)$ of length $n + m$. Let $\vec{\mathbb{L}}$ be the set of oriented links. By the Alexander theorem (see J. S. Birman [2]), every oriented link L is represented by the closure $\mathrm{cl}(b)$ of an s-string braid $b \in B_s$ for some $s \geqq 1$. The braiding algorithm of S. Yamada [23] would be useful to deform a link into a closed braid form. Let σ_i $(i = 1, 2, \ldots, s-1)$ be the standard generators of the s-string braid group B_s. By convention, we regard the sign of the crossing point of the diagram σ_i as $+1$. We consider that every braid b in B_s is written as a word on the letters σ_i $(i = 1, 2, \ldots, s-1)$. When b is not written as 1, we write

$$b = \sigma_{i_1}^{\epsilon_1}\sigma_{i_2}^{\epsilon_2}\ldots\sigma_{i_r}^{\epsilon_r}, \quad \epsilon_i = \pm 1 \quad (i = 1, 2, \ldots, r).$$

Then we define the *lattice point* $\mathbf{x}(b)$ *of the braid* b by the identity

$$\mathbf{x}(b) = (\epsilon_1 i_1, \epsilon_2 i_2, \ldots, \epsilon_r i_r) \in \mathbb{Z}^r \subset \mathbb{X}.$$

When b is written as 1, we understand that $\mathbf{x}(b) = 0 \in \mathbb{Z} \subset \mathbb{X}$. For a non-zero lattice point $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{X}$, let x_{i_j} $(j = 1, 2, \ldots, m; i_1 < i_2 < \cdots < i_m)$ be the set of the non-zero integers in the coordinates x_i $(i = 1, 2, \ldots, n)$ of $\mathbf{x}$. Then the lattice point $\tilde{\mathbf{x}} = (x_{i_1}, x_{i_2}, \ldots, x_{i_m})$ is called the *core* of $\mathbf{x}$. When $\mathbf{x}$ is a zero lattice point, we understand the core $\tilde{\mathbf{x}} = 0$. We note that for every non-zero lattice point $\mathbf{x}$, there is a unique braid $b \in B_s$ for every $s \geqq \max|\mathbf{x}| + 1$ such that $\mathbf{x}(b) = \tilde{\mathbf{x}}$. The braid b is called *the associated braid with index* s of $\mathbf{x}$ and is denoted by $\beta^{(s)}(\mathbf{x})$, and in particular, for $s = \max|\mathbf{x}| + 1$, it is called *the associated braid* of $\mathbf{x}$, denoted by $\beta(\mathbf{x})$. The associated braid with index s of any zero lattice point of $\mathbb{X}$ is understood as $1 \in B_s$, and in particular the associated braid as $1 \in B_1$. Taking the closure $\mathrm{cl}\beta(\mathbf{x})$ of the braid $\beta(\mathbf{x})$, we obtain a surjection

$$\mathrm{cl}\beta : \mathbb{X} \longrightarrow \vec{\mathbb{L}}.$$

Then every well-order Ω in $\mathbb{X}$ defines an injection (which is a right inverse of the map $\mathrm{cl}\beta$)

$$\vec{\sigma} : \vec{\mathbb{L}} \longrightarrow \mathbb{X}$$

by sending a link L to the initial element of the subset $\{\mathbf{x} \in \mathbb{X} | \mathrm{cl}\beta(\mathbf{x}) = L\}$ of $\mathbb{X}$ indicated by Ω. By definition, the closed braid $\mathrm{cl}\beta^{(s)}(\mathbf{x})$ with $s > \max|\mathbf{x}| + 1$ is obtained from the closed braid $\mathrm{cl}\beta(\mathbf{x})$ by adding a trivial link of $(s - \max|\mathbf{x}| - 1)$ components. We introduce an equivalence relation $\sim$ in $\mathbb{X}$ as follows:

Definition (2.1). Two lattice points $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{X}$ are related as $\mathbf{x} \sim \mathbf{y}$ if $\mathrm{cl}\beta(\mathbf{x}) = \mathrm{cl}\beta(\mathbf{y})$ in $\vec{\mathbb{L}}$ modulo split additions of trivial links.

Clearly the relation $\sim$ is an equivalence relation in $\mathbb{X}$. Let $\mathbb{X}/\sim$ be the quotient set of $\mathbb{X}$ by $\sim$, and $\langle\mathbf{x}\rangle$ the equivalence class of a lattice point $\mathbf{x} \in \mathbb{X}$ by $\sim$. The quotient map

$$\vec{\sigma}_\sim : \vec{\mathbb{L}} \longrightarrow \mathbb{X}/\sim$$

has the identity $\vec{\sigma}_\sim(\mathrm{cl}(b)) = \langle\mathbf{x}(b)\rangle$ and is a bijection from the quotient set of $\vec{\mathbb{L}}$ modulo split additions of trivial links onto $\mathbb{X}/\sim$. In particular, $\vec{\sigma}_\sim$ is independent of a choice of Ω. We can describe the equivalence relation $\sim$ only in terms of $\mathbb{X}$ by using the braid group presentation and the Markov theorem (see J. S. Birman [2]), as stated in the following lemma:

LEMMA (2.2). *The relations (1)-(6) below are in the equivalence relation $\sim$ in $\mathbb{X}$. Conversely, if we have $\mathbf{x} \sim \mathbf{y}$ in $\mathbb{X}$, then $\mathbf{y}$ is obtained from $\mathbf{x}$ by applying the relations (1)-(6) finitely often.*

1. $(\mathbf{x}, 0) \sim \mathbf{x},\ \mathbf{x} \sim (\mathbf{x}, 0)$ *for all* $\mathbf{x} \in \mathbb{X}$,
2. $(\mathbf{x}, \mathbf{y}, -\mathbf{y}^T) \sim \mathbf{x},\ \mathbf{x} \sim (\mathbf{x}, \mathbf{y}, -\mathbf{y}^T)$ *for all* $\mathbf{x}, \mathbf{y} \in \mathbb{X}$,
3. $(\mathbf{x}, y) \sim \mathbf{x},\ \mathbf{x} \sim (\mathbf{x}, y)$ *for all* $\mathbf{x} \in \mathbb{X}$ *and* $y \in \mathbb{Z}$ *such that* $|y| > \max|\mathbf{x}|$,

4 $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \sim (\mathbf{x}, \mathbf{z}, \mathbf{y})$ *for all* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{X}$ *such that* $\min|\mathbf{y}| > \max|\mathbf{z}| + 1$ *or* $\min|\mathbf{z}| > \max|\mathbf{y}| + 1$,

5. $(\mathbf{x}, \varepsilon y, y+1, y) \sim (\mathbf{x}, y+1, y, \varepsilon(y+1))$ *for all* $\mathbf{x} \in \mathbb{X}$ *and* $y \in \mathbb{Z}$ *such that* $y(y+1) \neq 0$ *and* $\varepsilon = \pm 1$,
6. $(\mathbf{x}, \mathbf{y}) \sim (\mathbf{y}, \mathbf{x})$ *for all* $\mathbf{x}, \mathbf{y} \in \mathbb{X}$.

Proof. The relation (1) is in $\sim$ since $\beta(\mathbf{x}, 0) = \beta(\mathbf{x})$. For (2), we take $\beta^{(s)}(\mathbf{x})$ and $\beta^{(s)}(\mathbf{y})$ in B_s for some s. Then we have

$$\beta^{(s)}(\mathbf{x}, \mathbf{y}, -\mathbf{y}^T) = \beta^{(s)}(\mathbf{x})\beta^{(s)}(\mathbf{y})\beta^{(s)}(\mathbf{y})^{-1} = \beta^{(s)}(\mathbf{x})$$

in B_s and hence

$$\mathrm{cl}\beta(\mathbf{x}, \mathbf{y}, -\mathbf{y}^T) = \mathrm{cl}\beta(\mathbf{x})$$

in $\vec{\mathbb{L}}$ modulo split additions of trivial links, showing that the relation (2) is in $\sim$. For (3), let $s = |y| + 1$. Then by the Markov theorem,

$$\mathrm{cl}\beta(\mathbf{x}, y) = \mathrm{cl}\beta^{(s)}(\mathbf{x})$$

in $\overrightarrow{\mathbb{L}}$ and the last link is equal to $\mathrm{cl}\beta(\mathbf{x})$ modulo split additions of trivial links, showing that the relation (3) is in $\sim$. For (4), we take $\beta^{(s)}(\mathbf{x})$, $\beta^{(s)}(\mathbf{y})$ and $\beta^{(s)}(\mathbf{z})$ in B_s for some s. By the assumption on $\mathbf{y}$ and $\mathbf{z}$, we have

$$\beta^{(s)}(\mathbf{x},\mathbf{y},\mathbf{z}) = \beta^{(s)}(\mathbf{x})\beta^{(s)}(\mathbf{y})\beta^{(s)}(\mathbf{z}) = \beta^{(s)}(\mathbf{x})\beta^{(s)}(\mathbf{z})\beta^{(s)}(\mathbf{y}) = \beta^{(s)}(\mathbf{x},\mathbf{z},\mathbf{y})$$

in B_s which shows that

$$\mathrm{cl}\beta(\mathbf{x},\mathbf{y},\mathbf{z}) = \mathrm{cl}\beta(\mathbf{x},\mathbf{z},\mathbf{y})$$

in $\overrightarrow{\mathbb{L}}$ modulo split additions of trivial links. Thus, the relation (4) is in $\sim$. For (5), consider $\beta^{(s)}(\mathbf{x})$ and $\sigma_j^{\varepsilon'}$ ($j = |y|$, $\varepsilon' = \mathrm{sign}(y)$) in B_s for some s. Let $\varepsilon' = +1$. Then

$$\beta^{(s)}(\mathbf{x},\varepsilon y, y+1, y) = \beta^{(s)}(\mathbf{x})\sigma_j^{\varepsilon}\sigma_{j+1}\sigma_j$$

and the last braid is equal to

$$\beta^{(s)}(\mathbf{x})\sigma_{j+1}\sigma_j\sigma_{j+1}^{\varepsilon} = \beta^{(s)}(\mathbf{x}, y+1, y, \varepsilon(y+1))$$

in B_s by a well-known braid relation. Hence we have

$$\mathrm{cl}\beta(\mathbf{x},\varepsilon y, y+1, y) = \mathrm{cl}\beta(\mathbf{x}, y+1, y, \varepsilon(y+1))$$

in $\overrightarrow{\mathbb{L}}$ modulo split additions of trivial links, showing that the relation (5) is in $\sim$. For $\varepsilon' = -1$, a similar argument gives the desired result since $\mathrm{sign}(y+1) = -1$ by assumption. For (6), let $\beta^{(s)}(\mathbf{x})$ and $\beta^{(s)}(\mathbf{y})$ in B_s for some s. Then we have

$$\mathrm{cl}\beta^{(s)}(\mathbf{x})\beta^{(s)}(\mathbf{y}) = \mathrm{cl}\beta^{(s)}(\mathbf{y})\beta^{(s)}(\mathbf{x})$$

by the Markov theorem and hence

$$\mathrm{cl}\beta(\mathbf{x},\mathbf{y}) = \mathrm{cl}\beta(\mathbf{y},\mathbf{x})$$

in $\overrightarrow{\mathbb{L}}$ modulo split additions of trivial links, showing that the relation (6) is in $\sim$.

Next, we assume $\mathbf{x} \sim \mathbf{y}$. By the relations (1) and (6), we assume $\tilde{\mathbf{x}} = \mathbf{x}$ and $\tilde{\mathbf{y}} = \mathbf{y}$. Let $b = \beta(\mathbf{x})$ and $b' = \beta(\mathbf{y})$ be the associated braids. We show that if $b = b'$ in B_s for an index s, then we can change $\mathbf{x}$ into $\mathbf{y}$ by finitely many applications of the relations (2), (4), (5) and (6). We use the group presentation of B_s with generators σ_i ($i = 1, 2, \ldots, s-1$) and relators

$$\text{(i) } \sigma_i\sigma_j\sigma_i^{-1}\sigma_j^{-1} \ (|i-j| \geqq 2) \quad \text{and} \quad \text{(ii) } \sigma_i\sigma_{i+1}\sigma_i\sigma_{i+1}^{-1}\sigma_i^{-1}\sigma_{i+1}^{-1} \ (1 \leqq i \leqq s-2)$$

(see [2]). Let F be the free group on the letters σ_i ($i = 1, 2, \ldots, s-1$). If $b = b'$ in F, then the solution of the word problem on F guarantees that we can change $\mathbf{x}$ into $\mathbf{y}$ by finitely many applications of the relations (2) and (6). If $b = b'$ in B_s, then the word $b(b')^{-1}$ is written in the form

$$b(b')^{-1} = \prod_{k=1}^{n} R_k^{\varepsilon_k W_k}$$

in F, where $R_k^{\varepsilon_k W_k} = W_k R_k^{\varepsilon_k} W_k^{-1}$ for $\varepsilon_k = \pm 1$ and R_k denotes a relator of the type (i) or (ii) and W_k is a word in F written on the letters σ_i ($i = 1, 2, \ldots, s-1$). Thus, $(\mathbf{x}, -\mathbf{y}^T)$ is changed into

$$\mathbf{a} = (\mathbf{x}(R_1^{\varepsilon_1 W_1}), \mathbf{x}(R_2^{\varepsilon_2 W_2}), \ldots, \mathbf{x}(R_n^{\varepsilon_n W_n}))$$

by finitely many applications of the relations (2) and (6). Since we can change $\mathbf{x}$ into $(\mathbf{x}, -\mathbf{y}^T, \mathbf{y}) = (\mathbf{a}, \mathbf{y})$ by the relation (2), we may consider $b(b')^{-1}b' = \beta(\mathbf{a}, \mathbf{y})$ instead of $b = \beta(\mathbf{x})$. We note that

$$\mathbf{x}(R_k) = (i, j, -i, -j), \quad \mathbf{x}(R_k^{-1}) = (j, i, -j, -i)$$

for the relator (i) and

$$\begin{aligned} \mathbf{x}(R_k) &= (i, i+1, i, -(i+1), -i, -(i+1)), \\ \mathbf{x}(R_k^{-1}) &= (i+1, i, i+1, -i, -(i+1), -i) \end{aligned}$$

for the relator (ii). Since

$$\mathbf{x}(R_k^{\varepsilon_k W_k}) = (\mathbf{x}(W_k), \mathbf{x}(R_k^{\varepsilon_k}), -\mathbf{x}(W_k)^T),$$

we see that $(\mathbf{a}, \mathbf{y})$ is changed into $\mathbf{y}$ by finitely many applications of the relations (2), (4), (5) and (6). Thus, in the case that $b = b'$ in B_s for an index s, we showed that $\mathbf{x}$ can be changed into $\mathbf{y}$ by finitely many applications of the relations (2), (4), (5) and (6).

Now we consider the general case of b and b'. Applying the relation (3) to $\mathbf{x}$ or $\mathbf{y}$, we can assume that $\mathrm{cl}(b) = \mathrm{cl}(b')$ in $\vec{\mathbb{L}}$. Then the Markov theorem says that we have $b = b'$ in B_s with a suitable index s after finitely many applications of the Markov equivalences:

$$\begin{aligned} b_1 b_2 &\leftrightarrow b_2 b_1 \quad (b_1, b_2 \in B_m), \\ b\sigma_m^{\pm 1} &\leftrightarrow b \quad (b \in B_m \subset B_{m+1}) \end{aligned}$$

for any m. This is equivalent to saying that $b = b' \in B_s$ after finitely many applications of the relations (3) and (6) besides the relations (2), (4), (5) and (6) to $\mathbf{x}$ and $\mathbf{y}$. Thus, $\mathbf{x}$ is changed into $\mathbf{y}$ by finitely many applications of the relations (2), (3), (4), (5) and (6). □

Composing the forgetful surjection $\iota : \vec{\mathbb{L}} \to \mathbb{L}$ with the map $\mathrm{cl}\beta$, we obtain a canonical surjection

$$\mathrm{cl}\beta_\iota : \mathbb{X} \to \mathbb{L}$$

and an injection which is a right inverse of $\mathrm{cl}\beta_\iota$

$$\sigma : \mathbb{L} \longrightarrow \mathbb{X}$$

sending an unoriented link L to the initial element of the subset $\{\mathbf{x} | \mathrm{cl}\beta_\iota(\mathbf{x}) = L\}$ of $\mathbb{X}$ indicated by Ω. The *length* of a link $L \in \mathbb{L}$ is the length of the lattice point $\sigma(L)$. By the rule that $L_1 < L_2$ if and only if $\sigma(L_1) < \sigma(L_2)$, a well-order in $\mathbb{L}$ is defined. Since the map σ is induced from Ω, we may say that this well-order in $\mathbb{L}$ is *induced by* Ω, and denote it also by Ω. We also introduce an equivalence relation $\approx$ in $\mathbb{X}$ more relaxed than $\sim$.

Definition (2.3). Two lattice points $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{X}$ are related as $\mathbf{x} \approx \mathbf{y}$ if we have $\mathrm{cl}\beta(\mathbf{x}) = \mathrm{cl}\beta(\mathbf{y})$ in $\mathbb{L}$ modulo split additions of trivial links.

It is straightforward to see that the relation $\approx$ is an equivalence relation in $\mathbb{X}$. The quotient map

$$\sigma_\approx : \mathbb{L} \longrightarrow \mathbb{X}/\approx$$

is independent of a choice of Ω and induces a bijection from the quotient set of $\mathbb{L}$ modulo split additions of trivial links onto $\mathbb{X}/\approx$. For the natural surjection $\mathbb{X}/\sim\ \to\ \mathbb{X}/\approx$, also denoted by ι, we have the following commutative square:

$$\begin{array}{ccc} \vec{\mathbb{L}} & \xrightarrow{\vec{\sigma}_\sim} & \mathbb{X}/\sim \\ {\scriptstyle\iota}\big\downarrow & & \big\downarrow{\scriptstyle\iota} \\ \mathbb{L} & \xrightarrow{\sigma_\approx} & \mathbb{X}/\approx . \end{array}$$

In this diagram, we denote $\iota\langle\mathbf{x}\rangle$ by $[\mathbf{x}]$. Then we have the identity $\sigma_\approx(\mathrm{cl}(b)) = [\mathbf{x}(b)]$. To determine the class $[\mathbf{x}] \in \mathbb{X}/\approx$, it is desirable to describe the equivalence relation $\approx$ only in terms of $\mathbb{X}$. At present, what we can say about $\approx$ is only the following lemma:

LEMMA (2.4). *We have the following (1) and (2):*
1. *For any* $\mathbf{x}, \mathbf{y} \in \mathbb{X}$ *such that* $\mathbf{x} \sim \mathbf{y}$, *we have* $\mathbf{x} \approx \mathbf{y}$.
2. *For all* $\mathbf{x} \in \mathbb{X}$, *we have* $\mathbf{x} \approx \mathbf{x}^T \approx -\mathbf{x} \approx -\mathbf{x}^T$.

Proof. (1) follows directly from the surjection $\iota : \mathbb{X}/\sim\to \mathbb{X}/\approx$. For (2), let $-L$ denote the inverse of an oriented link L, and $\pm\bar{L}$ the mirror image of $\pm L$. Then we have $L = -L = \bar{L} = -\bar{L}$ in $\mathbb{L}$. Taking $L = \mathrm{cl}(b)$ for a braid b, we have

$$\vec{\sigma}_\sim(L) = \langle\mathbf{x}(b)\rangle,\ \vec{\sigma}_\sim(-L) = \langle\mathbf{x}(b)^T\rangle,\ \vec{\sigma}_\sim(\bar{L}) = \langle-\mathbf{x}(b)\rangle,\ \vec{\sigma}_\sim(-\bar{L}) = \langle-\mathbf{x}(b)^T\rangle\,.$$

Then the commutative square preceding to Lemma (2.4) shows (2). ☐

The following remark means that (1) and (2) of Lemma (2.4) are sufficient to characterize the equivalence relation $\approx$ in the set of knots:

Remark (2.5). Let $\mathbb{X}_1$ be the subset of $\mathbb{X}$ consisting of lattice points $\mathbf{x}$ such that $\mathrm{cl}\beta(\mathbf{x})$ is a knot. Then every relation $\mathbf{x} \approx \mathbf{y}$ for $\mathbf{x}, \mathbf{y} \in \mathbb{X}_1$ is generated by the equivalence relation $\sim$ and the relations in (2) of Lemma (2.4). In fact, let $K = \mathrm{cl}\beta(\mathbf{x})$ and $K' = \mathrm{cl}\beta(\mathbf{y})$. If $\mathbf{x} \approx \mathbf{y}$, then we have $[K] = [K']$ modulo split additions of trivial links. Then there is an oriented knot K'', which is one of the knots $\pm K$ or $\pm\bar{K}$, such that $K'' = K'$ in $\vec{\mathbb{L}}$ modulo split additions of trivial links. Thus, we have $\mathbf{z} \sim \mathbf{y}$ for a lattice point $\mathbf{z}$ which is one of $\pm\mathbf{x}$, $\pm\mathbf{x}^T$. More generally, for oriented links L, L' in S^3, we have $L = L'$ in $\mathbb{L}$ modulo split additions of trivial links if and only if we have $L = L'$ in $\vec{\mathbb{L}}$ modulo split additions of trivial links after a suitable choice of orientations of L and S^3. By Lemma (2.4), this implies that in order to know the class $\sigma_\approx(L) \in \mathbb{X}/\approx$ of an oriented link L in S^3 with $r(\geqq 2)$-components K_i $(i = 1, 2, \ldots, r)$, it suffices to know a braid presentation of the link $(-L') \cup (L \setminus L')$ for every sublink L' of L with $1 \leqq \#L' \leqq \frac{r}{2}$ besides a braid presentation of L, where $\#L'$ denotes the number of components of L'.

We now define the *canonical order* Ω_c in $\mathbb{X}$. We define a well-order in $\mathbb{Z}$ by $0 < 1 < -1 < 2 < -2 < 3 < -3 < \ldots$ and extend it to a well-order in $\mathbb{Z}^n$ for every $n \geqq 2$ as follows: Namely, for $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{Z}^n$ we define $\mathbf{x}_1 < \mathbf{x}_2$ if we have one of the following conditions (1)-(3):

(1) $|\mathbf{x}_1|_N < |\mathbf{x}_2|_N$ by the lexicographic order (on the natural number order).

(2) $|\mathbf{x}_1|_N = |\mathbf{x}_2|_N$ and $|\mathbf{x}_1| < |\mathbf{x}_2|$ by the lexicographic order (on the natural number order).

(3) $|\mathbf{x}_1| = |\mathbf{x}_2|$ and $\mathbf{x}_1 < \mathbf{x}_2$ by the lexicographic order on the well-order of $\mathbb{Z}$ defined above.

Finally, for any two lattice points $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{X}$ with $\ell(\mathbf{x}_1) < \ell(\mathbf{x}_2)$, we define $\mathbf{x}_1 < \mathbf{x}_2$.

Then *this order* Ω_c *makes* $\mathbb{X}$ *a well-ordered set.* In fact, let S be any non-empty subset of $\mathbb{X}$. Let S_ℓ be the subset of S consisting of lattice points with the smallest length, say n. Since $\mathbb{Z}^n$ is a well-ordered set as defined above, we can find the initial lattice point of S_ℓ which is the initial lattice point of S by definition. The following lemma is useful in an actual tabulation of prime links.

LEMMA (2.6). *Let L be a link without a splittable component of the trivial knot. Then in the canonical order Ω_c, the lattice point $\sigma(L)$ is the initial element of the equivalence class $[\sigma(L)] \in \mathbb{X}/\approx$. In particular, we have $\mathrm{cl}\beta(\sigma(L)) = L$.*

Proof. Let $\mathbf{x}$ be the initial element of $[\sigma(L)]$. Suppose that $\mathrm{cl}\beta(\mathbf{x})$ has a splittable component of the trivial knot O. If a crossing point of the closed braid diagram $\mathrm{cl}\beta(\mathbf{x})$ is in O, then there is a shorter length lattice point $\mathbf{x}'$ such that $\mathrm{cl}\beta(\mathbf{x}')$ is obtained from the diagram $\mathrm{cl}\beta(\mathbf{x})$ by removing the component O, contradicting the minimality of $\mathbf{x}$. If there are no crossing point in O, then we see from the definition of β that there is a lattice point $\mathbf{x}'$ with $\mathbf{x}' < \mathbf{x}$ such that $\mathrm{cl}\beta(\mathbf{x}')$ is obtained from $\mathrm{cl}\beta(\mathbf{x})$ by removing the component O, contradicting the minimality of $\mathbf{x}$. Thus, we have $\mathrm{cl}\beta(\mathbf{x}) = L$. By definition, we have $\sigma(L) = \mathbf{x}$. $\square$

3. The range of prime links in the canonical order

In this section, we consider $\mathbb{X}$ ordered by the canonical order Ω_c unless otherwise stated. A lattice point $\mathbf{x} \in \mathbb{X}$ is *minimal* if $\mathbf{x}$ is the initial element of the class $[\mathbf{x}]$ in Ω_c. A *prime* link is a link which is neither a splittable link nor a connected sum of two non-trivial links. Let $\mathbb{L}^{\mathrm{p}}$ be the subset of $\mathbb{L}$ consisting of prime links. By Lemma (2.6), the lattice point $\sigma(L)$ is minimal for every prime link L. The following relations are consequences of the relations in Lemma (2.2) and useful in finding minimal lattice points:

LEMMA (3.1).

1. (Duality relation) *For any lattice point $\mathbf{x}$, we have $\mathbf{x} \sim \delta(\mathbf{x})$.*

2. (Flype relation) *For any lattice points $\mathbf{x}$, $\mathbf{y}$ with $\min|\mathbf{x}| \geqq 2$, $\min|\mathbf{y}| \geqq 2$, any integer $m \geqq 1$ and $\varepsilon', \varepsilon = \pm 1$, we have $(\varepsilon^m, \mathbf{x}, \varepsilon', \mathbf{y}) \sim (\varepsilon^m, \mathbf{y}, \varepsilon', \mathbf{x})$.*
For any lattice points $\mathbf{x}, \mathbf{z}$, any integers $m, y \in \mathbb{Z}$ with $m \geqq 1$, $y(y+1) \neq 0$ and $\varepsilon = \pm 1$, we have

$$\begin{aligned} (\mathbf{x}, \varepsilon y^m, y+1, y, \mathbf{z}) &\sim (\mathbf{x}, y+1, y, \varepsilon(y+1)^m, \mathbf{z}), \\ (\mathbf{x}, y, \varepsilon(y+1)^m, -y, \mathbf{z}) &\sim (\mathbf{x}, -(y+1), \varepsilon y^m, y+1, \mathbf{z}) . \end{aligned}$$

Proof. For (1), we note that the lattice point $\delta(\mathbf{x})$ is obtained by changing the usual indices $1, 2, \ldots, m$ of the strings of the associated braid $b = \beta(\mathbf{x})$ into $m, m-1, \ldots, 1$ and then overturning the braid diagram, where $m = \max|\mathbf{x}| + 1$ by definition. Since this deformation does not change the link type of $\mathrm{cl}(b)$ in $\vec{\mathbb{L}}$, we have $\mathbf{x} \sim \delta(\mathbf{x})$ by Definition (2.1). For (2), the closed braid diagrams of the lattice points $(\mathbf{y}, \varepsilon^m, \mathbf{x}, \varepsilon')$ and $(\mathbf{y}, \varepsilon', \mathbf{x}, \varepsilon^m)$ are in the braid-preserving flype relation (see J. S. Birman-W. W. Menasco [3]) [To understand this easier, we number the strings of the closed braid diagram so that the innermost string is labelled 1]. Hence they are related by the relation $\sim$. Since these lattice points are related to $(\varepsilon^m, \mathbf{x}, \varepsilon', \mathbf{y})$ and $(\varepsilon^m, \mathbf{y}, \varepsilon', \mathbf{x})$, respectively, by a relation in Lemma (2.2), the desired relation is obtained. For (3), the first equivalence is proved by induction on m using (5),(6) of Lemma (2.2). The second equivalence follows from (2),(6) of Lemma (2.2) and the first equivalence as follows:

$$\begin{aligned}(\mathbf{x}, y, \varepsilon(y+1)^m, -y, \mathbf{z}) &\sim (\mathbf{x}, -(y+1), y+1, y, \varepsilon(y+1)^m, -y, \mathbf{z})\\ &\sim (\mathbf{x}, -(y+1), \varepsilon y^m, y+1, y, -y, \mathbf{z})\\ &\sim (\mathbf{x}, -(y+1), \varepsilon y^m, y+1, \mathbf{z}).\end{aligned}$$

□

To limit the image $\sigma(\mathbb{L}^{\mathrm{p}}) \subset \mathbb{X}$, we introduce the delta set Δ as follows:

Definition (3.2). The *delta set* Δ is the subset of $\mathbb{X}$ consisting of

$$0 (\in \mathbb{Z}), \quad 1^n (n \geqq 2)$$

and all the lattice points $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ $(n \geqq 4)$ which satisfy all the following conditions (1)-(8):

(1) $x_1 = 1$, $|x_n| \geqq 2$, $n/2 \geqq \max|\mathbf{x}| \geqq 2$ and $\min|\mathbf{x}| \geqq 1$.

(2) For every integer k with $1 < k < \max|\mathbf{x}|$, there is an index i such that $|x_i| = k$.

(3) Every lattice point obtained from $\mathbf{x}$ by permuting the coordinates of $\mathbf{x}$ cyclically is not of the form $(\mathbf{x}', \mathbf{x}'')$ where $1 \leqq \max|\mathbf{x}'| < \min|\mathbf{x}''|$.

(4) If $|x_i| > |x_{i+1}|$, then $|x_i| - 1 = |x_{i+1}|$.

(5) If $|x_i| = |x_{i+1}|$, then $\mathrm{sign}(x_i) = \mathrm{sign}(x_{i+1})$.

(6) If $|(x_i, x_{i+1}, \ldots, x_{i+m+1})| = (k, (k+1)^m, k)$, $(k^m, k+1, k)$ or $(k, k+1, k^m)$ for some $k, m \geqq 1$ and $|x_j| \neq k$ for all $j < i$ and $j > i+m+1$, then $(x_i, x_{i+1}, \ldots, x_{i+m+1})$ is equal to $\pm(k, -\varepsilon(k+1)^m, k)$, $\pm(\varepsilon k^m, -(k+1), k)$ or $\pm(k, -(k+1), \varepsilon k^m)$ for some $\varepsilon = \pm 1$, respectively. Furthermore, if $m = 1$, then we have $\varepsilon = 1$.

(7) If $|(x_i, x_{i+1}, \ldots, x_{i+m+1})|$ is of the form $(k+1, k^m, k+1)$ for some $k, m \geqq 1$, then $(x_i, x_{i+1}, \ldots, x_{i+m+1}) = \pm(k+1, \varepsilon k^m, k+1)$ for some $\varepsilon = \pm 1$. Furthermore if $m = 1$, then we have $\varepsilon = -1$.

(8) $\mathbf{x}$ is the initial element (in the canonical order Ω_c) of the set of the lattice points obtained from every lattice point of $\pm\mathbf{x}$, $\pm\mathbf{x}^T$, $\pm\delta(\mathbf{x})$ and $\pm\delta(\mathbf{x})^T$ by permuting the coordinates cyclically.

See Example (6.2) for some small length lattice points in Δ. It follows directly from the definition of Ω_c that the lattice points in Δ smaller than any given lattice point $\mathbf{x} \in \mathbb{X}$ form a finite set. To analyze the image $\sigma(L) \in \mathbb{X}$ of a prime link $L \in \mathbb{L}^{\mathrm{p}}$, we use the following notion:

Definition (3.3). A lattice point $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is *reducible* if it satisfies one of the following conditions:

1. $\min|\mathbf{x}| = 0$ and $\ell(\mathbf{x}) > 1$.
2. There is an integer k such that $\min|\mathbf{x}| < k < \max|\mathbf{x}|$ and $k \neq |x_i|$ for all i.
3. There is a lattice point of the form $(\mathbf{x}', \mathbf{x}'')$ obtained from $\mathbf{x}$ by permuting the coordinates of $\mathbf{x}$ cyclically where $1 \leqq \max|\mathbf{x}'| < \min|\mathbf{x}''|$.

Otherwise, $\mathbf{x}$ is *irreducible.*

In Definition (3.3), we note the following points: In (1), the core $\tilde{\mathbf{x}}$ of $\mathbf{x}$ has a shorter length. In (2), the link $L = \mathrm{cl}\beta(\mathbf{x})$ is split. In (3), the closed braid diagram $L = \mathrm{cl}\beta(\mathbf{x})$ is a connected sum of two closed braid diagrams. Thus, L is a non-prime link or we have a shorter length lattice point $\mathbf{x}'$ with $\mathbf{x}' \sim \mathbf{x}$.

Since $\min|\mathbf{x}| = 0$ if and only if $\mathbf{x} = 0 \in \mathbb{Z}$ in Δ, we see from (2) and (3) of Definition (3.2) that every lattice point in Δ is irreducible. The following lemma is important to our argument:

LEMMA (3.4). *The lattice point $\sigma(L) \in \mathbb{X}$ of any prime link $L \in \mathbb{L}^p$ belongs to Δ.*

Proof. By Lemma (2.6), $\sigma(L) = \mathbf{x} = (x_1, x_2, \ldots, x_n)$ is a minimal lattice point and $L = \mathrm{cl}\beta(\mathbf{x})$. If $n = 1$, then $\mathbf{x} = 0 \in \Delta$ (and hence L is a trivial knot). In fact, if $\mathbf{x} \neq 0$, then

$$\mathbf{x} \sim (\mathbf{x}, 0) \sim (0, \mathbf{x}) \sim 0$$

by (1),(3) and (6) of Lemma (2.2), contradicting that $\mathbf{x}$ is minimal. Assume that $n > 1$. If $\mathbf{x}$ is reducible, then we see from the remarks following Definition (3.3) that we have a shorter length lattice point $\mathbf{x}'$ with $\mathbf{x}' \sim \mathbf{x}$ because L is a prime link except for the trivial knot, a contradiction. Hence $\mathbf{x}$ is irreducible. By the duality relation, we have $\mathbf{x}' \leqq \mathbf{x}$ with $\mathbf{x}' \sim \mathbf{x}$ and $\min|\mathbf{x}'| = 1$. Since $\mathbf{x}$ is minimal, we have $\mathbf{x}' = \mathbf{x}$ and $\min|\mathbf{x}| = 1$. By Lemmas (2.2) and (2.4), we must have $x_1 = 1$. If $\max|\mathbf{x}| = 1$, then $x_i = 1$ for all i, since otherwise $\mathbf{x}$ has a shorter length lattice point $\mathbf{x}'$ with $\mathbf{x}' \sim \mathbf{x}$, a contradiction. Let $\max|\mathbf{x}| > 1$. We show that $\mathbf{x}$ has the properties (1)-(8) of Definition (3.2). Using that $\mathbf{x}$ is irreducible, we see that $\mathbf{x}$ has (1), (2), (3) except that $|x_n| \geqq 2$. Suppose $|x_n| = 1$. Then by Lemma (2.2), there is a smaller lattice point $\mathbf{x}'$ with $\mathbf{x}' \sim \mathbf{x}$, a contradiction. Thus, the condition $|x_n| \geqq 2$ is also satisfied. If $|x_i| - 1 > |x_{i+1}|$, then the lattice point $\mathbf{x}'$ obtained from $\mathbf{x}$ by interchanging x_i and x_{i+1} has $\mathbf{x}' < \mathbf{x}$ and $\mathbf{x}' \sim \mathbf{x}$ by Lemma (2.2), a contradiction. Hence we have (4). We have also (5) since otherwise $\mathbf{x}$ has a shorter lattice point $\mathbf{x}'$ with $\mathbf{x}' \sim \mathbf{x}$ by Lemma (2.2). For (6), first let $(x_i, x_{i+1}, \ldots, x_{i+m+1}) = (\varepsilon k^m, \varepsilon'(k+1), \varepsilon'' k)$. When $\varepsilon'' = \varepsilon'$, we obtain from (3) of Lemma (3.1)

$$\mathbf{x} \sim \mathbf{x}' = (x'_1, x'_2, \ldots, x'_n)$$

where $(x'_i, x'_{i+1}, \ldots, x'_{i+m+1}) = (\varepsilon' k, \varepsilon' k, \varepsilon(k+1)^m)$ and $x'_j = x_j$ for all $j < i$ and $j > i + m + 1$. Since $|x'_j| \neq k$ for all $j < i$ and $j > i + m + 1$, we see that $\mathbf{x}'$ is reducible, contradicting the minimality of $\mathbf{x}$. Hence $\varepsilon'' = -\varepsilon'$. For $(x_i, x_{i+1}, \ldots, x_{i+m+1}) = (\varepsilon' k, \varepsilon''(k+1), \varepsilon k^m)$ or $(\varepsilon'' k, \varepsilon(k+1)^m, -\varepsilon' k)$, we see that $\varepsilon'' = -\varepsilon'$ by a similar argument using (3) of Lemma (3.1). In particular when $m = 1$, we have also $\varepsilon' = \varepsilon$. Thus, we have (6). For (7), we take $(x_i, x_{i+1}, \ldots, x_{i+m+1}) = (\varepsilon'(k+1), \varepsilon k^m, \varepsilon''(k+1))$. When $\varepsilon'' = -\varepsilon'$, we obtain from (3) of Lemma (3.1)

$$\mathbf{x} \sim \mathbf{x}' = (x'_1, x'_2, \ldots, x'_n)$$

where $(x'_i, x'_{i+1}, \ldots, x'_{i+m+1}) = (-\varepsilon' k, \varepsilon(k+1)^m, \varepsilon' k)$ and $x'_j = x_j$ for all $j < i$ and $j > i + m + 1$. Then $\mathbf{x}' < \mathbf{x}$, a contradiction. Hence $\varepsilon'' = \varepsilon'$. When $m = 1$

and $\varepsilon'' = \varepsilon' = \varepsilon$, we have

$$\mathbf{x} \sim \mathbf{x}' = (x'_1, x'_2, \ldots, x'_n),$$

where $(x'_i, x'_{i+1}, x'_{i+2}) = \varepsilon(k, k+1, k)$ and $x'_j = x_j$ for $j \neq i, i+1, i+2$. Then $\mathbf{x}' < \mathbf{x}$, a contradiction. Hence $\varepsilon' = \varepsilon'' = -\varepsilon$ and we have (7). Since $\mathbf{x}$ is minimal, we have (8). Thus, $\mathbf{x} = \sigma(L)$ is in Δ. □

We see from Lemma (2.6) that the length of a prime link (or more generally, a link without a splittable component of the trivial knot) L in Ω_c is nothing but the minimal crossing number among the crossing numbers of the closed braid diagrams representing L, so that there are only finitely many prime links with the same length. This property also holds for every well-order Ω of $\mathbb{X}$ such that $\ell(\mathbf{x}) < \ell(\mathbf{y})$ means $\mathbf{x} < \mathbf{y}$ for any $\mathbf{x}, \mathbf{y} \in \mathbb{X}$. There are long histories on making a table of knots and links, for example, by C. F. Gauss, T. P. Kirkman, P. G. Tait, C. N. Little, M. G. Haseman, J. W. Alexander-B. G. Briggs, K. Reidemeister for earlier studies (see [15] for references) and by J. H. Conway [5], D. Rolfsen [21], G. H. Dowker-M. B. Thistlethwaite [7], H. H. Doll- M. J. Hoste [6] and Y. Nakagawa [20] for relatively recent studies. In comparison with these tabulations, our tabulation method has three points which may be noted. The first point is that every prime link has a unique expression in canonically ordered lattice points, because $\mathbb{L}^{\mathrm{p}}$ is canonically identified with a subset of the well-ordered set Δ by σ. J. H. Conway's expression in [5] using basic polyhedra and algebraic tangles is excellent for enumerating knots and links together with some global features except for ordering them in a canonical way. C. H. Dowker and M. B. Thistlethwaite in [7] (for knots) and H. H. Doll- M. J. Hoste in [6] (for links) assigned integer sequences to oriented, ordered knot and link diagrams for a tabulation via computer use. As the second point, we can state in the context of their works that we can specify a unique integer sequence among lots of integer sequences representing every prime link, because our method specifies a unique closed braid diagram for every prime link. Using a result of R. W. Ghrist [9], Y. Nakagawa [20] defined an injection ϕ from the set of oriented knots into the set of positive integers so that the value $\phi(K)$ reconstructs K. Then the third point is that we can have a similar result for $\mathbb{L}^{\mathrm{p}}$ by our argument. In fact, in the forthcoming paper [17] (see [18]), we establish an embedding ζ from Δ into the set $\mathbb{Q}_+$ of positive rational numbers so that the value $\zeta(\mathbf{x})$ reconstructs $\mathbf{x}$. Thus, we can identify $\mathbb{L}^{\mathrm{p}}$ with a subset of $\mathbb{Q}_+$ in the sense that the value $\zeta\sigma(L) \in \mathbb{Q}_+$ reconstructs L. In §6, we explain how to make the table of prime links graded by the canonical order Ω_c and, as a demonstration, we make the table for the prime links with lengths up to 7.

4. π-minimal links

Let $K_i (i = 1, 2, \ldots, r)$ be the components of an oriented link L in S^3. A *coloring* f of L is a function

$$f : \{K_i |\, i = 1, 2, \ldots, r\} \longrightarrow \mathbb{Q} \cup \{\infty\}.$$

By a *meridian-longitude system* of L on $N(L)$, we mean a pair of a meridian system $m(L) = \{m_i |\, i = 1, 2, \ldots, r\}$ and a longitude system $\ell(L) = \{\ell_i |\, i = 1, 2, \ldots, r\}$ on $N(L)$ such that (m_i, ℓ_i) is the meridian-longitude pair of K_i on

$N(K_i)$ for every i. We can specify the orientations of $m(L)$ and $\ell(L)$ from those of L and S^3 uniquely. Let $f(K_i) = \frac{a_i}{b_i}$ for coprime integers a_i, b_i for every i where we take $a_i = \pm 1$ and $b_i = 0$ when $f(K_i) = \infty$. Then we have a (unique up to isotopies) simple loop s_i on $\partial N(K_i)$ with $[s_i] = a_i[m_i] + b_i[\ell_i]$ in the first integral homology $H_1(\partial N(K_i))$. We note that if the different choice $f(K_i) = \frac{-a_i}{-b_i}$ is made, then only the orientation of s_i is changed. The *Dehn surgery manifold* of a colored link (L, f) is the oriented 3-manifold

$$\chi(L, f) = E(L) \bigcup_{i=1}^{r} {}_{s_i = 1 \times \partial D_i^2} S^1 \times D_i^2$$

with the orientation induced from $E(L) \subset S^3$, where $\bigcup_{s_i=1\times\partial D_i^2}$ denotes a pasting of $S^1 \times \partial D_i^2$ to $\partial N(K_i)$ so that s_i is identified with $1 \times \partial D_i^2$. In this construction, the 3-manifold $\chi(L, f) \in \mathbb{M}$ is uniquely determined from the colored link (L, f). In this paper, we are particularly interested in the 0-surgery manifolds that are obtained, that is, in $\chi(L, f)$ with $f = 0$. For every link $L \in \mathbb{L}$, we consider the subset

$$\{L\}_\pi = \{L' \in \mathbb{L} |\ \pi_1 E(L') = \pi_1 E(L)\}$$

of $\mathbb{L}$. Here are some examples on $\{L\}_\pi$.

Example (4.1). (1) For every prime knot $K \in \mathbb{L}$, we have $\{K\}_\pi = \{K\}$ by the Gordon-Luecke theorem [10] and W. Whitten [22]. However, for example, if K is the trefoil knot, then $\{K\#K\}_\pi = \{K\#K, K\#\bar{K}\}$, where $\bar{K}$ denotes the mirror image of K.

(2) Let L be the Whitehead link obtained from the Hopf link $O \cup O'$ by replacing O' with the untwisted double D of O': $L = O \cup D$. Furthermore, let L_m be the link obtained by replacing D with the m-full twist D_m of D along O for every $m \in \mathbb{Z}$ where we take $L_0 = L$. Then we have

$$\{L\}_\pi = \{L_m \mid m \in \mathbb{Z}\}.$$

To see this, let $L' \in \{L\}_\pi$. Since $E(L)$ is a hyperbolic 3-manifold and hence $\pi_1 E(L) = \pi_1 E(L')$ means $E(L) = E(L')$ (see W. Jaco [12]), the meridian system on L' indicates a coloring f of L. Since the linking number of O and D is 0, we have $f(O) = \frac{1}{m}$ and $f(D) = \frac{1}{n}$ for some integers $m, n \in \mathbb{Z}$. If m or n is not 0, then we can assume that $m \neq 0$ since the components O and D are interchangeable. If $m \neq 0$, then we obtain L_m by taking m full twists along O. Since any twisted doubled knot K' is non-trivial and $\chi(K', \frac{1}{n}) \neq S^3$ for $n \neq 0$, we must have $n = 0$, giving the desired result. On this example, one may note that since the linking number of L_m is 0, the longitude system of L_m coincides with the longitude system of L in $\partial E(L)$, so that $\chi(L_m, 0) = \chi(L, 0)$ for every m.

We consider $\mathbb{L}$ as a well-ordered set by the well-order Ω (defined from the well-order Ω of $\mathbb{X}$ in §2). The following definition is needed to choose exactly one link in the set $\{L\}_\pi$ for a link $L \in \mathbb{L}$:

Definition (4.2). A link $L \in \mathbb{L}$ is π-*minimal* if L is the initial element of the set $\{L\}_\pi \cap \mathbb{L}^{\mathrm{p}}$ in the well-order Ω.

The following remark gives a reason why we restrict ourselves to a link in S^3:

Remark (4.3). For a certain torus knot $L \in \mathbb{L}$, there are homotopy torus knot spaces E', not the exterior of any knot in S^3, such that $\pi_1(E') = \pi_1 E(L)$ (see J. Hempel [11], p.152).

Let $\mathbb{L}^\pi$ be the subset of $\mathbb{L}$ consisting of π-minimal links. We note that

$$\mathbb{L}^\pi \subset \mathbb{L}^{\mathrm{p}} \subset \mathbb{L}.$$

For the map $\pi : \mathbb{L} \to \mathbb{G}$ sending a link to the link group, we have the following lemma:

LEMMA (4.4). *The restriction* $\pi|_{\mathbb{L}^\pi} : \mathbb{L}^\pi \longrightarrow \mathbb{G}$ *is injective.*

Proof. For $L, L' \in \mathbb{L}^\pi$, assume that $\pi_1 E(L) = \pi_1 E(L')$. Since both L and L' are π-minimal in $\{L\}_\pi = \{L'\}_\pi$, we have $L \leqq L'$ and $L \geqq L'$ by definition. Hence $L = L'$. □

The following question is related to determining when a given prime link is π-minimal:

Question (4.5). For $L, L' \in \mathbb{L}^{\mathrm{p}}$, does $\pi_1 E(L) = \pi_1 E(L')$ mean $E(L) = E(L')$?

The answer to this question is known to be yes for a large class of prime links, including all prime knots by W. Whitten [22], and prime links L such that $E(L)$ does not contain any essential embedded annulus, in particular, hyperbolic links, by the Johannson Theorem (see W. Jaco [12]). Here is another class of links.

PROPOSITION (4.6). *For links $L, L' \in \mathbb{L}$, assume that $E(L)$ is a special Seifert manifold (that is, a Seifert manifold without essential embedded tori) and that there is an isomorphism $\pi_1 E(L) \to \pi_1 E(L')$. Then there is a homeomorphism $E(L) \to E(L')$.*

Proof. By a classification result of G. Burde-K. Murasugi [4], the Seifert structure of $E(L)$ comes from a Seifert structure on S^3. By [12], the orbit surface of the Seifert manifold $E(L)$ is

(i) the disk with at most two exceptional fibers,

(ii) the annulus with at most one exceptional fiber, or

(iii) the disk with two holes and no exceptional fibers.

In particular, $\pi_1 E(L)$, and hence $\pi_1 E(L')$, are groups with non-trivial centers, so that $E(L')$ is also a special Seifert fibered manifold with the same orbit data as $E(L)$. In the case (i), both L and L' are torus knots and $\pi_1 E(L) \cong \pi_1 E(L')$ implies $L = L'$ (confirmed for example by the Alexander polynomials) and hence $E(L) = E(L')$. In the cases of (ii) without exceptional fiber and (iii), we have $E(L) = E(L') = S^1 \times C$ for C the annulus or the disk with two holes. Assume that $E(L)$ and $E(L')$ have, in the case of (ii), one exceptional fiber. Let (p, q) and (r, s) be the types of the exceptional fibers of $E(L)$ and $E(L')$, respectively, where $p, r \geqq 2$, $(p, q) = 1$, $(r, s) = 1$. Let

$$\begin{array}{rcl} \pi_1 E(L) &=& (t, a, b | ta = at, tb = bt, t^q = a^p) \quad \text{and} \\ \pi_1 E(L') &=& (t, a, b | ta = at, tb = bt, t^s = a^r) \end{array}$$

be the fundamental group presentations of $E(L)$ and $E(L')$, respectively, obtained from $S^1 \times C$ with C the disk with two holes by adjoining a fibered solid torus around the exceptional fiber. Let $\psi : \pi_1 E(L) \to \pi_1 E(L')$ be an isomorphism. Considering the central group which is the infinite cyclic group generated by t, we see that $\psi(t) = t^{\pm 1}$. Replacing $-s$ with s if necessary, we may have $\psi(t) = t$. In the quotient groups, ψ induces an isomorphism

$$\psi_* : (a|a^p = 1) * (b|-) \cong (a|a^r = 1) * (b|-)\,.$$

Hence $p = r$ and $\psi(a) = t^m a^\varepsilon$ for some integer m and $\varepsilon = \pm 1$. Then

$$t^q = \psi(a^p) = t^{mp} a^{\varepsilon p} = t^{mp} a^{\varepsilon r} = t^{mp+\varepsilon s}$$

and hence $q \equiv \pm s \pmod p$, which shows the types (p, q) and (r, s) are equivalent. Thus, there is a fiber-preserving homeomorphism $E(L) \to E(L')$. □

Here is a remark on π-minimal links.

Remark (4.7). Let L be the 2-fold connected sum of the Hopf link, and L' the $(3,3)$-torus link. Then we have $\sigma(L) = (1^2, 2^2)$ and $\sigma(L') = (1^2, 2, 1^2, 2)$ in the canonical order Ω_c (cf. §6). Although $E(L) = E(L')$ and $L < L'$, the link L' is a π-minimal link. We note that $\chi(L, 0) = S^1 \times S^2$ and $\chi(L', 0) = P^3$ (the projective 3-space).

5. Proof of Theorem (1.1)

The following lemma is a folklore result obtained by the Kirby calculus (see R. Kirby [19]):

LEMMA (5.1). *The map $\chi_0 : \mathbb{L} \to \mathbb{M}$ defined by $\chi_0(L) = \chi(L, 0)$ is a surjection.*

Proof. For every $M \in \mathbb{M}$, we have a colored link (L, f) with components K_i $(i = 1, 2, \ldots, r)$ such that $\chi(L, f) = M$ and $f(K_i) = m_i$ is an even integer for all i (see S. J. Kaplan [13]). We show that there is a link L_2' with $r + 2$ components such that $\chi(L_2', 0) = \chi(L, f)$. Let $L_2 = L \cup L_H$ be the split union of the oriented link L and an oriented Hopf link $L_H = O_1 \cup O_2$ with linking number $\mathrm{Link}(O_1, O_2) = -1$. Let f_2 be the coloring of L_2 obtained from f and the 0-coloring of L_H. If $m_i \neq 0$, then we take a fusion knot K_i' of K_i and $\frac{|m_i|}{2}$ parallels of $\mathrm{sign}(m_i)O_1$ and one parallel copy of O_2 in the 0-framings. If $m_i = 0$, then we take $K_i' = K_i$. Doing these operations for all i, we obtain from (L_2, f_2) a colored link (L_2', f_2') with $L_2' = (\cup_{i=1}^r K_i') \cup L_H$, a link with $r + 2$ components and a coloring f_2' such that

$$f_2'(K_i') = f_2(K_i) + 2\ \mathrm{Link}(\frac{m_i}{2} O_1, O_2) = m_i - m_i = 0.$$

Since $f_2'|_{L_H} = f_2|_{L_H} = 0$, we have $f_2' = 0$. By the Kirby calculus on handle slides ([19], [15,p.245]), we have $\chi(L_2', 0) = \chi(L_2, f_2) = M$. □

Let $\mathbb{L}^\pi(M)$ be the subset of $\mathbb{L}^\pi$ consisting of π-minimal links L such that $\chi(L, 0) = M$. When we consider a prime link $L \in \mathbb{L}$ with $\chi(L, 0) = M$ to find a π-minimal link in $\mathbb{L}^\pi(M)$ for a given $M \in \mathbb{M}$, the following points should be noted: If we take the initial element L_0 of the set $\{L\}_\pi$, then the link L_0 need not be a prime link, as it is noted in Remark (4.7). If L_0 is the initial element of the prime link subset of $\{L\}_\pi$, then L_0 is a π-minimal link in $\mathbb{L}^\pi(\chi(L_0, 0))$, but in general we cannot guarantee that $\chi(L_0,0)=M$, as we note in the following example:

Remark (5.2). There are hyperbolic links $L,L'\in\mathbb{L}$ such that $E(L)=E(L')$, $\chi(L,0)\neq\chi(L',0)$ and $\{L\}_\pi=\{L'\}_\pi=\{L,L'\}$. Thus, if $L<L'$ in the well-order Ω, then the link L is π-minimal, but L is not in $\mathbb{L}(\chi(L',0))$. To obtain this example, let $L_H=O_1\cup O_2$ be the Hopf link with coloring f such that $f(O_1)=0, f(O_2)=1$. Then $\chi(L_H,f)=S^3$ and the dual colored link (L'_H,f') of (L_H,f) is given by $L'_H=L_H$ and $f'(O_1)=-1$ and $f'(O_2)=0$. By Remark 4.7 of [16], we have a normal imitation $q:(S^3,L^*_H)\to(S^3,L_H)$ with $\chi(L^*_H,fq)=S^3$ and a dual normal imitation $q':(S^3,L'^*_H)\to(S^3,L'_H)$, that is a normal imitation such that $E(L^*_H)=E(L'^*_H)$, $q'|_{E(L^*_H)}=q|_{E(L^*_H)}$ and $(L'^*_H,f'q')$ is the dual colored link of (L^*_H,fq). As it is stated in Remark 4.7 of [16], we can impose on these normal imitations the following additional properties: namely, L^*_H and L'^*_H are totally hyperbolic, componentwise distinct links, and every homeomorphism $h: E(L'')\to E(L^*_H)$ extends to a homeomorphism $h^+:(S^3,L'')\to(S^3,L^*_H)$ or $h^{+\prime}:(S^3,L'')\to(S^3,L'^*_H)$. On the other hand, we see that $\chi(L'_H,0)=S^3$ and the dual colored link (L_H,f'') of $(L'_H,0)$ is given by $f''(O_1)=-1$ and $f'(O_2)=\infty$. Furthermore, we can assume from Theorem 4.1(2) of [16] that $\chi(L^*_H,0)$ and $\chi(L^*_H,f''q)=\chi(L'^*_H,0)$ are distinct because 0 and f'' are distinct from ∞,f. Thus, we can take L^*_H and L'^*_H as L and L', respectively. (We note that $\chi(L^*_H,0)$ and $\chi(L'^*_H,0)$ are homology 3-spheres, because they are normal imitations of $\chi(L_H,0)=\chi(L'_H,0)=S^3$.)

In spite of Example (5.2), we can show the following lemma:

LEMMA (5.3). *For every $M\in\mathbb{M}$, the set $\mathbb{L}^\pi(M)$ is an infinite set.*

Proof. By Lemma (5.1), we can take a disconnected link L in S^3 such that $\chi(L,0)=M$. Let $M\neq S^3$. By a result of [16], there are infinitely many normal imitations

$$q_i:(S^3,L^*_i)\;\longrightarrow\;(S^3,L)\quad(i=1,2,3,\dots)$$

such that

1. $\chi(L^*_i,0)=\chi(L,0)=M$,
2. L^*_i is (totally) hyperbolic, and
3. every homeomorphism $h:E(L^*_i)\to E(L')$ for a link L' in S^3 extends to a homeomorphism $h^+:(S^3,L^*_i)\to(S^3,L')$.

Then L^*_i is π-minimal by (2) and (3), so that $L^*_i\in\mathbb{L}^\pi(M)$, $i=1,2,3,\dots$. For $M=S^3$, let L be a Hopf link. Then $\chi(L,0)=S^3$ and the dual link L' of the Dehn surgery is also the Hopf link. By Remark 4.7 of [16], there are infinitely many pairs of normal imitations

$$\begin{aligned} q_i:(S^3,L^*_i)&\quad\longrightarrow\quad(S^3,L)\,,\\ q'_i:(S^3,L'^*_i)&\quad\longrightarrow\quad(S^3,L')\quad(i=1,2,3,\dots)\end{aligned}$$

such that

1. $\chi(L^*_i,0)=\chi(L,0)=S^3=\chi(L',0)=\chi(L'^*_i,0)$,
2. $E(L^*_i)=E(L'^*_i)$,
3. L^*_i and L'^*_i are (totally) hyperbolic,
4. every homeomorphism $h:E(L^*_i)\to E(L'')$ for a link L'' in S^3 extends to a homeomorphism $h^+:(S^3,L^*_i)\to(S^3,L'')$ or $h'^+:(S^3,L'^*_i)\to(S^3,L'')$.

Thus, $\{L_i^*\}_\pi = \{L_i^*, L_i'^*\}$ for every i, and we can take a π-minimal link, say L_i^* in $\{L_i^*\}_\pi$ for every i, so that $L_i^* \in \mathbb{L}^\pi(S^3)$, $i = 1, 2, 3, \ldots$. □

We are in a position to prove the first half of Theorem (1.1).

Proof of Theorem (1.1). Since $\mathbb{L}^\pi(M) \neq \emptyset$ by Lemma (5.3), we can take the initial element L_M of $\mathbb{L}^\pi(M)$ for every $M \in \mathbb{M}$. Using the fact that the set $\mathbb{L}^\pi(M)$ is uniquely determined by M and Ω, we see that the well-order Ω of $\mathbb{X}$ induces a map

$$\alpha : \mathbb{M} \longrightarrow \mathbb{L}^\pi \subset \mathbb{L}$$

sending a 3-manifold M to the link L_M. This map α must be injective, because the 0-surgery manifold $\chi(\alpha(M), 0) = M$. Combining this result with Lemma (4.4), we obtain the embeddings σ_α and π_α. If a lattice point $\mathbf{x} = \sigma_\alpha(M)$ is given, then we obtain the link $\alpha(M) = \mathrm{cl}\beta(\mathbf{x})$ with braid presentation, the 3-manifold $M = \chi(\mathrm{cl}\beta(\mathbf{x}), 0)$ with 0-surgery description, and the link group $\pi_1 E(\mathrm{cl}\beta(\mathbf{x}))$ with Artin presentation associated with the braid $\beta(\sigma_\alpha(M))$, completing the proof of the first half of the theorem. If a link group $G = \pi_\alpha(M)$ with a prime Artin presentation is given, then we have a braid b such that G is the link group of the prime link $\mathrm{cl}(b)$. Let $\mathbf{x}_i \in \Delta$ $(i = 1, 2, \ldots, n)$ be the lattice points smaller than or equal to the lattice point $\mathbf{x}(b)$. By Lemma (3.4), there is a lattice point $\mathbf{x}_i$ with $\mathbf{x}_i \approx \mathbf{x}(b)$. By using a solution of the problem in (3), let $\mathbf{x}_{i_0}$ be the smallest lattice point such that $\mathrm{cl}\beta(\mathbf{x}_{i_0})$ is a prime link and there is an isomorphism $\pi_1 E(\mathrm{cl}\beta(\mathbf{x}_{i_0})) \to G$ among $\mathbf{x}_i$ $(i = 1, 2, \ldots, n)$. Then the link $\mathrm{cl}\beta(\mathbf{x}_{i_0})$ is π-minimal by this construction. Thus, the desired lattice point $\sigma_\alpha(M) = \mathbf{x}_{i_0}$ is obtained, proving (3). If a π-minimal link L with $\chi(L, 0) = M$ is given, we take a braid b representing L. Let $\mathbf{x}_i \in \Delta$ $(i = 1, 2, \ldots, n)$ be the lattice points smaller than or equal to $\mathbf{x}(b)$. By Lemma (3.4), there is a lattice point $\mathbf{x}_i$ with $\mathbf{x}_i \approx \mathbf{x}(b)$. By using a solution of the problem in (4), we take the smallest lattice point $\mathbf{x}_{i_0}$ such that the link $\mathrm{cl}\beta(\mathbf{x}_{i_0})$ is a π-minimal link and $\chi(\mathrm{cl}\beta(\mathbf{x}_{i_0}), 0) = M$. Thus, the desired lattice point $\sigma_\alpha(M) = \mathbf{x}_{i_0}$ is obtained, proving (4). □

As a matter of fact, we can construct many variants of the embedding $\alpha : \mathbb{M} \to \mathbb{L}$. Here are some remarks on constructing other embeddings α.

Remark (5.4). Let $\mathbb{L}^{\mathrm{h}} \subset \mathbb{L}$ be the subset consisting of hyperbolic links L (possibly with infinite volume) such that L is determined by its exterior $E(L)$ (that is, $E(L) = E(L')$ for a link L' means $L = L'$), and $\mathbb{L}^{\mathrm{h}}(M) = \{L \in \mathbb{L}^{\mathrm{h}} \,|\, \chi(L, 0) = M\}$. Then we still have an embedding $\alpha : \mathbb{M} \to \mathbb{L}^{\mathrm{h}} \subset \mathbb{L}$ with $\chi_0\alpha = 1$ such that σ_α and π_α are embeddings by the proof of Theorem (1.1), using $\mathbb{L}^{\mathrm{h}}(M)$ instead of $\mathbb{L}^\pi(M)$. (For this proof, we use that $\mathbb{L}^{\mathrm{h}}(S^3)$ contains the Hopf link and the set $\mathbb{L}^{\mathrm{h}}(M)$ for $M \neq S^3$ is infinite by Lemma (5.3).) In this case, the links $\alpha(S^1 \times S^2)$, $\alpha(S^3)$ and $\alpha(M)$ for every $M \neq S^1 \times S^2, S^3$ are the trivial knot, the Hopf link, and a hyperbolic link of finite volume, respectively. If we take the subset $\mathbb{L}(M) \subset \mathbb{L}$ consisting of all links L with $\chi(L, 0) = M$, then the proof of Theorem (1.1), using $\mathbb{L}(M)$ instead of $\mathbb{L}^\pi(M)$, shows the existence of an embedding $\alpha : \mathbb{M} \to \mathbb{L}$ with $\chi_0\alpha = 1$. However, in this case, the map π_α is no longer injective in the canonical order Ω_c. In fact, if $K\#K$ is the granny knot and $K\#\bar{K}$ is the square knot, where K is a trefoil knot, then we see that $\alpha(\chi(K\#K,0)) = K\#K$ and $\alpha(\chi(K\#\bar{K},0)) = K\#\bar{K}$. Then we have $\pi_\alpha(\chi(K\#K,0)) = \pi_\alpha(\chi(K\#\bar{K},0))$, although $\chi(K\#K,0) \neq \chi(K\#\bar{K},0)$ (see [14, Example 3.2]).

Remark (5.5). The subsets $\mathbb{L}^{\mathrm{h}}(M) \subset \mathbb{L}^{\pi}(M) \subset \mathbb{L}(M)$ of $\mathbb{L}$ are defined up to automorphisms of M, but the Kirby calculus of [19] enables us to make "automorphism-free" definitions of them. In fact, for a given link L, let $\mathbb{L}(L)$ the set of links L' such that the 0-colored link $(L',0)$ is obtained from the 0-colored link $(L,0)$ or $(\bar{L},0)$ by a finite number of Kirby moves, and then define $\mathbb{L}^{\mathrm{h}}(L)$ and $\mathbb{L}^{\pi}(L)$ to be the restrictions of $\mathbb{L}(L)$ to the hyperbolic links determined by the exteriors and the π-minimal links, respectively. R. Kirby's theorem in [19] shows that for a link L with $\chi(L,0) = M$ we have the identities

$$\mathbb{L}(L) = \mathbb{L}(M), \quad \mathbb{L}^h(L) = \mathbb{L}^h(M) \quad \text{and} \quad \mathbb{L}^{\pi}(L) = \mathbb{L}^{\pi}(M),$$

where the right hand sides are the sets defined before for M. Thus, the embedding α is defined "automorphism-freely". In particular, in any use of $\mathbb{L}^{\mathrm{h}}(M)$ or $\mathbb{L}^{\pi}(M)$, the embedding π_α is defined "automorphism-freely". This is the precise meaning of the statement that the homeomorphism problem on $\mathbb{M}$ can be in principle replaced by the isomorphism problem on $\mathbb{G}$, stated in the introduction.

6. A classification program

In this section we take the canonical order Ω_c unless otherwise stated. We consider the following mutually related three embeddings already established in Theorem (1.1):

$$\begin{aligned} \alpha : \mathbb{M} &\longrightarrow \mathbb{L}, \\ \sigma_\alpha : \mathbb{M} &\longrightarrow \mathbb{X}, \\ \pi_\alpha : \mathbb{M} &\longrightarrow \mathbb{G}. \end{aligned}$$

Since $\sigma_\alpha(\mathbb{M}) \subset \Delta$ and every initial segment of Δ is a finite set, we can attach (without overlapping) to every 3-manifold M in $\mathbb{M}$ a label (n,i) where n denotes the length of M and i denotes that M appears as the ith 3-manifold of length n, so that we have

$$M_{n,1} < M_{n,2} < \cdots < M_{n,m_n}$$

for a positive integer $m_n < \infty$. Let

$$\alpha(M_{n,i}) = L_{n,i} \in \mathbb{L}, \quad \pi_\alpha(M_{n,i}) = G_{n,i} \in \mathbb{G} \quad \text{and} \quad \sigma_\alpha(M_{n,i}) = \mathbf{x}_{n,i} \in \Delta\,.$$

Our classification program is to enumerate the 3-manifolds $M_{n,i}$ for all $n = 1,2,\dots$ and $i = 1,2,\dots,m_n$ together with the data $L_{n,i}, G_{n,i}$ and $\mathbf{x}_{n,i}$, but by (2) of Theorem (1.1) it is sufficient to give the lattice point $\mathbf{x}_{n,i}$, because we can easily construct $L_{n,i}$, $M_{n,i}$ and $G_{n,i}$ by $L_{n,i} = \mathrm{cl}\beta(\mathbf{x}_{n,i})$, $M_{n,i} = \chi(L_{n,i},0)$ and $G_{n,i} = \pi_1 E(L_{n,i})$. We proceed with the argument by induction on the length n of the lattice points. Since the lattice points of lengths $1,2,3$ in Δ are 0, 1^2 and 1^3, we can do the classification of $\mathbb{M}$ with lengths $1,2,3$ as follows (where

$T^2 \times_A S^1$ denotes the torus bundle over S^1 with monodromy matrix A):

length 1: $m_1 = 1,\ M_{1,1} = S^1 \times S^2,\ L_{1,1} = O$ (the trivial knot),
$G_{1,1} = \mathbb{Z},\ \mathbf{x}_{1,1} = 0$.

length 2: $m_2 = 1,\ M_{2,1} = S^3,\ L_{2,1} = 2_1^1$ (the Hopf link),
$G_{2,1} = \mathbb{Z} \oplus \mathbb{Z},\ \mathbf{x}_{2,1} = 1^2$.

length 3: $m_3 = 1,\ M_{3,1} = T^2 \times_A S^1,\ A = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$,
$L_{3,1} = 3_1$ (the trefoil knot), $G_{3,1} = (x, y \,|\, xyx = yxy),\ \mathbf{x}_{3,1} = 1^3$.

To explain our classification of $\mathbb{M}$ with any length $n \geqq 4$, we assume that the classification of $\mathbb{M}$ with lengths $\leqq n-1$ is done. Let Δ_n be the subset of Δ consisting of lattice points of length n. The first step of our classification program is as follows:

Step 1. Make an ordered list $\Delta_n^* \subset \Delta_n$ containing all the minimal lattice points in Δ_n.

If we take a list smaller than Δ_n^*, then our work will be simpler. It is recommended to make first the ordered list $|\Delta_n^*| = \{|\mathbf{x}| \,|\, \mathbf{x} \in \Delta_n^*\}$ taking into account the property of Ω_c that $\mathbf{x} < \mathbf{y}$ if we have one of the following three conditions: (i) $\ell(\mathbf{x}) < \ell(\mathbf{y})$, (ii) $\ell(\mathbf{x}) = \ell(\mathbf{y})$ and $|\mathbf{x}|_N < |\mathbf{y}|_N$, and (iii) $|\mathbf{x}|_N = |\mathbf{y}|_N$ and $|\mathbf{x}| < |\mathbf{y}|$. To establish Step 1, we use the following notion:

Definition (6.1). A lattice point $\mathbf{x} \in \mathbb{X}$ is *locally-minimal* if it is the initial element of the subset of $[\mathbf{x}]$ consisting of the lattice points obtained from $\mathbf{x}$ by the duality relation, the flype relation, and the moves in Lemmas (2.2) and (2.4) except the length-increasing moves.

Every minimal lattice point is locally-minimal, but the converse is not true. It is realistic to take as Δ_n^* a list containing all the locally-minimal lattice points of Δ_n. The following list is such a list for Step 1.

Example (6.2). The following list contains all the minimal lattice points of lengths $\leqq 7$ in Δ:

Δ_1^*: 0,

Δ_2^*: 1^2,

Δ_3^*: 1^3,

Δ_4^*: $1^4, (1,-2,1,-2)$,

Δ_5^*: $1^5, (1^2,2,-1,2), (1^2,-2,1,-2)$,

Δ_6^*: $1^6, (1^3,2,-1,2), (1^3,-2,1,-2), (1^2,2,1^2,2)$,
$(1^2,2,(-1)^2,2), (1^2,-2,1^2,-2), (1^2,-2,1,(-2)^2)$,
$(1,-2,1,-2,1,-2), (1,-2,1,3,-2,3)$,

Δ_7^*: $1^7, (1^4,2,-1,2), (1^4,-2,1,-2)$,
$(1^3,2,1^2,2), (1^3,2,(-1)^2,2), (1^3,-2,1^2,-2)$,
$(1^3,-2,(-1)^2,-2), (1^3,2,-1,2^2), (1^3,-2,1,(-2)^2)$,

$(1^2, -2, 1^2, (-2)^2), (1^2, -2, 1, -2, 1, -2),$
$(1^2, 2, -1, -3, 2, -3), (1^2, -2, 1, 3, -2, 3), (1, -2, 1, -2, 3, -2, 3),$
$(1, -2, 1, 3, 2^2, 3), (1, -2, 1, 3, (-2)^2, 3).$

Let $\mathbb{L}_n^{\mathrm{p}}$ be the subset of $\mathbb{L}^{\mathrm{p}}$ consisting of prime links of length n. Let $\mathbb{D}_n^*$ be the set consisting of the link diagrams $\mathrm{cl}\beta(\mathbf{x})$ for all $\mathbf{x} \in \Delta_n^*$. By Lemma (3.4), we observe that *if $L = \mathrm{cl}\beta(\mathbf{x}) \in \mathbb{L}_n^p$ for a lattice point $\mathbf{x} \in \mathbb{X}$, then there is a minimal lattice point $\mathbf{x}' \in \Delta_n$ with $\mathbf{x}' \leqq \mathbf{x}$ such that $L = \mathrm{cl}\beta(\mathbf{x}')$.* This implies that the set $\mathbb{L}_n^{\mathrm{p}}$ consists of the prime links represented by link diagrams of $\mathbb{D}_n^*$ not belonging to $\mathbb{L}_j^{\mathrm{p}}$ $(j = 1, 2, \ldots, n-1)$ (which are assumed to have already constructed by our inductive hypothesis). Step 2 is the following procedure:

Step 2. Construct $\mathbb{L}_n^{\mathrm{p}}$ from $\mathbb{D}_n^*$.

The link $\mathrm{cl}\beta(\mathbf{x})$ of a lattice point $\mathbf{x}$ of length n such that $\tilde{\mathbf{x}} = \mathbf{x}$ admits a braided link diagram with crossing number n. Thus, if a list of prime links with crossing numbers up to n is available, then this procedure will not be so difficult. In the following example, our main work is only to identify the lattice points of length $n \leqq 7$ in Example (6.2) with the prime links in Rolfsen's table [21].

Example (6.3). The following list gives the elements of the sets $\mathbb{L}_n^{\mathrm{p}}$ for $n \leqq 7$ together with the corresponding lattice points.

$\mathbb{L}_1^{\mathrm{p}}$: O $\quad \sigma(O) = 0.$
$\mathbb{L}_2^{\mathrm{p}}$: 2_1^2 $\quad \sigma(2_1^2) = 1^2.$
$\mathbb{L}_3^{\mathrm{p}}$: 3_1 $\quad \sigma(3_1) = 1^3.$
$\mathbb{L}_4^{\mathrm{p}}$: $4_1^2 < 4_1$
$\sigma(4_1^2) = 1^4,$
$\sigma(4_1) = (1, -2, 1, -2).$
$\mathbb{L}_5^{\mathrm{p}}$: $5_1 < 5_1^2$
$\sigma(5_1) = 1^5,$
$\sigma(5_1^2) = (1^2, -2, 1, -2).$
$\mathbb{L}_6^{\mathrm{p}}$: $6_1^2 < 5_2 < 6_2 < 6_3^3 < 6_1^3 < 6_3 < 6_2^3 < 6_3^2$
$\sigma(6_1^2) = 1^6,$
$\sigma(5_2) = (1^3, 2, -1, 2),$
$\sigma(6_2) = (1^3, -2, 1, -2),$
$\sigma(6_3^3) = (1^2, 2, 1^2, 2),$
$\sigma(6_1^3) = (1^2, -2, 1^2, -2),$
$\sigma(6_3) = (1^2, -2, 1, (-2)^2),$
$\sigma(6_2^3) = (1, -2, 1, -2, 1, -2),$
$\sigma(6_3^2) = (1, -2, 1, 3, -2, 3).$
$\mathbb{L}_7^{\mathrm{p}}$: $7_1 < 6_2^2 < 7_1^2 < 7_7^2 < 7_8^2 < 7_4^2 < 7_2^2 < 7_5^2 < 7_6^2 < 6_1 < 7_6 < 7_7 < 7_1^3$
$\sigma(7_1) = 1^7,$
$\sigma(6_2^2) = (1^4, 2, -1, 2),$
$\sigma(7_1^2) = (1^4, -2, 1, -2),$
$\sigma(7_7^2) = (1^3, 2, 1^2, 2),$
$\sigma(7_8^2) = (1^3, 2, (-1)^2, 2),$
$\sigma(7_4^2) = (1^3, -2, 1^2, -2),$
$\sigma(7_2^2) = (1^3, -2, 1, (-2)^2),$
$\sigma(7_5^2) = (1^2, -2, 1^2, (-2)^2),$

$\sigma(7_6^2) = (1^2, -2, 1, -2, 1, -2)$,
$\sigma(6_1) = (1^2, 2, -1, -3, 2, -3)$,
$\sigma(7_6) = (1^2, -2, 1, 3, -2, 3)$,
$\sigma(7_7) = (1, -2, 1, -2, 3, -2, 3)$
$\sigma(7_1^3) = (1, -2, 1, 3, (-2)^2, 3)$.

The following lattice points of Example (6.2):

$(1^2, 2, -1, 2)$, $(1^2, 2, (-1)^2, 2)$, $(1^3, -2, (-1)^2, -2)$, $(1^3, 2, -1, 2^2)$, $(1, -2, 1, 3, 2^2, 3)$, are removed from the list, since these links are seen to be 4_1^2, 6_3^3, 7_7^2, 6_3^2, 6_3^3, respectively. The links $7_2, 7_3, 7_4, 7_5, 7_3^2$ in Rolfsen's table of [21] are also excluded from the list since these links turn out to have lengths greater than 7. In Steps 3 and 4, powers of low dimensional topology techniques will be seriously tested.

Step 3. Construct the subset $\mathbb{L}_n^\pi \subset \mathbb{L}_n^{\mathrm{p}}$ by removing every link $L \in \mathbb{L}_n^{\mathrm{p}}$ such that there is a link $L' \in \mathbb{L}_j^{\mathrm{p}}$ $(j \leqq n)$ with $L' < L$ and $\pi_1 E(L) = \pi_1 E(L')$.

By construction, we see that the set $\mathbb{L}_n^\pi$ consists of π-minimal links of length n. Among the links in Example (6.3), we see that $E(4_1^2) = E(7_7^2)$ and $E(5_1^2) = E(7_8^2)$ by taking one full twist along a component and that except these identities, all the links have mutually distinct link groups by using the following lemma on the Alexander polynomials:

LEMMA (6.4). *Let $A(t_1, t_2, \ldots, t_r)$ and $A'(t_1, t_2, \ldots, t_r)$ be the Alexander polynomials of oriented links L and L' with r components. If there is a homeomorphism $E(L) \to E(L')$, then there is an automorphism ψ of the multiplicative free abelian group $\langle t_1, t_2, \ldots, t_r\rangle$ with basis t_i $(i = 1, 2, \ldots, r)$ such that*

$$A'(t_1, t_2, \ldots, t_r) = \pm t_1^{s_1} t_2^{s_2} \ldots t_r^{s_r} A(\psi(t_1), \psi(t_2), \ldots, \psi(t_r))$$

for some integers s_i $(i = 1, 2, \ldots, r)$.

The proof of this lemma is direct from the definition of the Alexander polynomial(see [15]). Thus, we obtain the following example:

Example (6.5). We have $\mathbb{L}_n^\pi = \mathbb{L}_n^{\mathrm{p}}$ for $n \leqq 6$ and

$$\mathbb{L}_7^\pi : \quad 7_1 < 6_2^2 < 7_1^2 < 7_4^2 < 7_2^2 < 6_1 < 7_5^2 < 7_6^2 < 7_6 < 7_7 < 7_1^3 .$$

Let $\mathbb{M}_n$ be the subset of $\mathbb{M}$ consisting of 3-manifolds of length n, and $\mathbb{L}_n^{\mathbb{M}}$ the subset of $\mathbb{L}_n^\pi$ by removing a π-minimal link $L \in \mathbb{L}_n^\pi$ such that there is a π-minimal link $L' \in \mathbb{L}_j^\pi$ $(j \leqq n)$ with $L' < L$ and $\chi(L, 0) = \chi(L', 0)$. The following step is the final step of our classification program:

Step 4. Construct the set $\mathbb{L}_n^{\mathbb{M}}$.

Let $L_i\,(i = 1, 2, \ldots, r)$ be the π-minimal links in the set $\mathbb{L}_n^{\mathbb{M}}$, ordered by Ω_c. Then we have $M_{n,i} = \chi(L_i, 0)$, $\alpha(M_{n,i}) = L_i$ $(i = 1, 2, \ldots, r)$. An important observation is that every 3-manifold in $\mathbb{M}$ appears once as $M_{n,i}$ without overlaps. As we shall show later, the 0-surgery manifolds of the π-minimal links in Example (6.5) are mutually non-homeomorphic, so that we have the complete list of 3-manifolds in $\mathbb{M}$ with length $\leqq 7$ as stated in Example (6.6).

Example (6.6).
$M_{1,1} = \chi(O, 0)$, $\quad \mathbf{x}_{1,1} = 0$,
$M_{2,1} = \chi(2_1^2, 0)$, $\quad \mathbf{x}_{2,1} = 1^2$,
$M_{3,1} = \chi(3_1, 0)$, $\quad \mathbf{x}_{3,1} = 1^3$,
$M_{4,1} = \chi(4_1^2, 0)$, $\quad \mathbf{x}_{4,1} = 1^4$,

$M_{4,2} = \chi(4_1, 0)$, $\mathbf{x}_{4,2} = (1, -2, 1, -2)$,
$M_{5,1} = \chi(5_1, 0)$, $\mathbf{x}_{5,1} = 1^5$,
$M_{5,2} = \chi(5_1^2, 0)$, $\mathbf{x}_{5,2} = (1^2, -2, 1, -2)$,
$M_{6,1} = \chi(6_1^2, 0)$, $\mathbf{x}_{6,1} = 1^6$,
$M_{6,2} = \chi(5_2, 0)$, $\mathbf{x}_{6,2} = (1^3, 2, -1, 2)$,
$M_{6,3} = \chi(6_2, 0)$, $\mathbf{x}_{6,3} = (1^3, -2, 1, -2)$,
$M_{6,4} = \chi(6_3^3, 0)$, $\mathbf{x}_{6,4} = (1^2, 2, 1^2, 2)$,
$M_{6,5} = \chi(6_1^3, 0)$, $\mathbf{x}_{6,5} = (1^2, -2, 1^2, -2)$,
$M_{6,6} = \chi(6_3, 0)$, $\mathbf{x}_{6,6} = (1^2, -2, 1, (-2)^2)$,
$M_{6,7} = \chi(6_2^3, 0)$, $\mathbf{x}_{6,7} = (1, -2, 1, -2, 1, -2)$,
$M_{6,8} = \chi(6_3^2, 0)$, $\mathbf{x}_{6,8} = (1, -2, 1, 3, -2, 3)$.
$M_{7,1} = \chi(7_1, 0)$, $\mathbf{x}_{7,1} = 1^7$,
$M_{7,2} = \chi(6_2^2, 0)$, $\mathbf{x}_{7,2} = (1^4, 2, -1, 2)$,
$M_{7,3} = \chi(7_1^2, 0)$, $\mathbf{x}_{7,3} = (1^4, -2, 1, -2)$,
$M_{7,4} = \chi(7_4^2, 0)$, $\mathbf{x}_{7,4} = (1^3, -2, 1^2, -2)$,
$M_{7,5} = \chi(7_2^2, 0)$, $\mathbf{x}_{7,5} = (1^3, -2, 1, (-2)^2)$,
$M_{7,6} = \chi(7_5^2, 0)$, $\mathbf{x}_{7,6} = (1^2, -2, 1^2, (-2)^2)$,
$M_{7,7} = \chi(7_6^2, 0)$, $\mathbf{x}_{7,7} = (1^2, -2, 1, -2, 1, -2)$,
$M_{7,8} = \chi(6_1, 0)$, $\mathbf{x}_{7,8} = (1^2, 2, -1, -3, 2, -3)$,
$M_{7,9} = \chi(7_6, 0)$, $\mathbf{x}_{7,9} = (1^2, -2, 1, 3, -2, 3)$,
$M_{7,10} = \chi(7_7, 0)$, $\mathbf{x}_{7,10} = (1, -2, 1, -2, 3, -2, 3)$,
$M_{7,11} = \chi(7_1^3, 0)$, $\mathbf{x}_{7,11} = (1, -2, 1, 3, (-2)^2, 3)$.

To see that the 3-manifolds in Example (6.6) are mutually non-homeomorphic, we first check their first integral homology. It is computed as follows:

(1) $H_1(M) = \mathbb{Z}$ for $M = M_{1,1}$, $M_{3,1}$, $M_{4,2}$, $M_{5,1}$, $M_{6,2}$, $M_{6,3}$, $M_{6,6}$, $M_{7,1}$, $M_{7,8}$, $M_{7,9}$, $M_{7,10}$.

(2) $H_1(M) = \mathbb{Z} \oplus \mathbb{Z}$ for $M = M_{5,2}$, $M_{7,4}$, $M_{7,7}$.

(3) $H_1(M) = \mathbb{Z}_2$ for $M = M_{6,4}$, $M_{6,5}$, $M_{7,11}$.

(4) $H_1(M) = \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}$ for $M = M_{6,7}$.

(5) $H_1(M) = \mathbb{Z}_2 \oplus \mathbb{Z}_2$ for $M = M_{4,1}$, $M_{6,8}$, $M_{7,6}$.

(6) $H_1(M) = \mathbb{Z}_3 \oplus \mathbb{Z}_3$ for $M = M_{6,1}$, $M_{7,2}$.

(7) $H_1(M) = 0$ for $M = M_{2,1}$, $M_{7,3}$, $M_{7,5}$.

For (1), since the Alexander polynomial of a knot K is an invariant of the homology handle $\chi(K, 0)$, we see that the homology handles of (1) are mutually distinct. For (2), since the Alexander polynomial of an oriented link L with all the linking numbers 0 is also an invariant of $\chi(L, 0)$ in the sense of Lemma (6.4), these 3-manifolds are mutually distinct. For (3), we note that $M_{6,4} = P^3$ the projective 3-space, $M_{6,5} = \chi(3_1, -2)$ (where we take 3_1 as the positive trefoil knot) and $M_{7,11} = \chi(4_1, -2)$. We take the connected double covering spaces $\tilde{M}$ of $M = M_{6,4}$, $M_{6,5}$ and $M_{7,11}$. The homology $H_1(\tilde{M})$ for $M = M_{6,4}$, $M_{6,5}$ or $M_{7,11}$ is, respectively, computed as 0, $\mathbb{Z}_3$, $\mathbb{Z}_5$, showing that these 3-manifolds are mutually distinct. For (4), we have nothing to prove. Note that $M_{6,7} = T^3$. For (5), we compare the first integral homologies of the three kinds of connected double coverings of every $M = M_{4,1}$, $M_{6,8}$, $M_{7,6}$. For $M = M_{4,1}$, it is the quaternion space Q and we have $H_1(\tilde{M}) = \mathbb{Z}_4$ for every connected double covering space $\tilde{M}$ of M. For $M = M_{6,8}$, we have $H_1(\tilde{M}; \mathbb{Z}_3) = \mathbb{Z}_3$ for every connected double covering space $\tilde{M}$ of M. On the other hand, for $M = M_{7,6}$, we have $H_1(\tilde{M}) = \mathbb{Z}_{20}$ and $H_1(\tilde{M};\mathbb{Z}_3) = 0$ for some connected double covering space $\tilde{M}$ of M. Thus, these 3-manifolds are mutually distinct. For (6), we use the following lemma:

Lemma (6.7). *Let $H = \mathbb{Z}_p \oplus \mathbb{Z}_p$ for an odd prime $p > 1$. If the linking form $\ell : H \times H \longrightarrow \mathbb{Q}/\mathbb{Z}$ is hyperbolic, then the hyperbolic $\mathbb{Z}_p$-basis e_1, e_2 of H is unique up to unit multiplications of $\mathbb{Z}_p$.*

Proof. Let e_1', e_2' be another hyperbolic $\mathbb{Z}_p$-basis of H. Let $e_i' = a_{i1}e_1 + a_{i2}e_2$. Then

$$0 = \ell(e_i', e_i') = \frac{2a_{i1}a_{i2}}{p} \qquad (\mathrm{mod}\ 1),$$
$$\frac{1}{p} = \ell(e_1', e_2') = \frac{a_{11}a_{22} + a_{12}a_{21}}{p} \qquad (\mathrm{mod}\ 1).$$

By these identities, we have either $e_1' = a_{11}e_1$ and $e_2' = a_{22}e_2$ with $a_{11}a_{22} = 1$ in $\mathbb{Z}_p$ or $e_1' = a_{12}e_2$ and $e_2' = a_{21}e_1$ with $a_{12}a_{21} = 1$ in $\mathbb{Z}_p$. □

By Lemma (6.7), there are just two connected $\mathbb{Z}_3$-coverings $\tilde{M}$ of every $M = M_{6,1}$, $M_{7,2}$ associated with a hyperbolic direct summand $\mathbb{Z}_3$ of $H_1(M) = \mathbb{Z}_3 \oplus \mathbb{Z}_3$. In other words, the covering $\tilde{M}$ is associated with a $\mathbb{Z}_3$-covering covering of the exterior $E(L)$ lifting one torus boundary component trivially, where $L = 6_1^2, 6_2^2$. Since the link L is interchangeable, it is sufficient to check one covering for each M. For $M = M_{6,1}$ we have $H_1(\tilde{M}) = \mathbb{Z}_9 \oplus \mathbb{Z}_3$ and for $M = M_{7,2}$ we have $H_1(\tilde{M}) = \mathbb{Z} \oplus \mathbb{Z}$. Thus, these 3-manifolds are distinct. For (7), the Dehn surgery manifolds $\chi(7_1^2, 0)$ and $\chi(7_2^2, 0)$ are the boundaries of Mazur manifolds (which are normal imitations of S^3) and identified with the Brieskorn homology 3-spheres $\Sigma(2,3,13)$, $\Sigma(2,5,7)$ by S. Akbult and R. Kirby [1]. Hence, we have $M_{2,1} = S^3$, $M_{7,3} = \Sigma(2,3,13)$, and $M_{7,5} = \Sigma(2,5,7)$, and these 3-manifolds are mutually distinct. Thus, we see that the 3-manifolds of Example (6.6) are mutually distinct.

For the Poincaré homology 3-sphere $\Sigma = \Sigma(2,3,5)$, the prime link $\alpha(\Sigma)$ must have at least 10 components. [To see this, assume that $\alpha(\Sigma)$ has r components. Using that Σ is a homology 3-sphere and $\Sigma = \chi(\alpha(\Sigma), 0)$, we see that Σ bounds a simply connected 4-manifold W with an $r \times r$ non-singular intersection matrix whose diagonal entries are all 0. Since the Rochlin invariant of Σ is non-trivial, it follows that the signature of W is an odd multiple of 8 and r is even. Hence $r \geqq 8$. If $r = 8$, then the intersection matrix is a positive or negative definite matrix, which is not the case. Thus, we have $r \geqq 10$.] Since $\chi(3_1, 1) = \Sigma$ for the positive trefoil knot 3_1, an answer to the following question on Kirby calculus (see [13], [19], [21]) will help in understanding the link $\alpha(\Sigma)$:

Question (6.8). How is Ω_c understood among colored links?

We note that the cardinal numbers $l_n = \#\mathbb{L}_n^{\mathrm{p}}$ and $m_n = \#\mathbb{M}_n$ are independent of a choice of any well-order Ω of $\mathbb{X}$ with the condition that any lattice points $\mathbf{x}, \mathbf{y}$ with $\ell(\mathbf{x}) < \ell(\mathbf{y})$ has the order $\mathbf{x} < \mathbf{y}$. A sequence of non-negative integers k_n $(n = 1, 2, \dots)$ is a *polynomial growth sequence* if there is an integral polynomial $f(x)$ in one variable x such that $k_n \leqq f(n)$ for all n. Concerning the classifications of $\mathbb{L}^{\mathrm{p}}$ and $\mathbb{M}$, the following question may be of interest:

Question (6.9). Are the sequences l_n and m_n $(n = 1, 2, \dots)$ polynomial growth sequences?

Let p_n be the number of prime links with the crossing number n. C. Ernst and D.W. Sumners [8] showed that the sequence p_n $(n = 0, 1, 2, \dots)$ is not any polynomial growth sequence by counting the numbers of 2-bridge knots and links.

7. Notes on the oriented version and oriented 3-manifold invariants

Let $\vec{\mathbb{M}}$ be the set of closed connected oriented 3-manifolds. Using the injection $\vec{\sigma} : \vec{\mathbb{L}} \to \mathbb{X}$, we have a well-order in $\vec{\mathbb{L}}$ induced from a well-order Ω in $\mathbb{X}$ and also denoted by Ω. Writing

$$\vec{\mathbb{L}}^{\pi} = \iota^{-1}\mathbb{L}^{\pi} \subset \vec{\mathbb{L}},$$

we can show that the embedding $\alpha : \mathbb{M} \to \mathbb{L}$ in Theorem (1.1) lifts to an embedding

$$\vec{\alpha} : \vec{\mathbb{M}} \longrightarrow \vec{\mathbb{L}}$$

such that $\chi_0\vec{\alpha} = 1$ and $\vec{\alpha}(-M) = -\overline{\vec{\alpha}(M)}$ for every $M \in \vec{\mathbb{M}}$, where the map $\chi_0 : \vec{\mathbb{L}} \to \vec{\mathbb{M}}$ denotes the oriented version of the 0-surgery map $\chi_0 : \mathbb{L} \to \mathbb{M}$. To see this, for every $M \in \vec{\mathbb{M}}$, we note that the link $L_0 = \mathrm{cl}\beta\sigma_\alpha(M)$ is canonically oriented and $\chi(L_0, 0) = \pm M$, where $-M$ denotes M with opposite orientation. If $M = -M$, then we define $\vec{\alpha}(M) = L_0$. If $M \neq -M$, then we define $\vec{\alpha}(M)$ so as to satisfy

$$\{\vec{\alpha}(M), \vec{\alpha}(-M)\} = \{L_0, -\bar{L}_0\} \quad \text{and} \quad \chi(\vec{\alpha}(M), 0) = M.$$

As a related question, it would be interesting to know *whether or not there is an oriented link* $L \in \vec{\mathbb{L}}$ *with* $L = -\bar{L}$ and $\chi(L, 0) = M$ *for every* $M \in \vec{\mathbb{M}}$ *with* $M = -M$.

For an algebraic system Λ, an *oriented* 3-*manifold invariant* in Λ is a map $\vec{\mathbb{M}} \to \Lambda$ and *an oriented link invariant* in Λ is a map $\vec{\mathbb{L}} \to \Lambda$. Let $\mathrm{Inv}(\vec{\mathbb{M}}, \Lambda)$ and $\mathrm{Inv}(\vec{\mathbb{L}}, \Lambda)$ be the sets of oriented 3-manifold invariants and oriented link invariants in Λ, respectively. Then we have $\chi_0\vec{\alpha} = 1$. We consider the following sequence

$$\mathrm{Inv}(\vec{\mathbb{M}}, \Lambda) \xrightarrow{\chi_0^{\#}} \mathrm{Inv}(\vec{\mathbb{L}}, \Lambda) \xrightarrow{\vec{\alpha}^{\#}} \mathrm{Inv}(\vec{\mathbb{M}}, \Lambda)$$

of the dual maps $\vec{\alpha}^{\#}$ and $\chi_0^{\#}$ of $\vec{\alpha}$ and χ_0. Since the composite $\vec{\alpha}^{\#}\chi_0^{\#} = 1$, we see that $\chi_0^{\#}$ is injective and $\vec{\alpha}^{\#}$ is surjective, both of which imply that *every oriented* 3-*manifold invariant can be obtained from an oriented link invariant.* More precisely, if I is an oriented 3-manifold invariant, then $\chi_0^{\#}(I)$ is an oriented link invariant which takes the same value $I(M)$ on the subset $\vec{\mathbb{L}}(M) = \{L \in \vec{\mathbb{L}} | \chi(L, 0) = M\}$ for every $M \in \vec{\mathbb{M}}$. Conversely, if J is an oriented link invariant, then $\vec{\alpha}^{\#}(J)$ is an oriented 3-manifold invariant and every oriented 3-manifold invariant is obtained in this way. Here is an example creating an oriented 3-manifold invariant from an oriented link invariant when we use the right inverse $\vec{\alpha}$ of χ_0, defined by the canonical order Ω_c.

Example (7.1). We denote by V a Seifert matrix associated with a connected Seifert surface of the link (see [15]). Then the signature $\mathrm{sign}(V+V')$ and the determinant $\det(tV-V')$ give oriented link invariants, that is, the signature invariant $\lambda \in \mathrm{Inv}(\vec{\mathbb{L}}, \mathbb{Z})$ and the (one variable) Alexander polynomial $A \in \mathrm{Inv}(\vec{\mathbb{L}}, \mathbb{Z}[t,t^{-1}])$ (an oriented link invariant up to multiples of $\pm t^m$, $m \in \mathbb{Z}$). For the right inverse $\vec{\alpha}$ of χ_0 using the canonical order Ω_c, we have the oriented 3-manifold invariants

$$\lambda_{\vec{\alpha}} = \vec{\alpha}^{\#}(\lambda) \in \mathit{Inv}(\vec{\mathbb{M}}, \mathbb{Z}) \quad \mathit{and} \quad A_{\vec{\alpha}} = \vec{\alpha}^{\#}(A) \in \mathit{Inv}(\vec{\mathbb{M}}, \mathbb{Z})\,.$$

For some 3-manifolds, these invariants are calculated as follows:

(7.1.2) $\lambda_{\vec{\alpha}}(S^3) = -1, A_{\vec{\alpha}}(S^3) = t-1\,.$

(7.1.3) $\lambda_{\vec{\alpha}}(\pm Q) = \mp 3, A_{\vec{\alpha}}(\pm Q) = (t-1)(t^2+1)$ *(we note that* $Q \neq -Q$*)*.

(7.1.4) $\lambda_{\vec{\alpha}}(P^3) = -4, A_{\vec{\alpha}}(P^3) = (t-1)^2\,.$

(7.1.5) $\lambda_{\vec{\alpha}}(T^3) = 0, A_{\vec{\alpha}}(T^3) = (t-1)^4\,.$

Received January 14, 2003

Final version received February 23, 2004

Department of Mathematics
Osaka City University
Sumiyoshi-ku
Osaka 558-8585
Japan
kawauchi@sci.osaka-cu.ac.jp

References

[1] S. Akbult and R. Kirby, *Mazur manifolds*, Michigan Math. J. **26** (1979), 259–284.

[2] J. S. Birman, *Braids, links, and mapping class groups*, Ann. Math. Stud. **82**, Princeton Univ. Press, Princeton, NJ, 1974.

[3] J. S. Birman and W. W. Menasco, *Studying links via closed braids III, classifying links which are closed 3-braids*, Pacific J. Math. **161** (1993), 25–113.

[4] G. Burde and K. Murasugi, *Links and Seifert fiber spaces*, Duke Math. J. **37** (1970), 89–93.

[5] J. H. Conway, *An enumeration of knots and links and some of their related properties*, Computational problems in abstract algebra, 329–358, Pregamon Press, 1970.

[6] H. H. Doll and M. J. Hoste, *A tabulation of oriented links*, Math. Computation **57** (1991), 747–761.

[7] C. H. Dowker and M. B. Thistlethwaite, *Classifications of knot projections*, Topology Appl. **16** (1983), 19–31.

[8] C. Ernst and D. W. Sumners, *The growth of the number of prime knots*, Math. Proc. Cambridge Philos. Soc. **102** (1987), 303–315.

[9] R. W. Ghrist, *Branched two–manifolds supporting all links*, Topology **36** (1997), 423–448.

[10] C. Mc A. Gordon and J. Luecke, *Knots are determined by their complements*, J. Amer. Math. Soc. **2** (1989), 371–415.

[11] J. Hempel, *3-manifolds*, Ann. Math. Studies **86**, Princeton Univ. Press, Princeton, NJ, 1976.

[12] W. Jaco, Lectures on three-manifold topology, Conf. Board Math. **43**, Amer. Math. Soc. Providende, RI, 1980.

[13] S. J. KAPLAN, *Constructing framed 4-manifolds with given almost framed boundaries*, Trans. Amer. Math. Soc. **254** (1979), 237–263.

[14] A. KAWAUCHI, *$\tilde{H}$-cobordism I; The groups among three dimensional homology handles*, Osaka J. Math. **13** (1976), 567–590.

[15] A. KAWAUCHI, A survey of knot theory, Birkhäuser, 1996.

[16] A. KAWAUCHI, *Topological imitation of a colored link with the same Dehn surgery manifold*, Topology Appl. (to appear).

[17] A. KAWAUCHI, *Characteristic genera of closed orientable 3-manifolds*, preprint (http://www.sci.osaka-cu.ac.jp/~kawauchi/index.htm).

[18] A. KAWAUCHI AND I. TAYAMA, *Enumerating the prime knots and links by a canonical order*, Proc. First Asian School of Knots, Links, and Related Topics (Seoul, 2004), 307–316, http://knot.kaist.ac.kr/2004/(Electronic publication)(2004).

[19] R. KIRBY, *A calculus for framed links in S^3*, Invent. Math. **45** (1978), 35–56.

[20] Y. NAKAGAWA, *A family of integer-valued complete invariants of oriented knot types*, J. Knot Theory Ramifications **10** (2001), 1169–1199.

[21] D. ROLFSEN, Knots and links, Mathematics Lecture Series **7**, Publish or Perish, Inc., Berkeley, CA, 1976.

[22] W. WHITTEN, *Knot complements and groups*, Topology **26** (1987), 41–44.

[23] S. YAMADA, *The minimal number of Seifert circles equals the braid index of a link*, Invent. Math. **89** (1987), 347–356.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

SPLITTINGS OF $\mathbb{S}^4$

W.B. RAYMOND LICKORISH

A miniscule tribute to the pervasive mathematical influence of Fico González Acuña in appreciation of a long-standing friendship.

Abstract. It is shown that any two groups, with isomorphic abelianisations and finite balanced presentations, can be achieved as the fundamental groups of the two sides of a splitting of the 4-sphere by a 3-manifold. Furthermore here the two sides have handle presentations that produce the given group presentations up to Andrews-Curtis equivalence.

1. Introduction

In a previous paper [4] a method was developed of constructing, in the 4-sphere $\mathbb{S}^4$, contractible bounded 4-manifolds for which the complement had a given perfect balanced fundamental group. It was noted that the contractible manifold so formed consisted of 1-handles and 2-handles added to a 4-ball so that the resulting presentation of the trivial group could be trivialised by Andrews-Curtis moves. Thus the contractible manifold could also be doubled to form $\mathbb{S}^4$, so giving another distinct embedding of the manifold in $\mathbb{S}^4$. The method is here explored for groups other than the trivial group. Any pair of groups with balanced presentations, that give isomorphic groups when abelianised, are obtained as the fundamental groups of the two halves of some splitting of $\mathbb{S}^4$ by a 3-manifold. These two 4-manifolds have handle structures consisting of 1-handles and 2-handles added to a 4-ball and the resulting group presentations are Andrews-Curtis equivalent (but not in general equal to) the original presentations. If one reverts to the consideration of the trivial group one can for example achieve, up to Andrews-Curtis equivalence, the same presentation of the trivial group on either side of a splitting of $\mathbb{S}^4$. Doubling would also show this to be the case if the Andrews-Curtis conjecture were true. In one particular example, of a presentation which might be a counter example to this conjecture, the same manifold can be achieved on each of the two sides of a splitting. This will be explained below. If, as in this example, an embedding of a contractible manifold in $\mathbb{S}^4$ has a contractible complement, it is not easy to prove that there is any other inequivalent embedding (that is, that the manifold knots in $\mathbb{S}^4$). C.Livingston has one isolated example [5] of two such embeddings. When the complement is not simply connected he can, for carefully chosen groups, construct infinite sequences of embeddings. His method is to regard the 4-manifold as a regular neighbourhood of a contractible 2-complex in $\mathbb{S}^4$ and then change

2000 *Mathematics Subject Classification*: Primary 57M05; Secondary 57M25.
Keywords and phrases: 4-manifold in the 4-sphere, group presentation, Andrews-Curtis conjecture, matrix moves.

the embedding of that complex by taking its connected sum, in the middle of a 2-cell, with a knotted $\mathbb{S}^2$ in $\mathbb{S}^4$. After taking due care with the construction, distinct fundamental groups for the complement result. If the complement is simply connected no change can ever occur in the fundamental group of the complement by this method. In a final remark in this paper it is shown that for certain knots of $\mathbb{S}^2$ in $\mathbb{S}^4$ the 'new' embedding so constructed is actually isotopic to the original one.

The author thanks the Department of Mathematics and Statistics at the University of Melbourne for its hospitality whilst this paper was being written.

2. Group presentations

Firstly a few simple remarks, comparing matrix presentations of abelian groups with arbitrary presentations of (probably) non-abelian groups, will be recorded. Suppose a free *abelian* group E, with additive notation, is freely generated by $e_1, e_2, \ldots, e_m$. The quotient group $E/\{\ \sum_{j=1}^{m} A_{ij}e_j : i = 1, 2, \ldots, n\ \}$ is said to be presented by the $n \times m$ integer matrix $A = \{A_{ij}\}$. If A is changed, by a sequence of *matrix moves* of the following types, it is easy to see that there is no change, up to isomorphism, in the group presented by A.

Matrix moves:
(a) Add the ith row (or column) to the jth row (or column).
(b) Change the sign of the ith row (or column).
(c) Permute the rows (or columns).
(d) Enlarge the matrix A to the matrix $\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}$.

Note that (b) and (c) are self inverse, that the inverse of (a) is just a combination of moves of types (a) and (b), but that the inverse of (d) cannot be achieved by a combination of these matrix moves.

Suppose now that $\langle a_1, a_2, \ldots, a_m : r_1, r_2, \ldots, r_n \rangle$ is a presentation of a (not necessarily abelian) group in terms of generators and relators. There are various sorts of moves that can be performed on the presentation which do not change the group presented. Amongst these are the moves of J.J. Andrews and M.L. Curtis (sometimes called 'extended Nielsen transformations' or 'Q-transformations' or 'Markov operations') that are moves of the following types *and their inverses.*

Andrews-Curtis moves:
(i) Change r_j to $r_i r_j$ where $j \neq i$.
(ii) Change r_i to r_i^{-1}.
(iii) Add a new generator a_{m+1} and a new relator $a_{m+1}w$ where w is a word in $a_1, a_2, \ldots, a_m$.
(iv) Change r_i to $r_i a_j a_j^{-1}$ or $r_i a_j^{-1} a_j$.
(v) Change r_i to a cyclic permutation of r_i.

A presentation P of any group G, with multiplicative notation, obviously induces a presentation of the abelianisation of G. This is obtained by allowing the symbols to commute, cancelling all occurrences of a generator and its inverse in the same relator, and then recording the generators' exponents in each relator as a matrix A.

Lemma (2.1). *Suppose that P is a finite presentation of group G and A is the corresponding presentation matrix of its abelianisation $G/[G,G]$. If A is changed to B by a matrix move (as described above) then P can be changed by Andrews-Curtis moves to a presentation Q so that B is the matrix corresponding to Q.*

Proof. Adding the ith row to the jth row corresponds to the Andrews-Curtis move of changing r_j to $r_i r_j$. The analogue for columns is a little harder to describe but it just corresponds to the geometric idea of sliding a 1-handle over a 1-handle: Use (iii) to add a new generator a_{m+1} and relator $a_{m+1}a_j a_i^{-1}$. Then use (i) and (v) to change every occurrence of a_i in $r_1, r_2, \ldots, r_n$ to an occurrence of $a_{m+1}a_j$, then remove a_i and relator $a_{m+1}a_j a_i^{-1}$ using the inverse of (iii) and finally relabel a_{m+1} as a_i. The corresponding matrix move is that of adding the ith column to the jth column. Any relabelling of the generators throughout the presentation can be achieved by using the same idea, with the new relator being $a_{m+1}{a_i}^{-1}$, to change every occurrence of a_i to one of a_{m+1} and then to remove a_i. Thus moves inducing permutation of matrix columns and the sign change of a column can be created. Consideration of the remaining matrix moves is straightforward. □

The classification theorem for finitely generated abelian groups asserts that if an abelian group is presented by a square matrix A, then A can be changed by a sequence of the above matrix moves to a 'canonical' diagonal matrix Δ which has only prime powers, ones or zeros on the diagonal. Furthermore, up to a reordering, the non-unit elements on the diagonal of such a Δ are uniquely determined by the isomorphism class of the group presented.

Corollary (2.2). *Suppose that P is a balanced finite presentation of a group G and B is some square presentation matrix of its abelianisation $G/[G,G]$. Then P can be changed by Andrews-Curtis moves to a presentation Q for which the corresponding matrix is $\begin{pmatrix} B & 0 \\ 0 & I_r \end{pmatrix}$ for some $r \geq 0$, where I_r is the identity $r \times r$ matrix.*

Proof. The word 'balanced' means that P has the same number of generators as relators. Let the matrix A correspond to P. By the above mentioned classification theorem, each of A and B can be changed by matrix moves to become the same diagonal matrix Δ. Thus, by matrix moves, A can be changed to Δ which can be changed to $\begin{pmatrix} B & 0 \\ 0 & I_r \end{pmatrix}$, for some $r \geq 0$, by inverting the moves that change B to Δ but refusing to implement an inverse of a type (d) move. Then each move can, by lemma (2.1), be imitated in the presentation P. □

Note that Andrews-Curtis moves of type (iv) have not so far actually been used in any proof. Of course they could be regarded as intimately related to the process of producing the matrix A from a presentation P.

3. 4-manifolds corresponding to group presentations

Next is the main result about splitting $\mathbb{S}^4$ into two handlebodies each of a 0-handle, 1-handles and 2-handles to obtain designated fundamental groups for these two parts.

THEOREM (3.1). *Let P_1 and P_2 be balanced presentations of groups G_1 and G_2 having the property that $G_1/[G_1, G_1] \cong G_2/[G_2, G_2]$. Then $\mathbb{S}^4$ can be separated (by a closed connected 3-manifold) into 4-manifolds M_1 and M_2 with $\pi_1(M_1) \cong G_1$ and $\pi_1(M_2) \cong G_2$. Each of M_1 and M_2 has a handle structure consisting of one 0-handle, n 1-handles and n 2-handles, for some n, with the associated group presentation for M_1 being Andrews-Curtis equivalent to P_1 and that for M_2 being Andrews-Curtis equivalent to P_2.*

Proof. Let $P_1 = \langle a_1, a_2, \dots, a_n : r_1, r_2, \dots, r_n \rangle$. Let P_1 correspond to the $n \times n$ matrix C presenting the abelian group $G_1/[G_1, G_1]$. The transpose matrix C^τ also presents this group as, for example, follows from the symmetry with respect to rows and columns of the above mentioned classification theorem. Regard C^τ as a presentation matrix for $G_2/[G_2, G_2]$. By corollary (2.2) there is, for some $r \geq 0$, a presentation $\Pi = \langle \alpha_1, \alpha_2, \dots, \alpha_{n+r} : \rho_1, \rho_2, \dots, \rho_{n+r} \rangle$, Andrews-Curtis equivalent to P_2, so that $\begin{pmatrix} C^\tau & 0 \\ 0 & I_r \end{pmatrix}$ is the presentation matrix of $G_2/[G_2, G_2]$ corresponding to Π. Let $A = \begin{pmatrix} C & 0 \\ 0 & I_r \end{pmatrix}$. Add to P_1 generators $a_{n+1}, a_{n+2}, \dots, a_{n+r}$ and relators $r_{n+1} = a_{n+1}, r_{n+2} = a_{n+2}, \dots, r_{n+r} = a_{n+r}$ so that now A is the matrix corresponding to this new P_1. Suppose that in the relator r_i there are $n_+^{i,j}$ occurrences of the symbol a_j and $n_-^{i,j}$ of a_j^{-1}. Similarly suppose that in ρ_j there are $\nu_+^{i,j}$ occurrences of the symbol α_i and $\nu_-^{i,j}$ of α_i^{-1}. Then $n_+^{i,j} - n_-^{i,j} = A_{ij} = \nu_+^{i,j} - \nu_-^{i,j}$. If $n_+^{i,j} > \nu_+^{i,j}$ alter Π by changing ρ_j to $\rho_j(\alpha_i \alpha_i^{-1})^{n_+^{i,j} - \nu_+^{i,j}}$. If $n_+^{i,j} < \nu_+^{i,j}$ alter P_1 by changing r_i to $r_i(a_j a_j^{-1})^{\nu_+^{i,j} - n_+^{i,j}}$. These are, of course, Andrews-Curtis moves. By repeating this for every pair (i, j), it may be assumed that $n_+^{i,j} = \nu_+^{i,j}$ and hence $n_-^{i,j} = \nu_-^{i,j}$ for all (i, j).

Let $\mathbb{S}^4 = B_1 \cup B_2$, the union of two 4-balls intersecting in their common boundary S^3. In S^3 construct a link as follows. Let $D_1, D_2, \dots, D_{n+r}$ and $\Delta_1, \Delta_2, \dots, \Delta_{n+r}$ be mutually disjoint oriented discs. For each pair (i, j) for $1 \leq i, j \leq n + r$ take a collection $H_+^{i,j}$ of $n_+^{i,j}$ copies of the positive Hopf link of two ordered, oriented components and a collection $H_-^{i,j}$ of $n_-^{i,j}$ copies of the negative, ordered, oriented Hopf link. Each of these Hopf links is to be in a (small) ball in which each of the two components bounds an oriented disc meeting the other component in one point. These balls are to be all mutually disjoint and disjoint from the original discs. Now join the boundary of Δ_i once to the first component of each link in $\bigcup_j (H_+^{i,j} \cup H_-^{i,j})$ with (long thin) bands. Do this in the order around $\partial \Delta_i$ specified by the relator r_i. When a_j occurs in the relator connect to the first component of one of the links in $H_+^{i,j}$, when a_j^{-1} occurs in the relator connect to the first component of one of the links in $H_{-}^{i,j}$. Similarly when $\alpha_i^{\pm 1}$ occurs in ρ_j connect the boundary of D_j to the second components of $H_{\pm}^{i,j}$. For an occurrence of α_i any unused second component of any Hopf link in $H_{+}^{i,j}$ may be selected and similarly for α_i^{-1}. It can easily be ensured that all the bands used are mutually disjoint and that they respect all orientations (but there is enormous scope for varying the route taken by a band). Note that the numbers of links in the $H_{\pm}^{i,j}$ have been chosen so that each link in each $H_{\pm}^{i,j}$ has its first component banded to Δ_i and its second component banded to D_j. This banding process changes the original discs to two new collections $D'_1, D'_2, \ldots, D'_{n+r}$ and $\Delta'_1, \Delta'_2, \ldots, \Delta'_{n+r}$, each of mutually disjoint discs, by adding to the original discs the bands and the discs spanning the components of the Hopf links.

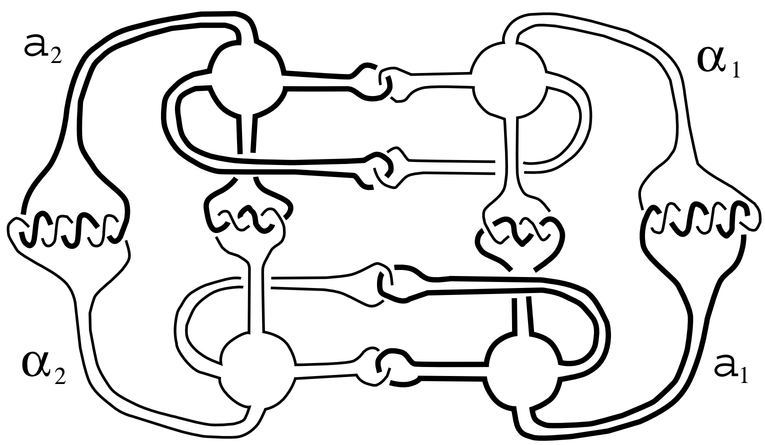
Now let S^3 be embedded in a standard way in $\mathbb{S}^4$, separating $\mathbb{S}^4$ into two 4-balls B_1 and B_2. From the 4-ball B_1 remove neighbourhoods of $n+r$ standard properly embedded discs with boundaries $\partial D'_1, \partial D'_2, \ldots, \partial D'_{n+r}$ and add them to B_2. Take these discs to be the D'_i pushed a little into B_1. This creates from B_1 a ball with $(n+r)$ 1-handles added (a technique fully described in [3]) and changes B_2 into a 4-ball with 2-handles added. Next, similarly, remove from B_2 neighbourhoods of $(n+r)$ standard properly embedded discs with boundaries $\partial\Delta'_1, \partial\Delta'_2, \ldots, \partial\Delta'_{n+r}$ and add them to B_1. Then each of B_1 and B_2 has been changed into a ball with $(n+r)$ 1-handles and $(n+r)$ 2-handles; the resulting manifolds are to be denoted M_1 and M_2.

The presentation of $\pi_1(M_1)$ coming from the handle decomposition is obtained by labelling each 1-handle with a generator and taking a relator for each 2-handle. Thus, allocate the symbol a_i to the 1-handle of M_1 corresponding to D'_i and let r_j be the relator from the 2-handle corresponding to Δ'_j. Then r_j has an entry $a^{\pm 1}$ for every signed point of $\partial\Delta'_j \cap D'_i$ taken in order along $\partial\Delta'_j$. Of course the construction has been engineered so that this r_j is indeed the jth relator of the presentation P_1. Similarly the presentation for $\pi_1(M_2)$ coming from the handle structure is indeed the presentation Π. $\square$

4. The trivial group

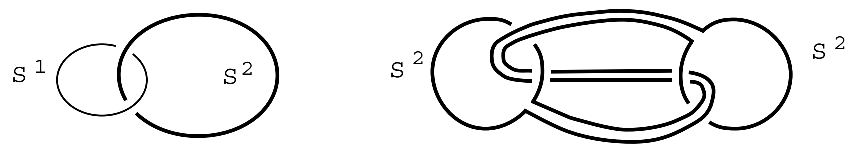
The general idea of the above proof was used in [4] to show that the 4-sphere can be split so that $\pi_1(M_1)$ is any given perfect group G with a balanced presentation and M_2 is contractible. Of course $G/[G,G]$ is then the trivial group so the trivial presentation could be used for P_2. This allowed the proof in [4] to be, in several ways, simpler than that given above. In this context, when $\pi_1(M_2)$ is to be trivial, the present theorem allows things to be chosen so that the presentation of $\pi_1(M_2)$ coming from its handle structure belongs to any given Andrews-Curtis equivalence class of presentations of the trivial group. Of course, Andrews and Curtis conjectured there to be only one such class although R.E.Gompf [2] makes the conjecture (based on much experience but little evidence) that there are infinitely many such classes.

Example (4.1). $\langle a_1, a_2 \colon a_2^{-1}a_1^{-2}a_2a_1^3,\ a_1^{-1}a_2^{-2}a_1a_2^3\rangle$. This is a famous presentation of the trivial group which is often conjectured to be inequivalent to the trivial presentation by Andrews-Curtis moves. In the notation of the above proof, the $n_{+}^{i,j}$ and $n_{-}^{i,j}$ are the terms of the *symmetric* matrices $\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ and $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. If this presentation is taken for P_1 *and* P_2 the theorem can be applied to split $\mathbb{S}^4$ into two contractible manifolds M_1 and M_2. However the symmetry of the matrices means that none of the Andrews-Curtis moves used in the proof is necessary. Thus M_1 and M_2 will have handle presentations that correspond exactly to the the given (unmoved) group presentation. After a little experimentation it can be seen that, in this case, M_1 and M_2 can be taken to be homeomorphic manifolds.



Consider for example the four simple closed curves, labelled a_1, a_2, α_1 and α_2, shown in the diagram. These bound the discs D_1', D_2', Δ_1' and Δ_2' of the proof of the theorem. Reading off the word in a_1 and a_2 from the boundaries of Δ_1' and Δ_2' does give the required presentation. However, there is a π-rotation of S^3, about the 'horizontal' bisector of the diagram, which sends a_1 to α_1 and a_2 to α_2. There is then an orientation reversing involution of $\mathbb{S}^4$ which interchanges M_1 and M_2. (Of course, the other obvious symmetry, from a rotation about an axis perpendicular to the diagram, gives an involution preserving M_1 and M_2 setwise.) It is not known whether $M_1 \times I$ is the 5-ball (it is if the presentation is Andrews-Curtis equivalent to a trivial presentation). If it is, then $\mathbb{S}^4 = \partial(M_1 \times I)$ has an orientation reversing involution that interchanges $M_1 \times 0$ and $M_1 \times 1$ and is fixed on $\partial\left(M_1 \times \frac{1}{2}\right)$. It does not seem likely that the involution of $\mathbb{S}^4$ constructed in the example is equivalent to such a homeomorphism, but it does show that M_1 can be glued to a copy of itself to give $\mathbb{S}^4$.

Example (4.2). In [5] examples were given in which a contractible 4-manifold M was embedded in $\mathbb{S}^4$ in infinitely many different knotted ways as distinguished by the fundamental group of the complement of the embeddings. The idea was to add to the interior of a disc core of a 2-handle of M a knotted $\mathbb{S}^2$ in $\mathbb{S}^4$ in the manner of connected sums. Careful choices enabled the examples to be valid. It is easy to see, using the Van Kampen theorem that, if this is done when $\mathbb{S}^4 - M$ is simply connected, the fundamental group of the complement of the embedding remains trivial. However, does that mean that the modified embedding is isotopic to the original one? Can the M_1 of the above example be embedded in $\mathbb{S}^4$ in an inequivalent way? The next diagram is meant to indicate, in the following way, that the embedding does not change if the connected sum technique is used with certain types of knotted $\mathbb{S}^2$ in $\mathbb{S}^4$.



An S^1 and $\mathbb{S}^2$ link in a standard (homological) way in $\mathbb{S}^4$. This S^1 can just be considered as the boundary of a meridional disc of an unknotted $\mathbb{S}^2$. Now take in $\mathbb{S}^4$ two copies of $\mathbb{S}^2$ that are unknotted and unlinked and pipe them together by a tube. The tube is a copy of $S^1 \times I$ that is contained in the boundary of an arc joining the two spheres. The arc is to be chosen so that it follows the path of an S^1 that links each $\mathbb{S}^2$ as indicated schematically in the diagram. It if easy to ensure (see [7] for example) that the resulting 2-sphere is knotted. However in the case of a 4-manifold M_1 in $\mathbb{S}^4$ with contractible complement, the S^1 meridional to a 2-handle is isotopic to a trivial S^1. That is because simple connectivity and general position ensure that it bounds a disc in the complement of M_1 with but isolated point self-intersections. In moving S^1 across the disc ensure that it passes through a self-intersection twice *at two different times*. Thus if to the core of the 2-handle a connected sum is taken with the above knotted $\mathbb{S}^2$, the tube can be isotopped so that it becomes a standard tube (not linking the original pair of $\mathbb{S}^2$s). Thus, up to isotopy, the construction has created no change in the embedding.

Questions (4.3). Does every contractible 4-manifold, other than the 4-ball, that embeds in $\mathbb{S}^4$ always knot in $\mathbb{S}^4$? If a 4-manifold knots in $\mathbb{S}^4$ does it always have infinitely many knots in $\mathbb{S}^4$? If unbalanced presentations of groups G_1 and G_2 are given, still with $G_2/[G_2, G_2] \cong G_1/[G_1, G_1]$, the above theorem could be applied *after* balancing the presentations with extra generators (so changing the groups by free products with free groups) and empty relators (a change not possible with Andrews-Curtis moves). That adds no deep understanding. However Livingston [6] has shown that certain perfect groups without balanced presentations are the fundamental groups of $(\mathbb{S}^4 - M^4)$ for some contractible 4-manifolds M^4. In what way does his technique generalise to splitting $\mathbb{S}^4$ into two parts with prescribed fundamental groups without balanced presentations?

Received February 19, 2003

Final version received July 29, 2003

Department of Pure Mathematics and Mathematical Statistics
University of Cambridge
Wilberforce Road
Cambridge CB3 OWB
UK
wbrl@dpmms.cam.ac.uk

References

[1] J. J. Andrews and M. L. Curtis, *Free groups and handlebodies*, Proc. Amer. Math. Soc. **16** (1965), 192–195.
[2] R. E. Gompf, *Killing the Akbulut-Kirby* 4*-sphere, with relevance to the Andrews-Curtis and Schoenflies problems*, Topology **30** (1991), 97–115.
[3] R. E. Gompf and A. I. Stipsicz, 4-manifolds and Kirby calculus, Grad. Stud. Math. **20**, Amer. Math. Soc., Providence, RI, 1999.
[4] W. B. R. Lickorish, *Knotted contractible 4–manifolds in the 4–sphere*, Pacific J. Math. **208** (2003), 283–290.
[5] C. Livingston, *Observations on Lickorish knotting of contractible 4-manifolds*, Pacific J. Math. **209** (2003), 319–323.
[6] C. Livingston, *Four-manifolds of large negative deficiency*, to appear.
[7] D. Rolfsen, *Knots and links*, Publish or Perish (1976).

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

TOROIDAL DEHN FILLINGS AND GENERALIZED SCHARLEMANN CYCLES

DANIEL MATIGNON AND ELSA MAYRAND[1]

ABSTRACT. This paper concerns Dehn fillings on 3-manifolds which produce an essential 2-torus. Let M be an irreducible and atoroidal 3-manifold, and T be an essential 2-torus created by a Dehn filling on M. Generically in M, the intersection of the punctured 2-torus $T \cap M$ with an arbitrary surface F, is a 1-complex which can be viewed as a graph in either T or F. A good way to get obstructions to the existence of Dehn fillings producing essential 2-spheres, or projective planes (T is switched with an essential 2-sphere or a projective plane) is to find generalized Scharlemann cycles in the graph in F (see [10, 4] respectively). This paper is devoted to find similar obstructions concerning the creation of essential tori. This obstruction is considered as a step towards bounding the finite number of exceptional Dehn fillings.

1. Introduction

Let M be a connected, compact and orientable 3-manifold such that a boundary component $\partial_1 M$ is a 2-torus. We assume that M is *irreducible* (i.e. all 2-spheres bound a 3-ball) and *atoroidal* (i.e. all 2-tori bound a solid torus or are boundary parallel).

A *slope* on $\partial_1 M$ is an isotopy class of essential unoriented simple closed curves on $\partial_1 M$. To each slope r on $\partial_1 M$ we associate the unique closed manifold $M(r)$ obtained by attaching a solid torus to M along $\partial_1 M$ in such a way that the gluing homeomorphism identifies the meridional slope of the solid torus with r. The core of the solid torus is a knot in $M(r)$, called *the core of the Dehn filling*, denoted by K_r.
If $M(r)$ contains an essential 2-torus, we say that r is a *toroidal slope* and that the r-Dehn filling is *toroidal*. Toroidal Dehn fillings are the topics of a large amount of investigations, see the nice surveys of Gordon [6, 7]. Let us say a few words about this. Since M has a non-empty boundary, it is clearly a *Haken manifold* (irreducible 3-manifold containing an essential surface). Therefore, by [15, 16], M satisfies the Thurston Geometrisation Conjecture. Thus either

M is *hyperbolic*, i.e. int$M$ admits a complete Riemannian metric of constant sectional curvature -1; or

M is a Seifert fibered space, i.e. an S^1-bundle over a surface, such that the tubular neighbourhood of the circle fibers are trivial fibered solid tori, except for

2000 *Mathematics Subject Classification*: 57N10, 57M25, 57M15.
Keywords and phrases: essential 2-torus, Dehn filling, Scharlemann cycle, generalized Scharlemann cycle .
[1] Supported by the Chair of Research of Canada grant.

a finite number of fibers, whose tubular neighbourhood are non-trivial fibered solid tori; or

M contains *an essential surface* (i.e. an incompressible, properly embedded surface, non-parallel to the boundary of M) of non-negative Euler characteristic.

Moreover by [15, 16], if M is hyperbolic then only a finite number of Dehn fillings can produce a non-hyperbolic 3-manifold. Such Dehn fillings are called *exceptional Dehn fillings*, and the toroidal family is a special class among them. So as to bound the finite number of *exceptional slopes* (slopes which correspond to exceptional Dehn fillings) we refer to the *distance* between distinct slopes. Let α and β be two distinct slopes on $\partial_1 M$. The *distance* $\Delta(\alpha, \beta)$ between the slopes α and β is the minimal geometric intersection number between two simple closed curves representing respectively α and β. Gordon has shown [5] that the distance between two toroidal slopes is bounded by 8, and has given explicitely the four 3-manifolds which admit two distinct toroidal slopes with distance apart 6, 7 and 8. Therefore, generically the distance between two toroidal slopes is bounded by 5. Moreover, Gordon has conjectured [6, 7] that if M is not one of these four special 3-manifolds, and M is hyperbolic, then the distance between two distinct exceptional slopes is bounded by 5. Until now, the bound 5 is reached only by the distance between two toroidal slopes.

So far, one of the best way to bound the distance between two slopes r, s that produce *small surfaces* (i.e. essential surfaces of non-negative Euler characteristic) $\widehat{P}$ and $\widehat{Q}$ respectively, is to study the *intersection graphs* which come from the intersection $P \cap Q$, where $P = \widehat{P} \cap M$ and $Q = \widehat{Q} \cap M$ are assumed to be in general position. We can see the graph G_P (respectively G_Q) in $\widehat{P}$ (respectively $\widehat{Q}$) considering the arc-components of $P \cap Q$ as edges and the components of $\widehat{P} - P$ (respectively $\widehat{Q} - Q$) as "fat"vertices. The *Scharlemann cycles*, which are particular disk-faces in these graphs (see later in Section 2, for a precise definition) play a key-role in the study of intersection graphs. For example, Scharlemann has shown ([14]) that if G_P contains a Scharlemann cycle and if the corresponding edges of G_Q are in a disk in $\widehat{Q}$ then $M(s)$ contains a non-trivial lens space. An efficient way to get obstructions to the existence of small surfaces, in order to bound $\Delta(r, s)$, is to find *generalized Scharlemann cycles*, which is a special subgraph in a disk, containing a Scharlemann cycle. They lead to the construction of small surfaces intersecting the core of the corresponding Dehn filling (K_r or K_s respectively) less than the original surfaces ($\widehat{P}$ or $\widehat{Q}$ respectively). They have first appeared in [4], where $\widehat{Q}$ was a projective plane. The authors proved that if G_P contains a generalized Scharlemann cycle, then $\widehat{Q}$ is not a *minimal* projective plane. A surface $\widehat{F}$ in $M(r)$ is *minimal* if the number of intersections between $\widehat{F}$ and the core of the Dehn filling is minimal amongst all the surfaces isotopic to $\widehat{F}$ in $M(r)$. Similar constructions are used by Hoffman in [10] to prove the following : if $\widehat{Q}$ is a minimal essential 2-sphere, then G_P cannot contain a generalized Scharlemann cycle (called *closed cluster*). Other recent works (see [3, 11, 12, 13]) concern generalized Scharlemann cycles and "minimal constructions".

The goal of the present paper is to discuss the existence of generalized Scharlemann cycles when one of $\widehat{P}$ or $\widehat{Q}$ is a minimal essential 2-torus. We may note that the edges of a Scharlemann cycle are not necessarily in a disk in the 2-torus. Let us fix the notations for the following.

Let α, β be two distinct slopes in $\partial_1 M$, such that $M(\beta)$ is irreducible and contains an essential 2-torus $\widehat{T}$. Let $\widehat{F}$ be an embedded surface in $M(\alpha)$, and denote by $F = \widehat{F} \cap M$ and $T = \widehat{T} \cap M$, the punctured surfaces properly embedded in M. After isotopy, we may assume that T and F are in general position.

We define the intersection graphs, which come from $F \cap T$, in the usual way (for more details, see [8] for example). Let G_F (respectively G_T) be the graph in $\widehat{F}$ (respectively $\widehat{T}$) obtained by taking the arc components of $T \cap F$ as edges and taking the components of $\widehat{F} - F$ (respectively $\widehat{T} - T$) as fat vertices.

One important property of the intersection graphs is that the edges are *labelled* by a numeration of the boundary-components of F and T in $\partial_1 M$ (see the next section for more details). This gives a label set to Scharlemann cycles and generalized Scharlemann cycles. Notice that a generalized Scharlemann cycle contains a Scharlemann cycle. We shall say that two generalized Scharlemann cycles Σ_1 and Σ_2 are *quasi-disjoint* if the label set of Σ_1 is disjoint from the label set of a Scharlemann cycle in Σ_2. We may note that the definition is symmetric and that the Scharlemann cycles in a generalized Scharlemann cycle all have the same label set. Now, we can formulate the main result of the paper.

THEOREM (1.1). *Assume that no arc-component of $F \cap T$ is boundary parallel in either F or T. If G_F contains two quasi-disjoint generalized Scharlemann cycles, then $\widehat{T}$ is not minimal.*

The remaining of the paper is organized as follows.

In the next section, we recall the basic definitions and constructions about intersection graphs. Then, we will give preliminary results.

In Section 3, we look at the topological effects of the existence of generalized Scharlemann cycles. First, we focus on the existence of a single generalized Scharlemann cycle. Then, we add a Scharlemann cycle whose label set is disjoint from the label set of the generalized Scharlemann cycle. And finally, we give the proof of Theorem (1.1).

2. Classical combinatorics on intersection graphs

Let $t = |\widehat{T} \cap K_\beta|$ be the number of intersections between $\widehat{T}$ and K_β, and $f = |\widehat{F} \cap K_\alpha|$ be the number of intersections between $\widehat{F}$ and K_α.

Recall that the vertices of G_F are the meridian disks of the α-Dehn filling that cap off the boundary-components of F in $\partial_1 M$, to obtain $\widehat{F}$. Similarly, the vertices of G_T are the meridian disks of the β-Dehn filling that cap off the boundary-components of T, to obtain $\widehat{T}$. Thus, if v is a vertex of G_F (respectively G_T), v corresponds to a component of $\widehat{F} \cap N(K_\alpha)$ (respectively $\widehat{T} \cap N(K_\beta)$), and ∂v is a boundary component of $F \cap \partial_1 M$ (respectively of T). After giving an orientation to K_α and K_β, we number the vertices of G_T : $v_1, v_2, \ldots, v_t$ so that they correspond to consecutive meridian discs of $N(K_\beta)$ in $\widehat{T} \cap N(K_\beta)$. Similarly, we number the vertices of G_F : $w_1, w_2, \ldots, w_f$ in the order that $\widehat{F}$ cuts $N(K_\alpha)$.

Each component ∂v_i of ∂T intersects each component ∂w_j in exactly $\Delta(\alpha, \beta)$ points. The endpoints of the edges of G_F (respectively G_T) can be labelled by an integer $i \in \{1, 2, ..., t\}$ (respectively $j \in \{1, 2, ..., f\}$) as follows. Each endpoint of an edge corresponds to a point in $\partial T \cap \partial F \cap \partial_1 M$. Consider one endpoint of an edge e, corresponding to the point $* \in \partial v_i \cap \partial w_j$. If e is seen as lying in G_T, the endpoint $*$ is labelled j, and if e is seen in G_F, then $*$ has the label i. Thus when travelling around ∂v_i, we see the labels $1, 2, .., f$ appearing in cyclic order, and around ∂w_j we see the labels $1, 2, \dots, t$; these sequences being repeated $\Delta(\alpha, \beta)$ times.

In the following, we assume for convenience, that F is orientable. We fix an orientation on T and F, and let the components of ∂T and $\partial F \cap \partial_1 M$ have the induced orientations. So we can assign a sign $+$ or $-$ to each component of $\partial F \cap \partial_1 M$ and each component of ∂T according to the orientation on $\partial_1 M$. Then we refer to a vertex of sign $+$ or $-$ according to whether the corresponding boundary component is of sign $+$ or $-$.

Let G be either the graph G_F or the graph G_T. Two vertices of G are *parallel* if they have the same sign, otherwise they are called *antiparallel*. Since M, F and T are orientable, we have the well known property :

Parity rule. An edge joining parallel vertices or the same vertex in G_T, joins antiparallel vertices in G_F and vice versa.

Now let G be either the graph G_F or G_T ; and $\widehat{Q}$ (resp. q) be either the surface $\widehat{F}$ or $\widehat{T}$, (resp. $q = f$ or $q = t$) according to whether $G = G_F$ or $G = G_T$. Then $Q = \widehat{Q} \cap M$. The graph G has the label set $\{1, 2, \dots, r\}$ where $\{q, r\} = \{f, t\}$.

If D is a disc-face of G, then ∂D consists of an alternating sequence of edges and *corners*, where corners are arcs between consecutive labels on the boundary of a vertex of G. An $\langle i, i+1 \rangle$-corner of G is an arc on ∂Q between two consecutively labelled components i, $i+1$ (modulo r) of $\partial R \cap \partial_1 M$, where R is the other surface. The corners of G_T (respectively G_F) are called the *T-corners* (respectively *F-corners*).

An *n-sided disc-face* of G is a disc-face whose boundary is the union of n edges and n corners. A *trivial loop* in G is a one-sided disc-face of G. Note that, if no arc-component of $F \cap T$ is boundary parallel in either F or T, then the graphs G_T and G_F contain no trivial loop.

A $\{x, y\}$-edge is an edge with one endpoint labelled x, and the other labelled y.

A *cycle* in G is a subgraph homeomorphic to a circle when shrinking its vertices to points. The *length* of a cycle is the number of edges which it contains.

An x-cycle in G is a cycle Σ bounding a disk D_Σ in $\widehat{Q}$, such that all the vertices of Σ are parallel and which can be oriented so that the tail of each edge has label x. *A great x-cycle* (see Figure 1a) is an x-cycle, such that all the vertices in the closed disc $\overline{D_\Sigma}$ are parallel.

A *Scharlemann cycle* (see Figure 1b) is an x-cycle σ that bounds a disc-face D_σ of G. Note that ∂D_σ is an alternating sequence of $\{x, x+1\}$-edges and $\langle x+1, x \rangle$-corners, so we assign the set of labels $\{x, x+1\}$ to σ, and σ is called *an $\{x, x+1\}$-Scharlemann cycle*. We can note that a Scharlemann cycle is a great cycle. *A strict great cycle* (see Figure 2a) is a great cycle which is not a

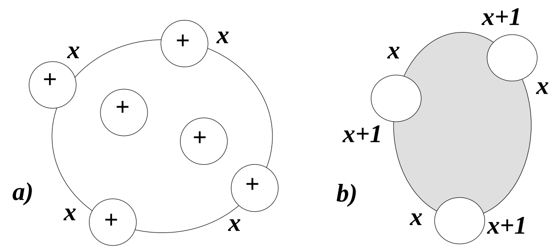


Figure 1

Scharlemann cycle. We shall say that a strict great cycle Σ is *innermost* if D_Σ contains no other strict great cycle.

Remark (2.1). *Since an x-cycle Σ of G_F is defined in a disk D_Σ in $\widehat{F}$, it is not necessary to assume that $\widehat{F}$ is orientable. Indeed, we can attribute a sign to vertices in D_Σ, and define parallel or antiparallel vertices in D_Σ. Therefore, x-cycles, great cycles and Scharlemann cycles are well defined, even if $\widehat{F}$ is not orientable. But the parity rule is thus satisfied only in D_Σ.*

The existence of great cycles guarantees the existence of Scharlemann cycles, by the following result.

Lemma (2.2) ([2, Lemma 2.6.2]). *If G contains a strict great cycle Σ, then G contains a Scharlemann cycle in D_Σ.*

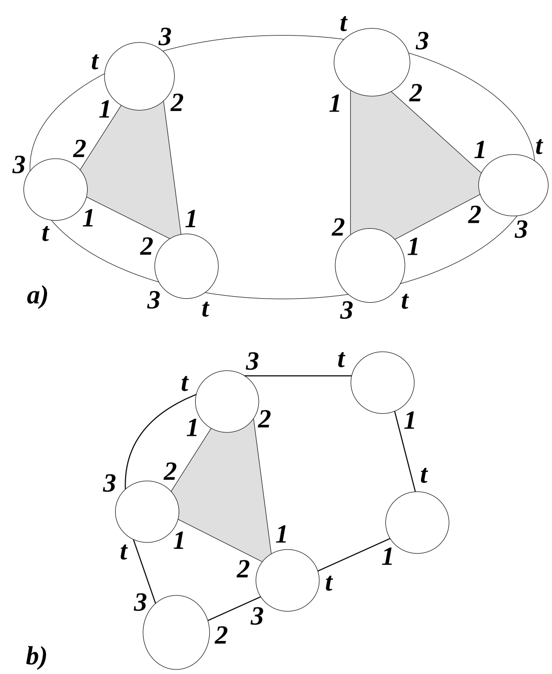


Figure 2

A generalized Scharlemann cycle in G (see Figure 2) is a subgraph Λ of G in a disk in $\widehat{Q}$, such that :

(i) Λ contains $\{x, x+1\}$-Scharlemann cycles;

(ii) all the Scharlemann cycles in Λ have the same label set;

(iii) if D is an adjacent face to a Scharlemann cycle in Λ then D is a disk-face, and its corners are exactly $\langle x-1, x\rangle$-corners and $\langle x+1, x+2\rangle$-corners;

(iv) Λ consists of its Scharlemann cycles and all their adjacent faces;

(v) every $\{x, x+1\}$-edge of Λ belongs to a Scharlemann cycle;

(vi) Λ is connected.

(vii) Λ has no cut vertex.

Note that in [10, 12], generalized Scharlemann cycles are called *(closed) clusters* and the faces defined in (*iii*) 2-*cornered faces*. The condition (*v*) implies that the subgraph in Figure 3 is not a generalized Scharlemann cycle, since its boundary contains a $\{1, 2\}$-edge.

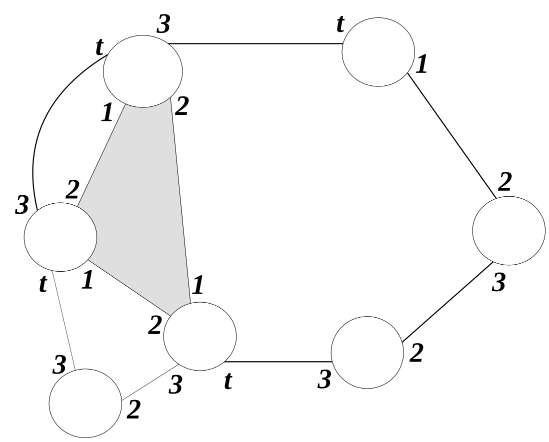


Figure 3

The label set $\{x-1, x, x+1, x+2\}$ is called *the label set of* Λ. Note that each edge in Λ has its both endpoints in $\{x-1, x, x+1, x+2\}$, and that for each $y \in \{x-1, x, x+1, x+2\}$, there exists an y-edge in Λ. Let Λ_x, Λ_y be two generalized Scharlemann cycles of label sets $\{x-1, x, x+1, x+2\}$ and $\{y-1, y, y+1, y+2\}$ respectively. Then Λ_x, Λ_y are *quasi-disjoint* if $\{x-1, x, x+1, x+2\} \cap \{y, y+1\} = \emptyset$. Note that $\{x-1, x, x+1, x+2\} \cap \{y, y+1\} = \emptyset$ if and only if $\{y-1, y, y+1, y+2\} \cap \{x, x+1\} = \emptyset$.

If Λ is a generalized Scharlemann cycle, then D_Λ denotes the union of the disk faces bounded by the Scharlemann cycles in Λ with their adjacent faces. Thus (by (*vii*)) D_Λ is a disk such that $\Lambda = G \cap \overline{D_\Lambda}$. For convenience, we sometimes refer to ∂D_Λ to be Λ; in this case, we would rather note Σ (as a cycle) instead of Λ to avoid confusion.

Remark (2.3). *As in the previous remark, generalized Scharlemann cycles are well defined even if* $\widehat{Q} = \widehat{F}$ *is non-orientable.*

An x-face in G is a disk D in $\widehat{Q}$ bounded by a cycle of G, such that all the vertices in $\overline{D}$ are parallel, and all the edges in ∂D are x-edges. *A strict x-face* is an x-face which is not a Scharlemann cycle. Actually, the existence of strict great cycles or strict x-faces guarantees the existence of generalized Scharlemann cycles, by the following result.

Lemma (2.4) ([10, Lemma 4.1], [12, Lemma 3.1]). *Assume that the Scharlemann cycles in G all have the same label set. If G contains a strict great cycle or a strict x-face, bounding a disk D, then G contains a generalized Scharlemann cycle in $IntD$.*

3. Generalized Scharlemann cycles

This section is devoted to general results concerning the effects of the existence of both Scharlemann cycle and generalized Scharlemann cycle with disjoint labels sets on the minimality of $\widehat{T}$. There are three subsections. The first one is focused on the effects of the existence of the generalized Scharlemann cycle. The second one is interested in the obstructions given by the existence of both of them. Last subsection is the proof of Theorem 1.

In the following, we assume that G_F contains a generalized Scharlemann cycle Σ and a Scharlemann cycle σ_x with disjoint label sets. Then $t \geq 6$. After changing the labelling if necessary, we may assume that the label set of Σ is $\{t, 1, 2, 3\}$ and the label set of σ_x is $\{x, x+1\}$, with $\{x, x+1\} \cap \{t, 1, 2, 3\} = \emptyset$. We keep the previous notations. Recall that the vertices of $G_T(\sigma_x)$ are v_x and v_{x+1}.

Let L be a subgraph of G_F. We denote by $G_T(L)$ the subgraph of G_T whose edges correspond to the edges of L, and whose vertices are the vertices of G_T incident to these edges. As example, $G_T(\Sigma)$ is the subgraph of G_T whose edges correspond to the edges of Σ, and whose vertices are the vertices v_t, v_1, v_2 and v_3.

Two edges are said to be *parallel* if they cobound a 2-sided disk-face. The *reduced graph* $\widehat{G}$ of a graph G is obtained from G by replacing each family of parallel edges by a single edge.

Figures 4 or 5 give examples of possible graphs for $\widehat{G_T(\Sigma)}$, after some homeomorphism of $\widehat{T}$. Note that $G_T(\Sigma)$ always contain $\{t, 3\}$-edges.

Lemma (3.1) ([1, Lemma 2.8], [9, Lemma 3.1]). *If G_F contains a Scharlemann cycle, then its edges cannot lie in a disc in $\widehat{T}$. Furthermore, $\widehat{T}$ is separating, and then t is even.*

Lemma (3.2). *If $G_T(\Sigma)$ has neither a $\{2,3\}$-edge nor a $\{t,1\}$-edge, then the $\{3,t\}$-edges of $G_T(\Sigma)$ form an essential loop on $\widehat{T}$.*

Proof. Assume that $G_T(\Sigma)$ has neither a $\{2,3\}$-edge nor a $\{t,1\}$-edge. In this case, the edges of Σ are all $\{1,2\}$-edges or $\{t,3\}$-edges. So the boundary of D_Σ is a 3-cycle Σ^* such that all its edges are $\{t,3\}$-edges. By Lemma (3.1), these edges lie in an annulus on $\widehat{T}$, which is disjoint from $G_T(\sigma)$, where σ is a Scharlemann cycle on Σ. Now, suppose for a contradiction that there is a disk D in this annulus, which contains the vertices v_t, v_3 and all the $\{t,3\}$-edges

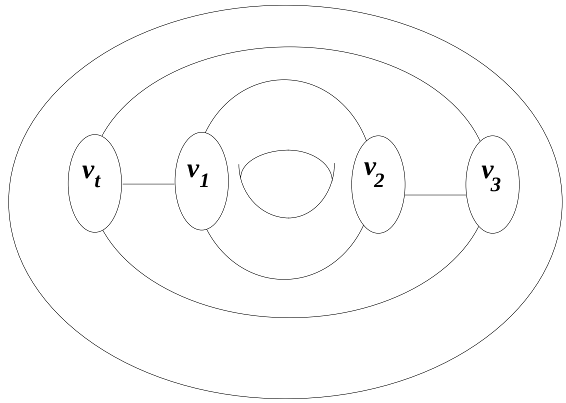


Figure 4

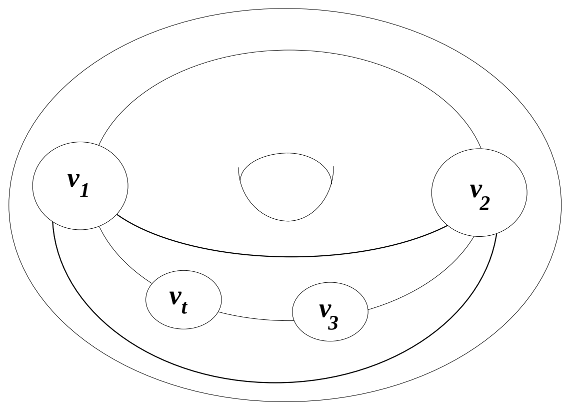


Figure 5

of $G_T(\Sigma)$. Let H be the 3-ball in $N(K_\beta)$ between v_t and v_3, and containing v_1. Then $N(D \cup H \cup D_\Sigma)$ is a punctured lens space. Therefore, since $M(\beta)$ is irreducible, it is a lens space, and then $M(\beta)$ is atoroidal; a contradiction. □

By Lemma (3.1), the existence of σ_x implies that $G_T(\Sigma)$ lies in an essential annulus A_1 in $\widehat{T}$.

Moreover, $\widehat{T}$ is separating. So we may color the faces of Σ black and white, so that the Scharlemann cycle faces are colored black, and all the others are white. Let X_B, X_W be respectively the black and white sides of $M(\beta) - \widehat{T}$, i.e. $M(\beta) = X_B \cup_{\widehat{T}} X_W$.

In the remainder of the paper, let $H_{i,i+1}$ be the 3-ball which is the portion of $N(K_\beta)$ between the vertices v_i and v_{i+1} that contains no other vertex.

Let V_1 be the solid torus $N(A_1)$ and $Y = N(A_1 \cup H_{2,3} \cup H_{t,1})$, pushed slightly inside X_W so that A_1 lies in ∂Y. Then Y is a genus three handlebody in X_W.

(3.3) Construction from Σ. A white face g of Σ is said to be *interior* if g is adjacent to at least two black faces (i.e., $\{1,2\}$-Scharlemann cycles); otherwise we say that g is a *boundary face*. We say that a black face σ in Σ is *outermost* if all the faces adjacent to σ are boundary faces, except at most one.

LEMMA (3.3.1). *The subgraph Σ contains an outermost black face in D_Σ.*

Proof. We construct a dual graph, in the following way. For each black face and interior white face g, we attribute a dual vertex v in Int$g$. For each $\{1,2\}$-edge common to a black face and an interior white face, we fix a transversal dual edge joining the corresponding dual vertices. The dual graph Γ consists in the dual vertices and the dual edges. Assume that all the black faces have at least two edges adjacent to white faces in D_Σ. Then Γ clearly contains a cycle, i.e. a subgraph homeomorphic to a circle. Thus, this cycle bounds a disk in D_Σ. Since $\Sigma = G_F \cap \overline{D_\Sigma}$, all the F-corners occur in Σ, in contradiction with $t \geq 6$. Therefore, Γ cannot contain a cycle, and so G_F contains an outermost black face in D_Σ. □

Let σ be an outermost black face in D_Σ. We have made the confusion between the disk-face σ and its boundary denoted also by σ, which is a $\{1,2\}$-Scharlemann cycle. By Lemma (3.1), there are two edges e_1, e_2 in the cycle σ so that the simple closed curve $\gamma = e_1 \cup e_2$, obtained by shrinking the vertices v_1 and v_2 to points, is essential on $\widehat{T}$. Let f_1, f_2 be the white faces adjacent to σ along the edges e_1 and e_2 respectively. As in the proof of Lemma (3.3.1), since $\Sigma = G_F \cap \overline{D_\Sigma}$, $t \geq 6$ and there are only four corners that occur in Σ : $f_1 \neq f_2$. Note that σ is outermost, so at least one face among f_1, f_2 is a boundary face, say f_1.

LEMMA (3.3.2). *If g is a white face of Σ then ∂g is an essential and non-separating simple closed curve in ∂Y.*

Proof. Let g be a white face of Σ. Orient arbitrarily its boundary ∂g. Therefore, the meridians of $\partial H_{t,1}$ are always intersected by ∂g in the same direction (and similarly for the meridians of $\partial H_{2,3}$). That implies that ∂g is essential and non-separating on ∂Y. □

PROPOSITION (3.3.3). *The annulus A_1 lies in a 2-torus component of $\partial N(Y \cup f_1 \cup f_2)$.*

Proof. Let $G = G_T(\sigma \cup f_1 \cup f_2)$ be the graph consisting of the edges of $\sigma \cup \partial f_1 \cup \partial f_2$ and the vertices v_t, v_1, v_2 and v_3. One can note that each white face has $\{1,2\}$-edges and $\{t,3\}$-edges on its boundary. More precisely, in the boundary of a white face, the number of $\{1,2\}$-edges is the same as the number of $\{t,3\}$-edges (for more details, see [4, Lemma 2.1]). Thus G has $\{1,2\}$-edges and $\{t,3\}$-edges, and possibly $\{2,3\}$-edges and $\{t,1\}$-edges.

Recall that the F-corners correspond to arcs on $\partial H_{i,i+1}$ and the T-corners are the arcs on the components ∂v_i of T. An *n-gon* (respectively, a *bigon*) in G is a disc-face with n sides (respectively, with 2 sides) . We shall call again *bigon* a disc on $\partial H_{i,i+1}$ whose boundary is the union of two F-corners on $\partial H_{i,i+1}$, a T-corner on ∂v_i, and a T-corner on ∂v_{i+1} (note that $i = t$ or 2).

Let $\gamma_1 = \partial f_1$, and $\gamma_2 = \partial f_2$. Then γ_1 and γ_2 are essential circles on ∂Y, by Lemma (3.3.2). A circle γ_i is a union of edges and F-corners. Each connected component of $\partial Y - \{\gamma_1, \gamma_2\}$ is a union of faces of G, bigons on $\partial H_{t,1}$ and bigons on $\partial H_{2,3}$. For convenience, in the following we call n-gon an n-gon for which $n > 2$; otherwise we explicitely say 'bigon'.

Claim (3.3.4). *An annulus component of $\partial Y - \{\gamma_1, \gamma_2\}$ is a union of bigons.*

Proof. Let C be an annulus component of $\partial Y - \{\gamma_1, \gamma_2\}$. Each boundary component of C is a union of consecutive F-corners and edges.

Assume that C is not a union of bigons. It is easy to see that C cannot be obtained by gluing n-gons ($n > 2$) and bigons together, since each T-corner of each n-gon must be glued to a bigon in $\partial H_{i,i+1}$ (for $i = t$ or 2). Similarly, C cannot contain an annulus face or a punctured annulus face. *(Claim (3.3.4))* □

Claim (3.3.5). *The simple closed curves γ_1 and γ_2 are not parallel on ∂Y.*

Proof. Assume for contradiction that γ_1 and γ_2 are parallel on ∂Y. Then they cobound an annulus component of $\partial Y - \{\gamma_1, \gamma_2\}$. Thus by Claim (3.3.4), it is an union of bigons. Then γ_1 and γ_2 have the same number of $\{1, 2\}$-edges; consequently exactly one (as γ_1, since f_1 is a boundary face). Recall that $e_1 \subset \gamma_1$ and $e_2 \subset \gamma_2$. Therefore, e_1 and e_2 cobound a bigon on G; a contradiction. *(Claim (3.3.5))* □

Note that this claim can be proved by using the fundamental group of Y; such an argument was used in [11].

By Lemma (3.3.2), each component of $\partial Y - \{\gamma_1, \gamma_2\}$ has at least two boundary components. By Claim (3.3.5), that implies the components of $\partial Y - \{\gamma_1, \gamma_2\}$ are punctured tori. Then each boundary component of $W = Y \cup f_1 \cup f_2$ is a 2-torus, in particular there exists a 2-torus T' which contains the annulus face A_1. *(Proposition (3.3.3))* □

Recall that $W = Y \cup f_1 \cup f_2$. Let :

- f_σ be the disk-face bounded by σ in G_F;
- $M_1 = N(A_1 \cup H_{1,2} \cup f_\sigma)$, $T_1 = \partial M_1$ (2-torus)
- T_3 be the 2-torus component of $\partial N(Y \cup f_1 \cup f_2)$ that contains A_1;
- M_3 be the component of W whose boundary is the torus T_3;
- $B_1 = T_1 - A_1$ and $B_3 = T_3 - A_1$ (see Figure 6).

If S is a surface in $M(\beta)$, we define $n(S)$ to be the number of intersections between S and the knot K_β. As an example, we have $n(\widehat{T})$=t.

Let A_2 be the annulus $\widehat{T} - A_1$. Then $t = n(A_1) + n(A_2)$. Moreover, since $n(T_1) = 2n(A_1) - 2$ and $n(\partial W) = 2n(A_1) - 4$, then we have

$$n(B_1) = n(A_1) - 2 \text{ and } n(B_3) \le n(A_1) - 4.$$

Let $Z_1 = \overline{X_B - M_1}$ and $Z_3 = \overline{X_W - M_3}$. Then

$$\partial Z_1 = B_1 \cup A_2 \text{ and } \partial Z_3 = B_3 \cup A_2.$$

Lemma (3.3.6). *If $\widehat{T}$ is an essential minimal 2-torus in $M(\beta)$, then Z_1 and Z_3 are solid tori.*

Proof. Since $n(\partial Z_1) = n(B_1) + n(A_2) = n(A_1) - 2 + n(A_2) < t$, then ∂Z_1 compresses in $M(\beta)$. But $M(\beta)$ is irreducible, implying that ∂Z_1 bounds a solid torus. Since $\widehat{T}$ is essential, ∂Z_1 must bound a solid torus in the side that

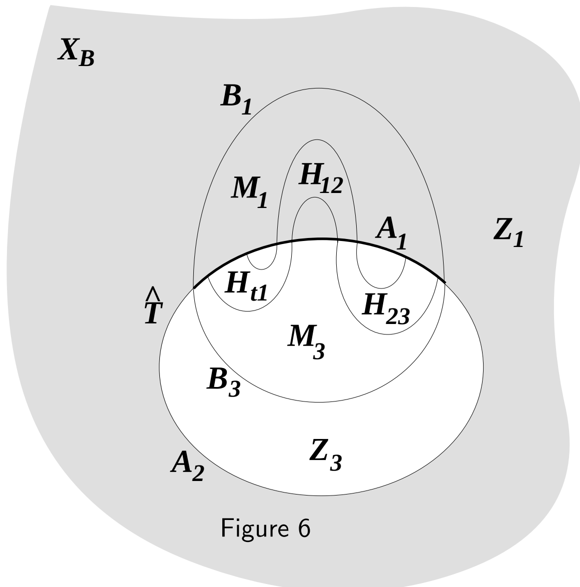


Figure 6

does not contain $\widehat{T}$. This means Z_1 is a solid torus. Similarly, since $n(\partial Z_3) \leq n(A_1) - 4 + n(A_2)$, Z_3 is a solid torus. ☐

Applying the same argument, we can prove the following.

LEMMA (3.3.7). *If $n(T_i) < t$ then M_i is a solid torus, where $i = 1$ or 3.*

LEMMA (3.3.8). *If $\widehat{T}$ is an essential minimal 2-torus in $M(\beta)$ then the two following assertions are true :*
i) *$n(A_1) \geq n(A_2) + 2$;*
ii) *if M_1 is a solid torus then $n(A_1) \geq n(A_2) + 4$.*

Proof. i) Suppose that $n(A_1) < n(A_2) + 2$, then $n(T_1) < t$; therefore, M_1 is a solid torus (Lemma (3.3.7)). Moreover, we have also that $n(T_3) < t$, which implies that M_3 is also a solid torus (Lemma (3.3.7)). We have

$X_B = M_1 \cup_{B_1} Z_1$, and

$X_W = M_3 \cup_{B_3} Z_3$.

It follows that X_W and X_B are both the union of two solid tori along an annulus. Since $\widehat{T}$ is essential, X_B and X_W must be Seifert fibered spaces which are not solid tori. Thus the core of B_i turns at least twice around the cores of M_i and Z_i respectively, for $i = 1$ and 3.

Let $T' = B_1 \cup B_3$. Then T' is an essential 2-torus. Indeed, we can decompose $M(\beta)$ along T' in the following way :

$$M(\beta) = (M_1 \cup_{A_1} M_3) \bigcup_{T'} (Z_1 \cup_{A_2} Z_3).$$

The core of B_1 turns at least twice around the core of M_1, and the core of B_3 turns at least twice around the core of M_3; thus the core of A_1 turns at least twice around the cores of M_3 and M_1 respectively. Therefore, $M_1 \cup_{A_1} M_3$ is a Seifert fibered space over a disc, with two exceptional fibers, which are respectively the core of M_1 and the core of M_3. In the same way, $Z_1 \cup_{A_2} Z_3$ is a Seifert fibered space over a disc, with two exceptional fibers, which are the cores of Z_1 and Z_3 respectively. Thus T' is essential in $M(\beta)$. Since $n(T') = 2n(A_1) - 6 < t$, then $\widehat{T}$ is not a minimal essential 2-torus in $M(\beta)$.

ii) Now, if M_1 is a solid torus and $n(A_1) < n(A_2)+4$, then M_3 is again a solid torus; we can repeat the same argument as above, and $n(T') = 2n(A_1) - 6 < t$ gives the same conclusion. □

(3.4) Constructions from Σ and σ_x. We keep the previous notations.

LEMMA (3.4.1). *The graph $G_T(\Sigma)$ lies in an essential annulus A_1 in $\widehat{T}$, and $G_T(\sigma_x)$ lies in the annulus $A_2 = \widehat{T} - A_1$.*

Proof. Let σ be a Scharlemann cycle in D_Σ. By Lemma (3.1), the edges of σ do not lie in a disc of $\widehat{T}$; and similarly for the edges of σ_x.

Recall that the edges of σ join v_1 to v_2. Let e_1, e_2 be two edges in the cycle σ so that the simple closed curve $\gamma = e_1 \cup e_2$, obtained by shrinking the vertices v_1 and v_2 to points, is essential on $\widehat{T}$. If we do the same with the Scharlemann cycle σ_x, we obtain γ', an essential simple closed curve on $\widehat{T}$, disjoint from γ. Thus γ and γ' are parallel on $\widehat{T}$. Therefore, we may assume that the graph $G_T(\Sigma)$ lies in an essential annulus A_1 in $\widehat{T}$, and $G_T(\sigma_x)$ lies in the annulus $A_2 = \widehat{T} - A_1$. □

Recall that $W = Y \cup f_1 \cup f_2$. Let :

f_{σ_x} be the disk-face bounded by σ_x in G_F;
$M_2 = N(A_2 \cup H_{x,x+1} \cup f_{\sigma_x})$, $T_2 = \partial M_2$ (2-torus);
$B_2 = T_2 - A_2$.

Note that :

$$n(T_2) = 2n(A_2) - 2 \text{ and } n(B_2) = n(A_2) - 2.$$

Let $Z_2 = \overline{X - M_2}$, where X is the side X_W or X_B which contains M_2. Then $\partial Z_2 = B_2 \cup A_1$.

LEMMA (3.4.2). *If M_1 and M_3 are both solid tori, then $\widehat{T}$ is not an essential minimal 2-torus in $M(\beta)$.*

Proof. Assume that M_1 and M_3 are both solid tori. Thus,
$X_B = M_1 \cup_{B_1} Z_1$, and
$X_W = M_3 \cup_{B_3} Z_3$,
meaning X_W and X_B are both the union of two solid tori along an annulus. Since $\widehat{T}$ is essential, X_B and X_W are Seifert fibered spaces. The core of B_i turns at least twice around the cores of M_i and Z_i respectively, for $i = 1$ and 3.

We consider two cases according to whether M_2 lies in X_W or X_B. Then M_2 lies in Z_3 or Z_1 respectively, which implies M_2 is a solid torus.

First, assume that $M_2 \subset X_W$, so M_2 lies in Z_3. Following the proof of Lemma (3.3.8) we obtain that the 2-torus $T' = B_1 \cup B_2$ is an essential 2-torus, which satisfies $n(T') < t$.

Now, assume that $M_2 \subset X_B$, then it lies in Z_1. Following the proof of Lemma (3.3.8), we obtain that the 2-torus $T'' = B_3 \cup B_2$ is an essential 2-torus, which satisfies $n(T'') < t$.

In both cases $\widehat{T}$ is not a minimal essential 2-torus in $M(\beta)$. □

Lemma (3.4.3). *If M_1 is a solid torus and f_{σ_x} lies in X_W, then $\widehat{T}$ is not an essential minimal 2-torus in $M(\beta)$.*

Proof. Assume that $\widehat{T}$ is an essential minimal 2-torus in $M(\beta)$. If f_{σ_x} lies in X_W, then the 2-torus $T' = B_2 \cup A_1$ bounds a solid torus V' in X_W, because $n(T') < t$. Since $M_3 \subset V'$, we obtain that M_3 is also a solid torus. The result follows by Lemma (3.4.2). □

Lemma (3.4.4). *If M_3 is a solid torus and f_{σ_x} lies in X_B, then $\widehat{T}$ is not an essential minimal 2-torus in $M(\beta)$.*

Proof. It is the same argument, by symmetry. Assume that $\widehat{T}$ is an essential minimal 2-torus in $M(\beta)$. If f_{σ_x} lies in X_B, then the 2-torus $T' = B_2 \cup A_1$ bounds a solid torus V' in X_B, because $n(T') < t$. Since $M_1 \subset V'$, we obtain that M_1 is also a solid torus. The result follows by Lemma (3.4.2). □

Proposition (3.4.5). *Let $\widehat{T}$ be a minimal essential 2-torus. Assume that $t =$ 6 and thet G_F contains a generalized Scharlemann cycle Σ and a Scharlemann cycle with disjoint label sets. Let σ be a Scharlemann cycle of Σ, and A be an annulus in $\widehat{T}$ such that $G_T(\sigma)$ lies in A, then $n(A) > 2$ (see Figure 5).*

Proof. Assume that $\widehat{T}$ is a minimal essential 2-torus. We keep the previous notations.

By Lemma (3.3.8), $n(A_2) \leq 2$. But A_2 contains v_x and v_{x+1}, thus $n(A_2) =$ 2. It follows that $n(A_1) = 4$, and $n(T_3) = 4$. Then M_3 is a solid torus (Lemma (3.3.7)). Let C_1 be a minimal annulus in A_1 which contains $G_T(\sigma)$, and $N_1 = N(C_1 \cup H_{1,2} \cup f_\sigma)$.

If $n(C_1 = 2)$ then $n(\partial N_1) = 2$ and so N_1 is a solid torus. But $M_1 = N_1 \cup N(A_1 - C_1)$ is isotopic to N_1. Therefore, M_1 is also a solid torus, which contradicts Lemma (3.4.2). □

Corollary (3.4.6). *If $t = 6$ and $\widehat{T}$ is a minimal essential 2-torus, then G_F cannot contain two quasi-disjoint generalized Scharlemann cycles.*

Proof. We assume that G_F contains two quasi-disjoint generalized Scharlemann cycles Σ and Σ_x. After changing the labelling if necessary, we may assume that the label set of Σ is $\{6, 1, 2, 3\}$ and the label set of Σ_x is $\{x-1, x, x+1, x+2\}$, with $\{x, x+1\} \cap \{6, 1, 2, 3\} = \emptyset$ so $x = 4$.

Let σ (resp. σ_4) be a Scharlemann cycle of Σ (resp. Σ_4). Let C_1 (resp. C_4) be a minimal annulus which contains $G_T(\sigma)$ (resp. $G_T(\sigma_4)$). By Lemma (3.2), $n(C_1) = 2$ or $n(C_4) = 2$. Therefore, the contradiction follows by Proposition (3.4.5). □

(3.5) Proof of Theorem (1.1). We assume that G_F contains two quasi-disjoint generalized Scharlemann cycles Σ and Σ_x. After changing the labelling if necessary, we may assume that the label set of Σ is $\{t, 1, 2, 3\}$ and the label set of Σ_x is $\{x-1, x, x+1, x+2\}$, with $\{x, x+1\} \cap \{t, 1, 2, 3\} = \emptyset$.

We keep the previous notations. Let σ_x be a Scharlemann cycle of G_F in D_{Σ_x}. Let $G_T(\Sigma_x)$ be the subgraph of G_T whose edges correspond to the edges of Σ_x, and whose vertices are the vertices v_{x-1}, v_x, v_{x+1} and v_{x+2}.

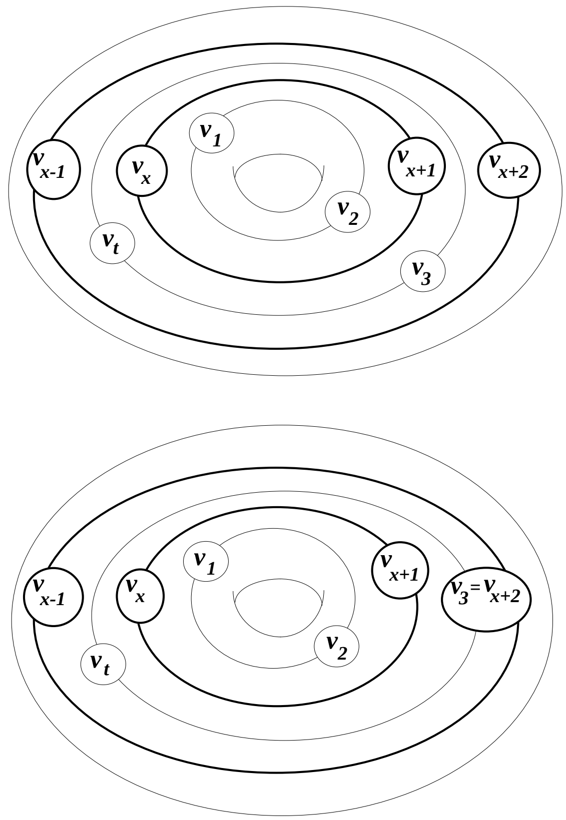


Figure 7

By Corollary (3.4.6), we may assume that $t \geq 8$. We want to prove that $\widehat{T}$ is not a minimal essential 2-torus in $M(\beta)$.

By Lemma (3.2) $G_T(\Sigma)$ lies in an annulus A_1 in $\widehat{T}$; similarly $G_T(\Sigma_x)$ lies in an annulus $A_{1,x}$ in $\widehat{T}$. Note that the if σ (resp. σ_x) is a Scharlemann cycle of Σ (resp. Σ_x) then $G_T(\sigma)$ (resp. $G_T(\sigma_x)$) is disjoint to $A_{1,x}$ (resp. A_1).

There are three cases. One is the case where the interior of A_1 and $A_{1,x}$ are disjoint; the second is that $A_1 \cap A_{1,x}$ is a disk with one vertex; the last is the case as in Figure 7.

First, we consider the two former cases. Then $A_{1,x} = A_2 \cup E$, where E is the empty set or a disc in A_1 which contains exactly one vertex. So, changing the labelling if necessary, we may assume that $n(A_1) \leq n(A_2) + 1$. Indeed, since $t \geq 8$, the label sets have at most one common label. Thus $\widehat{T}$ is not a minimal essential 2-torus, by Lemma (3.3.8).

To complete the proof of Theorem (1.1), we have to consider the case where the reduced graph of $G_T(\Sigma) \cup G_T(\Sigma_x)$ corresponds to one of the graphs on Figure 7. We assume for contradiction that $\widehat{T}$ is minimal.

We choose for A_1 a minimal annulus containing $G_T(\Sigma)$ and disjoint from the edges of any Scharlemann cycle in D_{Σ_x} (see Figure 8a).

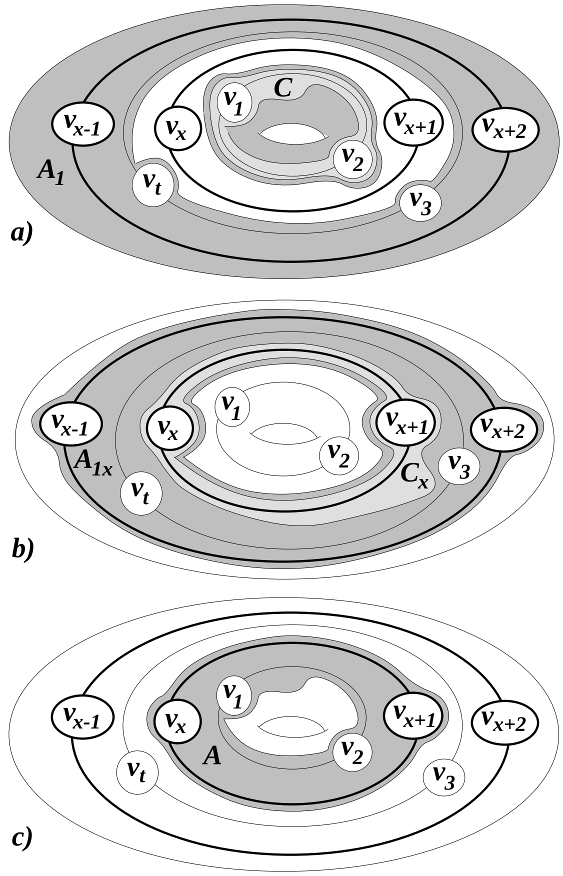


Figure 8

Similarly, let A_{1x} be a minimal annulus containing $G_T(\Sigma_x)$ and disjoint from the edges of any Scharlemann cycle in D_Σ. Let $A_{2x} = \widehat{T} - A_{1x}$.

Now, let C be an annulus in A_1 containing $G_T(\sigma)$ and disjoint from the $\{t, 3\}$-edges. Similarly, let C_x be an annulus in A_{1x} containing $G_T(\sigma_x)$ and disjoint from the $\{x-1, x+2\}$-edges (see Figure 8b).

Let $M_x = N(A_{1x} \cup H_{x,x+1} \cup f_{\sigma_x})$. Let $M'_x = N(C_x \cup H_{x,x+1} \cup f_{\sigma_x})$ and $M'_1 = N(C \cup H_{1,2} \cup f_\sigma)$; see Figure 9.

Note that C_x is disjoint from A_1.

If f_{σ_x} lies in X_B (see Figure 9a) then $M'_x \subset X_B$. But M'_x is disjoint from M_1, thus $M'_x \subset Z_1$. If f_{σ_x} lies in X_W (see Figure 9b) then $M'_x \subset X_W$; but M'_x is disjoint from M_3, so $M'_x \subset Z_3$. Therefore, in both cases M'_x is a solid torus, by Lemma (3.3.6).

By symmetry, we obtain that M'_1 is also a solid torus, which implies that M_1 is a solid torus, since it is isotopic to M_1. Therefore, f_{σ_x} lies in X_B by Lemma (3.4.3). Let A be a minimal annulus in $\widehat{T}$ (see Figure 8c) containing both $G_T(\sigma)$ and $G_T(\sigma_x)$. We have $A \subset A_{2x} \cup_\partial B$, where B is a small annulus

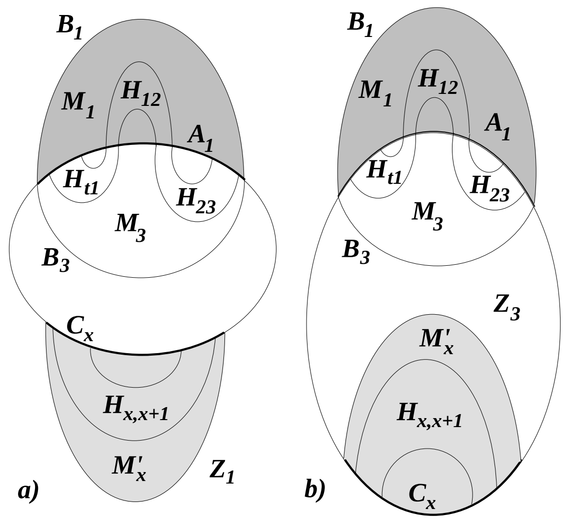


Figure 9

containing v_x and v_{x+1}. Then $n(A) \leq n(A_{2x}) + 2$. By Lemma (3.3.8), $n(A_{1x}) \geq n(A_{2x}) + 4$, so $n(A_{2x}) \leq t/2 - 2$. Then $n(A) \leq t/2$.

Let $Z = N(A \cup H_{1,2} \cup H_{x,x+1} \cup f_\sigma \cup f_{\sigma_x})$. Then Z is a Seifert fibered space over a disk with two exceptional fibers (the cores of M_1' and M_x').

Therefore ∂Z is an essential 2-torus, since $M(\beta) - Z$ contains $\widehat{T}$. Since $n(\partial Z) \leq t - 4$, it contradicts the fact that the essential 2-torus $\widehat{T}$ is minimal in $M(\beta)$. This completes the proof of Theorem (1.1).

Acknowledgements

We would like to thank the referee for helpful comments, which lead to make some parts more readable and some proofs more simple.

Received June 17, 2003

Final version received February 27, 2004

Université de Provence
CMI 39 rue Joliot-Curie
13453 Marseille Cedex 13
FRANCE
matignon@cmi.univ-mrs.fr

elsa.mayrand@freesurf.fr

References

[1] S. Boyer and X. Zhang, *Reducing Dehn filling and Toroidal Dehn filling*, Topology Appl. **68** (1996), 385–409.

[2] M. Culler, C. McA. Gordon, J. Luecke and P. Shalen, *Dehn surgery on knots*, Ann. Math. **125** (1987), 237–300.

[3] A. Deruelle and D. Matignon, *Lens spaces and generalized Scharlemann cycles*, Preprint.

[4] M. Domergue and D. Matignon, *Dehn surgeries and P^2-reducible 3-manifolds*, Topology and its App. **72** (1996), 135–148.

[5] C. McA. Gordon, *Boundary slopes of punctured tori in 3-manifolds*, Trans. Amer. Math. Soc. **350** (1998), 1713–1790.

[6] C. McA. Gordon, *Dehn filling: a survey*, Knot Theory, Banach Center Pub. **42**, Institute of Math. Warszawa (1998), 129–144.

[7] C. McA. Gordon, *Small surfaces and Dehn filling*, Geometry and Topology Monographs **2** (1999), 177-199.

[8] C. McA. Gordon and J. Luecke, *Dehn surgeries on knots creating esential tori*, I, Comm. Anal. Geom. **4** (3) (1995), 597–644.

[9] C. McA. Gordon and J. Luecke, *Reducible manifolds and Dehn surgery*, Topology **35** (1996), 385–409.

[10] J. A. Hoffman, *There are no strict great x-cycles after a reducing or P^2 surgery on a knot*, J. Knot Theory and its Ramifications **7** (5) (1998), 549–569.

[11] K. Ichihara and M. Teragaito, *Klein bottles surgery and genera of knots*, Pacific J. Math. **210** (2) (2003), 317–333.

[12] S. Lee, S. Oh and M. Teragaito, *Reducing Dehn fillings and x-faces*, Proc. of the conference "On Heegaard Splittings and Dehn surgeries of 3-manifolds", RIMS#1229, Kyoto Univ, June 11-June 15 (2001) 50–65.

[13] S. Lee, S. Oh and M. Teragaito, *Reducing Dehn fillings and small surfaces*, preprint (available from Math. ArXiv).

[14] M. Scharlemann, *Unknotting number one knots are prime*, Invent. Math. **82** (1985), 37–55.

[15] W. Thurston, *Three dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc **6** (1982), 357–381.

[16] W. Thurston *Hyperbolic structures on 3-manifolds, I: Deformation of acylindrical manifolds*, Ann. of Math. **124** (1986), 203–246.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

SPLITTING OF CERTAIN SINGULAR FIBERS OF GENUS TWO

YUKIO MATSUMOTO

Dedicated to Professor Francisco Javier Gonzáles Acuña on his sixtieth birthday.

Abstract. This is a detailed account of the results announced in our previous paper [Y.Matsumoto, Lefschetz fibrations of genus two - a topological approach, in the Proceedings of the 37th Taniguchi Symposium on Topology and Teichmüller spaces, World Scientific (1996)]. Via computer calculations, we will observe how certain genus two singular fibers of specific types split into Lefschetz type singular fibers, which are atomic in the sense of G. Xiao and M. Reid [14]. Also, we will give explicitly the positions of the vanishing cycles corresponding to the atomic fibers.

1. Introduction

By the *splitting* of a singular fiber we mean the phenomenon that a singular fiber in a holomorphic one-parameter family of Riemann surfaces splits into several less complicated singular fibers when the family is modified by a certain perturbation. Following G. Xiao and M. Reid [14], we will call a singular fiber that does not split any further an *atomic fiber*. In the case of genus two, atomic fibers are now completely understood thanks to the work of Horikawa [7], Xiao [18], Reid [14], Persson [13], and Arakawa and Ashikaga [1]. Arakawa and Ashikaga [1] extended the investigation to hyperelliptic families of genus ≥ 2. More recently, Takamura [15, 16, 17] has started a systematic study on splitting of more general singular fibers, not necessarily hyperelliptic.

Our study in this paper, however, is very restricted. We will be confined to two concrete examples of singular fibers. We will take up two specific types of genus two singular fibers and, via computer calculations, observe concretely how they split into atomic fibers.

We are interested not only in splittability of these singular fibers but also in the precise positions of the vanishing cycles corresponding to those atomic fibers that occur at the splitting. Since the topological monodromy of the original singular fiber is decomposed into a product of the right-handed Dehn twists [6] along the vanishing cycles of the atomic fibers, the knowledge of the precise positions of the vanishing cycles will give a precise decomposition formula of the original monodromy homeomorphism. Thus, the splitting of singular fibers is

2000 *Mathematics Subject Classification*: Primary 14D06; Secondary 14D05, 57M50, 55R55.

Keywords and phrases: Singular fibers, splitting, Lefschetz fibrations, vanishing cycles, monodromy.

This research is supported by Grant-in-Aid for Scientific Research (No.12440013),JSPS.

expected to provide a heuristic method to find relations in the mapping class group.

In the case of genus two, it is known ([1]) that there are precisely two topological types of atomic fibers. Both of them are Lefschetz type singular fibers shown in Figure (1). We will call them an atomic fiber of type I, and of type II, respectively. (This notation is different from that in [7] or in [1].) An atomic fiber of type I is obtained by pinching a *non-separating* simple closed curve on a genus two Riemann surface into a point, and that of type II is obtained by pinching a *separating* simple closed curve. The singular point of these fibers is a node. Thus a singular fiber of type II consists of two tori intersecting transversely in a point.

The precise statements of our main results will be given in the next section.

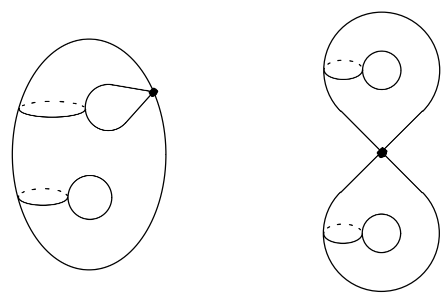


Figure 1. Genus two atomic fibers of types I (left) and II (right)

The author was informed by L. Balke [2] that our splitting in Theorem (2.1) below can be neatly reconstructed by an algebraic-geometric method. The author is grateful to him for communicating his construction. All the figures in this paper were drawn by I. Hasegawa, K. Tanaka and K. Yoshida. The author thanks them for their beautiful work and kind help. Finally but not at all least, the author greatly appreciates the referees' careful reading and useful comments, which improved the paper very much.

2. Main results

First we will describe the singular fibers which we want to study.

Let Σ_2 denote an oriented closed surface of genus two, and consider an involution $\omega : \Sigma_2 \to \Sigma_2$ shown in Figure 2. Given a complex structure on Σ_2, we may assume that ω is holomorphic. Let $\Delta = \{\xi||\xi| < 1\}$ be the unit disk on the complex plane. To obtain a singular fiber having the topological monodromy ω, consider the quotient $\Delta \times \Sigma_2/(-1) \times \omega$ and blow up the two singular points. The resulting complex manifold V fibers over a disk $D = \Delta/(-1)$ with the projection $f : V \to D$ induced by the first projection $\Delta \times \Sigma_2/(-1) \times \omega \to \Delta/(-1)$. The family $f : V \to D$ has a single singular fiber $f^{-1}(0)$ and its topological monodromy is ω. According to [11], the topological monodromy around a singular fiber determines the topological type of (the fibered neighborhood of) the singular fiber. We denote the singular fiber $f^{-1}(0)$ (or rather its topological type) by F^ω.

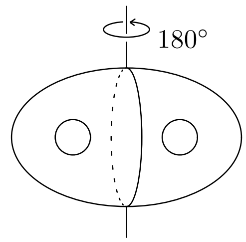


Figure 2. The involution $\omega : \Sigma_2 \to \Sigma_2$

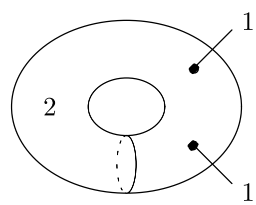


Figure 3. Singular fiber F^ω

The singular fiber F^ω consists of a torus of multiplicity 2 and two 2-spheres of multiplicity 1. Each 2-sphere intersects the torus transversely in a point (see Figure (3)). This is the first singular fiber we shall study.

The second one is the singular fiber F^ι shown in Figure (4). This fiber consists of seven 2-spheres intersecting transversely as shown in the figure. The monodromy corresponding to this singular fiber is the hyperelliptic involution $\iota : \Sigma_2 \to \Sigma_2$ (see Figure (5)). The construction of F^ι is similar to that of F^ω.

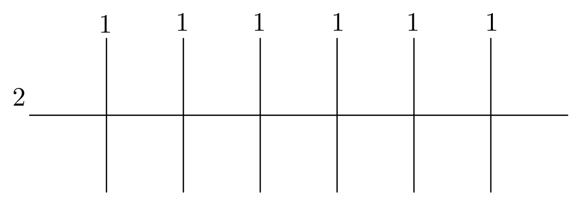


Figure 4. Singular fiber F^ι

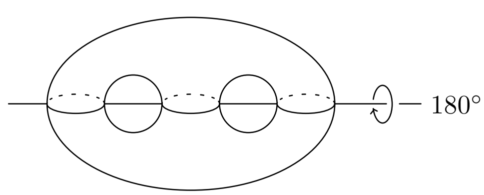


Figure 5. The hyperelliptic involution ι

The following theorem is our first main result. In this theorem the family $\phi : N \to D$ is topologically equivalent to the family $f : V \to D$ constructed above, but their complex structures are not necessarily the same.

THEOREM (2.1). *There exists a holomorphic family of compact genus two Riemann surfaces over a disk, $\phi : N \to D$, having a single singular fiber over the origin $0 \in D$ whose topological type is F^{ω}, such that if one perturbs the family in a certain way by a real parameter ϵ then, in the perturbed family $\phi_\epsilon : N_\epsilon \to D$, the singular fiber F^{ω} disappears, and in place of it four atomic fibers occur. The positions of their vanishing cycles $\beta_1, \beta_2, \beta_3, \beta_4$ are as shown in Figure (6).*

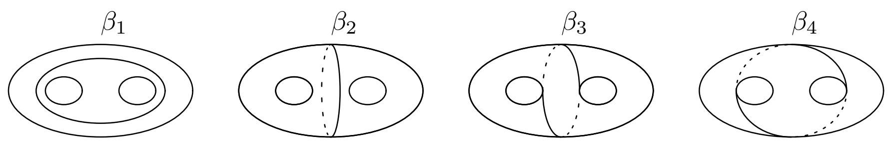


Figure 6. Vanishing cycles $\beta_1, \beta_2, \beta_3, \beta_4$

As a corollary, the monodromy ω is decomposed as follows:

$$\omega = \beta_1\beta_2\beta_3\beta_4. \tag{2.2}$$

Note that here we use identical notation for a simple closed curve on Σ_2 and the right-handed Dehn twist along the curve. Also note that the mapping class group $\mathcal{M}_2$ is assumed to act on Σ_2 from the *right*: the composition $\beta_1\beta_2\beta_3\beta_4$ means that first we apply β_1 and then β_2, and so on. Among these vanishing cycles, only β_2 is separating. Thus the atomic fiber corresponding to β_2 is of type II. The other three atomic fibers are of type I, and the splitting of Theorem (2.1) is simply written as

$$F^{\omega} \Rightarrow 3I + II. \tag{2.3}$$

This splitting seems to be known to specialists except for the precise positions of vanishing cycles (cf. [13]).

Although Theorem (2.1) is merely an experimental observation (and the author has a little hesitation about calling it "a theorem"), it has turned out to be quite useful. For example, in [10], we made use of this splitting to construct a Lefschetz fibration of genus two

$$S^2 \times T^2 \# 4\overline{\mathbb{C}P^2} \to S^2 \tag{2.4}$$

whose singular fibers are of types $6I + 2II$ and whose total monodromy is $(\beta_1\beta_2\beta_3\beta_4)^2 = 1$. With this fibration, we were able to calculate the local signature $\sigma(II)$ of a type II atomic fiber [10]. That is, substituting the known values $\sigma(I) = -\frac{3}{5}$ (which was known from another example) and $\mathrm{Sign}(S^2 \times T^2 \# 4\overline{\mathbb{C}P^2}) = -4$ in the local signature formula

$$\mathrm{Sign}(S^2 \times T^2 \# 4\overline{\mathbb{C}P^2}) = 6\sigma(I) + 2\sigma(II) \tag{2.5}$$

we obtained

$$\sigma(II) = -\frac{1}{5}.$$

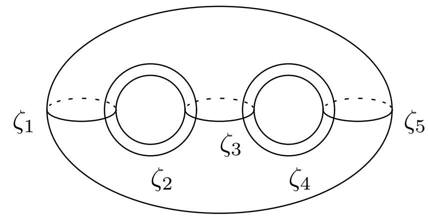


Figure 7. Standard curves on Σ_2

Using algebraic methods, Endo [5] extended this result and calculated local signature of singular fibers in hyperelliptic Lefschetz fibrations of genus $g \geq 3$. Arakawa and Ashikaga [1] studied the local signature from an algebraic-geometric viewpoint using the Horikawa index. Moreover, Ozbagci and Stipsicz [12], starting from the fibration (2.4) and applying Gompf's theorem (see [6]) on the existence of symplectic structures on Lefschetz fibrations, constructed infinitely many examples of closed symplectic 4-manifolds which do not have the homotopy type of any complex surface. The decomposition (2.2) was extended to higher genera by Cadavid [4] and Korkmaz [9].

Our second main result is the following

THEOREM (2.6). *There exists a holomorphic family of compact genus two Riemann surfaces over a disk, $\varphi : M \to D$, having a single singular fiber over the origin $0 \in D$ whose topological type is F^ι, such that if one perturbs the family in a certain way by a real parameter ϵ, then in the perturbed family $\varphi_\epsilon : M_\epsilon \to D$ the singular fiber F^ι disappears, and in its place 10 atomic fibers of type I occur. Their vanishing cycles are*

$$(\zeta_1, \zeta_2, \zeta_3, \zeta_4, \zeta_5, \zeta_5, \zeta_4, \zeta_3, \zeta_2, \zeta_1)$$

where ζ_i, $i = 1, 2, \ldots, 5$, are the standard simple closed curves on Σ_2 given in Figure (7).

Both in Theorems 1 and 2, we tacitly assume that a general fiber F_{t_0} is fixed as a reference fiber in the family of Riemann surfaces, and that a set of loops $\{l_1, l_2, \ldots, l_s\}$ on $D - \{\text{critical values}\}$ corresponding to the occurring atomic fibers $\{F_1, F_2, \ldots, F_s\}$ are chosen as follows (see Figure (8)): The loop l_i corresponding to an atomic fiber F_i starts from the locus t_0 of the reference fiber F_{t_0}, follows a path γ_i, and reaches a point on the boundary of a small disk containing the critical value b_i of F_i, then moves counter-clockwise along the boundary of this small circle, and finally comes back to t_0 along γ_i^{-1}. We assume that the loops l_i $(i = 1, 2, \ldots, s)$ are mutually disjoint except at the base point t_0, and taking a small disk D_0 centered at t_0, we assume that the paths $\gamma_1, \gamma_2, \ldots, \gamma_s$ intersect the boundary ∂D_0 counter-clockwise in this order.

The vanishing cycle corresponding to an atomic fiber F_i is considered to be a simple closed curve C_i on the reference fiber F_{t_0}. This cycle C_i shrinks to the nodal point on the atomic fiber F_i as one "moves" the reference fiber F_{t_0} along the path γ_i to F_i. The corresponding monodromy, which is the Dehn twist along C_i, is the returning diffeomorphism obtained by moving the fiber F_{t_0} along the loop l_i. The order and the positions of the vanishing cycles $(\beta_1, \beta_2, \beta_3, \beta_4)$ in

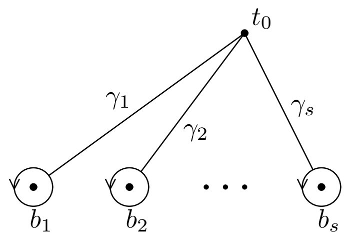


Figure 8. Loops $l_1, l_2, \ldots, l_s$

Theorem 1 and $(\zeta_1, \zeta_2, \zeta_3, \zeta_4, \zeta_5, \zeta_5, \zeta_4, \zeta_3, \zeta_2, \zeta_1)$ in Theorem (2.6) assume certain choices for the loops l_i, as indicated above.

By Theorem (2.6), the topological monodromy ι of the singular fiber F^ι decomposes into

$$\iota = \zeta_1\zeta_2\zeta_3\zeta_4\zeta_5\zeta_5\zeta_4\zeta_3\zeta_2\zeta_1. \tag{2.7}$$

This is of course a well-known relation in the mapping class group of genus two (see [3]). Ito [8] extended Theorem (2.6) to higher genera.

3. Construction of a fibered neighborhood of the singular fiber F^ω

The construction of the family $f : V \to D$ given in §2 is simple, but for the purpose of computer calculation it is not necessarily adequate, because the complex structure of V is not explicitly described. Thus we must construct a family $\phi : N \to D$ which is topologically equivalent to $f : V \to D$ using concrete equations.

We will start with a torus in the complex projective plane $\mathbb{C}P^2$, defined by the following cubic homogeneous polynomial

$$x^2z - y^3 - z^3 = 0. \tag{3.1}$$

The affine space $\mathbb{C}P^2 - \{z = 0\}$ will be identified with the (complex) xy-plane. Define a polynomial $f(x, y)$ as follows:

$$f(x, y) = x^2 - y^3 - 1. \tag{3.2}$$

Then in the xy-plane the torus (3.1) is given by $f(x, y) = 0$. This is actually a punctured torus.

Let N' be a tubular neighborhood of the punctured torus $f = 0$ in the xy-plane with "constant thickness":

$$|f(x, y)| < \delta, \tag{3.3}$$

where δ is a positive constant. Note that in the projective plane $\mathbb{C}P^2$, N' is no longer of constant thickness: it becomes thinner and thinner as it approaches the point $(1 : 0 : 0)$ at infinity.

f is a well-defined function on N'.

We introduce another well-defined function ϕ on N' by setting as

$$\phi(x, y) = yf(x, y)^2. \tag{3.4}$$

The punctured torus $f = 0$ and the x-axis ($y = 0$) intersect in the two points $x = \pm 1$ (cf. (3.2)). Thus the divisor $\phi = 0$ in N' is a union of a punctured torus and two disks D_1, D_2, as shown in Figure (9).

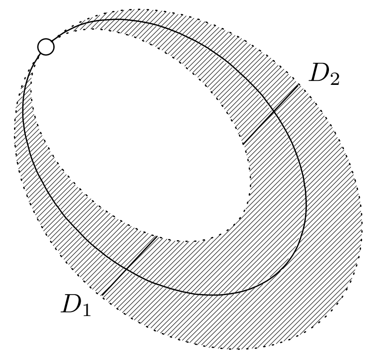


Figure 9. The divisor $\phi = 0$

We attach two 2-handles H_1 and H_2 to N' along the boundaries of the 2-disks D_1, D_2 so that $\phi : N' \to \mathbb{C}$ extends to a well-defined function $N' \cup H_1 \cup H_2 \to \mathbb{C}$. This is explained more precisely as follows. Let p_1 be the intersection point of the disk D_1 and the punctured torus $f = 0$. Introduce local coordinates (s, t) whose origin is p_1. We assume that D_1 is locally defined by $t = 0$, and $f = 0$ by $s = 0$. In these coordinates, ϕ is locally given by

$$\phi = s^2 t. \tag{3.5}$$

Take a polydisk $\Delta^2 = \{(\sigma, \tau) \in \mathbb{C}^2 | |\sigma| < \delta', |\tau| < \delta''\}$, and glue Δ^2 to N' by setting

$$\sigma = s^{-1}, \ \ \tau = s^2 t. \tag{3.6}$$

Then the polydisk Δ^2 is attached to N' as a 2-handle H_1, and $\phi : N' \to \mathbb{C}$ extends to $N' \cup H_1 \to \mathbb{C}$. The 2-disk D_1 together with the core disk ($\tau = 0$) of the 2-handle H_1 make a 2-sphere S^2.

We perform the same construction at the intersection point p_2 of the other disk D_2 and the punctured torus $f = 0$, that is, we attach a 2-handle H_2 to N' using another polydisk Δ^2. Then the 2-disk D_2 closes up to a 2-sphere.

We have attached two 2-handles H_1, H_2 to N'. Let us denote the resulting manifold by N'':

$$N'' = N' \cup H_1 \cup H_2. \tag{3.7}$$

Then the map $\phi : N' \to \mathbb{C}$ extends to $N'' \to \mathbb{C}$. This extended map is denoted again by ϕ.

Observe that N'' is an open complex manifold, but no longer an open set in $\mathbb{C}^2$.

We have almost finished the construction of a fibered neighborhood of the singular fiber F^ω. But, at this stage, the singular fiber is a punctured torus (rather than a torus) of multiplicity 2 stuck with two 2-spheres, and it only lacks the point $(1 : 0 : 0) \in \mathbb{C}P^2$ to become F^ω.

A general fiber $\phi = k$ (k a non-zero constant) of $\phi : N'' \to \mathbb{C}$ is obtained by closing the surface $N' \cap \{\phi = k\}$ with two 2-disks, each being defined by $\tau = k$ in each polydisk Δ^2 (i.e., a disk parallel to the core of the 2-handle H_1 or H_2). This general fiber is essentially a 2-fold branched covering of the punctured torus $f = 0$ branched at the two points p_1 and p_2. Over the point at infinity $(1:0:0)$ the covering is unbranched, and the general fiber $\phi = k$ is a twice punctured surface of genus two (twice punctured since it lacks two points over $(1:0:0)$).

Now we will look at a neighborhood of the point $(1:0:0)$ more closely.

The affine space $\mathbb{C}P^2 - \{x = 0\}$ has complex coordinates (u, v), where

$$u = \frac{z}{x}, \quad v = \frac{y}{x}. \tag{3.8}$$

The origin of this coordinate system is the point $(1:0:0)$. Since the polynomial function f (defined on the xy-plane) is given in homogeneous coordinates by $f(x:y:z) = (\frac{x}{z})^2 - (\frac{y}{z})^3 - 1$, f is given in the uv-plane by the formula

$$u^{-2} - u^{-3}v^3 - 1 = u^{-3}(u - v^3 - u^3) \tag{3.9}$$

which we denote by $g(u, v)$. In the intersection $N' \cap uv$-plane, the function ϕ is equal to

$$\phi(u, v) = \frac{v}{u} g(u, v)^2. \tag{3.10}$$

From (3.3) we have that, on $N' \cap$ (a neighborhood of $(1:0:0)$),

$$|g(u, v)| < \delta. \tag{3.11}$$

The function $g(u, v)$ is not defined at the point $(1:0:0)$. But the complex curve $g = 0$ ($\Leftrightarrow u - v^3 - u^3 = 0$) is becoming tangent to the v-axis ($u = 0$) as it approaches to $(1:0:0)$ and, in the neighborhood of $(1:0:0)$, the curve $g = 0$ does not intersect neither the u-axis nor the v-axis. (Since the domain of the function $g(u, v)$ does not contain the origin $(u, v) = (0, 0)$, the curve $g = 0$ does not contain this point either.) Since N' becomes thinner and thinner near the origin and it has $g = 0$ as "core", we have $u \neq 0$ and $v \neq 0$ on N'. See Figure (10).

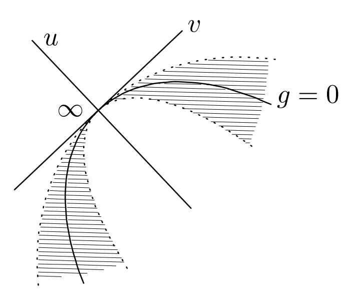


Figure 10. Situation near $(1:0:0)$

In order to extend the function $g(u, v)$ over the point at infinity, $(u, v) = (0, 0)$, we paste a new coordinate neighborhood $U = \{(\xi, \eta)\}$ to N' by assuming that

$U \cap N' = N' \cap$ (a neighborhood of $(1:0:0)$) and by defining the pasting map as follows:

$$\xi = v, \quad \eta = \frac{\sqrt{1-(1+g(u,v))u^2}}{v} g(u,v). \tag{3.12}$$

Recall that in the intersection of N' and a small neighborhood of $(u,v) = (0,0)$, we have $|g(u,v)| < \delta$ and $(u,v) \risingdotseq (0,0)$. Thus, in the same intersection, $u^2(1+g(u,v))$ is close to 0, and $\sqrt{1-(1+g(u,v))u^2}$ in (3.12) may be considered to be a well-defined complex number close to 1. The manifold N' does not contain the point $(1:0:0) \in \mathbb{C}P^2$, but the coordinate neighborhood U contains this point, namely, $(\xi, \eta) = (0,0)$.

Note that we do not regard U as an open set of $\mathbb{C}P^2$: we paste the two open manifolds U and N'', abstractly. In this pasting, N'', which was infinitely thin near $(1:0:0)$, recovers its finite thickness. Thus the pasting $N'' \cup U$ may be considered to be a kind of blow up process.

We have that

$$\begin{aligned}
\eta^2 &= \frac{1-(1+g(u,v))u^2}{v^2} g(u,v)^2 \\
&= \frac{1-(1+u^{-3}(u-v^3-u^3))u^2}{v^2} g(u,v)^2 \\
&= \frac{1-(u^2+1-\frac{v^3}{u}-u^2)}{v^2} g(u,v)^2 \\
&= \frac{v}{u} g(u,v)^2 \\
&= y f(x,y)^2 \\
&= \phi.
\end{aligned} \tag{3.13}$$

Thus, if we denote the manifold $N'' \cup U = N' \cup H_1 \cup H_2 \cup U$ by N, then ϕ extends to a well-defined function on N, denoted by the same symbol $\phi : N \to \mathbb{C}$. From (3.13), we have

$$\phi \mid U = \eta^2. \tag{3.14}$$

The divisor $\phi = 0$ in N is obtained from the divisor $\phi = 0$ in N'' (which was a punctured torus of multiplicity 2 stuck with two 2-spheres) by capping off the puncture with the 2-disk in U of multiplicity 2 defined by $\eta^2 = 0$. Thus we have obtained a torus of multiplicity 2 stuck with 2 spheres, that is, a singular fiber of type F^ω (see Figure (11)).

A general fiber of $\phi : N \to \mathbb{C}$ is obtained from a general fiber $\phi = t$ (t a non-zero constant) in N'' (which was a twice punctured surface of genus 2 as we remarked above) by capping off the punctures with two 2-disks in U defined by $\eta = \pm\sqrt{t}$ (see (3.13)). The resulting general fiber is a closed surface of genus 2.

Taking a small open disk D centered at 0 of $\mathbb{C}$, and denoting $\phi^{-1}(D)$ again by N, we have a holomorphic family of Riemann surfaces of genus 2, $\phi : N \to D$, which has a unique singular fiber of type F^ω over 0. The construction is now complete.

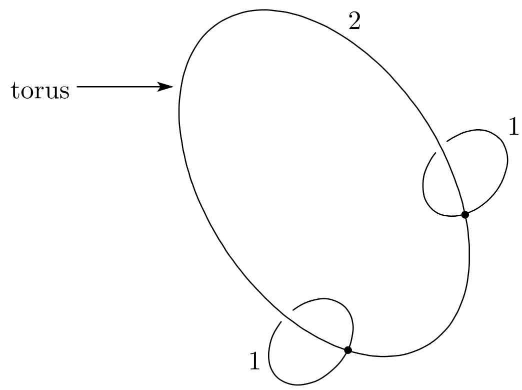


Figure 11. Divisor $\phi = 0$ in N

4. Perturbation of the projection $\phi : N \to D$

Recall that $N = U \cup N''$ and $N'' = N' \cup H_1 \cup H_2$. The attaching maps of the 2-handles H_1, H_2 are automatically determined by the requirement that the projection map $\phi : N' \to \mathbb{C}$ should extend to $N' \cup H_1 \cup H_2 \to \mathbb{C}$. Thus the main body of the information on the projection $\phi : N \to \mathbb{C}$ is contained in $\phi : U \cup N' \to \mathbb{C}$. In what follows, we will study this part of the projection closely.

The projection ϕ is given on $U \cup N'$ as follows:

$$\begin{aligned} \phi(x,y) &= yf(x,y)^2 \quad \text{on } N' \\ \phi(\xi,\eta) &= \eta^2 \qquad \text{on } U. \end{aligned} \tag{4.1}$$

We perturb $\phi : N' \to \mathbb{C}$ to ϕ_ϵ as follows:

$$\phi_\epsilon(x,y) = (yf(x,y) - \epsilon)f(x,y) \tag{4.2}$$

where ϵ is the parameter of perturbation, and is a non-zero small real number.

Let us examine the divisor $\phi_\epsilon = 0$ on $U \cup N'$.

On N', the divisor $\phi_\epsilon = 0$ has two irreducible components:

$$yf(x,y) - \epsilon = 0 \quad \text{and} \quad f(x,y) = 0. \tag{4.3}$$

These components do not intersect each other in N'.

On U, the perturbed map $\phi_\epsilon|U$ is automatically determined by $\phi_\epsilon|N'$:

$$\begin{aligned} \phi_\epsilon &= (yf(x,y) - \epsilon)f(x,y) \\ &= (\frac{v}{u}g(u,v) - \epsilon)g(u,v) \\ &= \frac{v}{u}g(u,v)^2 - \epsilon g(u,v) \\ &= \eta^2 - \frac{\epsilon v}{\sqrt{1-(1+g(u,v))u^2}}\eta \\ &= \eta(\eta - \frac{\epsilon\xi}{\sqrt{1-(1+g(u,v))u^2}}) \\ &\fallingdotseq \eta(\eta - \epsilon\xi). \end{aligned} \tag{4.4}$$

In the above, we used the fact that the $\sqrt{}$ in (4.4) is close to 1. Thus, in U, the divisor $\phi_\epsilon = 0$ consists approximately of the two (complex) lines $\eta = 0$ and $\eta - \epsilon\xi = 0$, which intersect transversely at the point $(\xi, \eta) = (0, 0)$.

Deforming the projection $\phi : N' \to \mathbb{C}$ to $\phi_\epsilon : N' \to \mathbb{C}$ necessarily changes the attaching maps of the 2-handles H_1, H_2, because we require that the projection $\phi_\epsilon : N' \to \mathbb{C}$ should extend on $N' \cup H_1 \cup H_2$.

We look at the attaching map of a 2-handle more closely. Previously the 2-handle H_1 was attached by the pasting map (3.6). Now we attach it by the following map:

$$\sigma = s^{-1}, \quad \tau = (ts - \epsilon)s. \tag{4.5}$$

(Recall that the 2-handle H_1 is a polydisk $\Delta^2 = \{(\sigma, \tau) | |\sigma| < \delta', |\tau| < \delta''\}$ glued to N'.) Then the projection $\phi_\epsilon | N'$ is extended on the polydisk by setting $\phi_\epsilon = \tau$, and the component $yf(x, y) - \epsilon = 0$ of the divisor $\phi_\epsilon | N' = 0$ is capped off by the disk $\tau = 0$ in Δ^2. The same thing happens in the other 2-handle H_2, where the component $yf(x, y) - \epsilon = 0$ is capped off again to make a punctured torus. The other component of $f(x, y) = 0$ was already a punctured torus. Thus, if we denote the extended projection by the same symbol ϕ_ϵ, the divisor $\phi_\epsilon = 0$ in N''_ϵ ($:= N' \cup_\epsilon H_1 \cup_\epsilon H_2$) is a union of two punctured tori. We denote the perturbed manifold by $N_\epsilon (:= U \cup N''_\epsilon)$. Of course, if $\epsilon = 0$, the N_ϵ coincides with the original manifold N: $N_0 = N$. We are assuming, however, that $\epsilon > 0$, and in this case N_ϵ and N are diffeomorphic, but have different complex structures.

The divisor $\phi_\epsilon = 0$ in N''_ϵ was a disjoint union of two punctured tori. In U, the punctures are capped off by the two disks $\eta = 0$ and (approximately speaking) $\eta - \epsilon\xi = 0$. Thus the divisor $\phi_\epsilon = 0$ in N_ϵ consists of two 2-tori intersecting each other transversely at the point $(\xi, \eta) = (0, 0)$. This is the "central" singular fiber of $\phi_\epsilon : N_\epsilon \to D$ over 0.

We will now look for the other singular fibers of ϕ_ϵ. For this purpose, we will study the critical points of ϕ_ϵ.

We compute on N'.

Recall that

$$\begin{aligned} \phi_\epsilon &= (yf(x, y) - \epsilon)f(x, y) \\ &= (y(x^2 - y^3 - 1) - \epsilon)(x^2 - y^3 - 1). \end{aligned} \tag{4.6}$$

Thus

$$\frac{\partial \phi_\epsilon}{\partial x} = 2x(2y(x^2 - y^3 - 1) - \epsilon) \tag{4.7}$$

and

$$\frac{\partial \phi_\epsilon}{\partial y} = (x^2 - y^3 - 1)^2 - 6y^3(x^2 - y^3 - 1) + 3\epsilon y^2. \tag{4.8}$$

Solving $\frac{\partial \phi_\epsilon}{\partial x} = 0$ we have

$$x = 0 \quad \text{or} \quad x^2 - y^3 - 1 = \frac{\epsilon}{2y}. \tag{4.9}$$

Substituting $x = 0$ in $\frac{\partial \phi_\epsilon}{\partial y} = 0$ we have

$$7y^6 + 8y^3 + 3\epsilon y^2 + 1 = 0, \tag{4.10}$$

while substituting $x^2 - y^3 - 1 = \frac{\epsilon}{2y}$ in $\frac{\partial \phi_\epsilon}{\partial y} = 0$ we have

$$\frac{\epsilon^2}{4y^2} = 0. \tag{4.11}$$

The latter is clearly impossible, for $\epsilon > 0$.

Therefore, every critical point is on the line $x = 0$.

If we put $\epsilon = 0$, then the equation (4.10) has six solutions $y = -1$, $(1/2)(1 \pm \sqrt{-3})$, $-\sqrt[3]{1/7}$, $(1/2\sqrt[3]{7})(1 \pm \sqrt{-3})$. Among them, the three solutions $y = -1$, $(1/2)(1 \pm \sqrt{-3})$ satisfy $\phi_0(0, y) = 0$, while the solutions $y = -\sqrt[3]{1/7}$, $(1/2\sqrt[3]{7})(1 \pm \sqrt{-3})$ do not.

Now we look for new-born singular fibers. Since ϵ is a sufficiently small positive real number in this case, we look for three solutions y of (4.10) which are close to -1, and $(1/2)(1 \pm \sqrt{-3})$.

Put $\epsilon = 0.1$. Solving $7y^6 + 8y^3 + 0.3y^2 + 1 = 0$ with *Mathematica* we obtain six solutions

$$\begin{aligned}
A &= -0.982582 \\
B &= -0.540281 \\
C &= 0.245133 - 0.452077\sqrt{-1} \\
\overline{C} &= 0.245133 + 0.452077\sqrt{-1} \\
D &= 0.516298 - 0.866582\sqrt{-1} \\
\overline{D} &= 0.516298 + 0.866582\sqrt{-1}.
\end{aligned} \tag{4.12}$$

Among them, the three solutions A, D, $\overline{D}$ are close to -1, $(1/2)(1 \pm \sqrt{-3})$. Thus we see that the critical points A, D, and $\overline{D}$ are on the new-born singular fibers at the splitting of the original singular fiber F^ω (besides the singular fiber ${\phi_\epsilon}^{-1}(0)$ which is a bouquet of two tori).

5. New-born singular fibers and their monodromies

It is well known that any family of genus two Riemann surfaces is hyperelliptic in the sense that it is obtained by taking a double branched covering of a sphere bundle along a branch locus which meets a general sphere-fiber in 6 points. In the concrete situation which we are dealing with, we are very naturally led to such a branched covering. In this section, we will explain this. Balke [2] explains the appearance of the double branched covering in our situation more clearly.

In what follows, we will distinguish several complex planes $\mathbb{C}$ by the symbol used for the variable in the plane, as in $\mathbb{C}_t$, $\mathbb{C}_y$, etc.

Since ϕ_ϵ is equal to $(y(x^2 - y^3 - 1) - \epsilon)(x^2 - y^3 - 1)$, we may think of $\phi_\epsilon(x, y) : \mathbb{C}^2 \to \mathbb{C}_t$ as the pull-back of

$$\psi_\epsilon(X, y) = (y(X - y^3 - 1) - \epsilon)(X - y^3 - 1) : \mathbb{C}^2 \to \mathbb{C}_t \tag{5.1}$$

under the branched-covering map $\Pi : \mathbb{C}^2 \to \mathbb{C}^2$ defined by $\Pi(x, y) = (x^2, y) = (X, y)$. The branch locus of Π is the y-axis: $X = 0$.

In other words, we consider the following commutative diagram:

$$\begin{array}{ccc} \mathbb{C}^2 & \xrightarrow{\phi_\epsilon} & \mathbb{C}_t \\ \Pi \downarrow & & \downarrow \text{identity} \\ \mathbb{C}^2 & \xrightarrow[\psi_\epsilon]{} & \mathbb{C}_t. \end{array}$$

Let us prove that any fiber $f_t^\epsilon (:= \{(X,y)|\psi_\epsilon(X,y) = t\})$ of $\psi_\epsilon : \mathbb{C}^2 \to \mathbb{C}_t$ has no singular points.

In fact, from

$$\frac{\partial \psi_\epsilon}{\partial X} = 2y(X - y^3 - 1) - \epsilon = 0 \tag{5.2}$$

we have $X - y^3 - 1 = \frac{\epsilon}{2y}$. Substituting this into

$$\frac{\partial \psi_\epsilon}{\partial y} = (X - 4y^3 - 1)(X - y^3 - 1) + (y(X - y^3 - 1) - \epsilon)(-3y^2) = 0 \tag{5.3}$$

we have the equation

$$\begin{aligned} 0 &= (\frac{\epsilon}{2y} - 3y^3)\frac{\epsilon}{2y} + (-\frac{\epsilon}{2})(-3y^2) \\ &= (\frac{\epsilon}{2y})^2 \end{aligned} \tag{5.4}$$

which is clearly impossible, because $\epsilon \neq 0$. Thus each fiber f_t^ϵ of the projection ψ_ϵ is a smooth curve in the Xy-plane.

$\Pi : \mathbb{C}^2_{(x,y)} \to \mathbb{C}^2_{(X,y)}$ is a double branched covering branched along the y-axis. Thus, if a fiber f_t^ϵ of ψ_ϵ is tangent to the y-axis, the preimage $F_t^\epsilon = \Pi^{-1}(f_t^\epsilon)$ is a singular fiber of $\phi_\epsilon = \psi_\epsilon \circ \Pi$. A point of f_t^ϵ tangent to the y-axis lifts to a singular point of F_t^ϵ. Since f_t^ϵ is always nonsingular, a singular fiber F_t^ϵ occurs only in this way. Thus, let us look for points of the fibers f_t^ϵ tangent to the y-axis.

For this, we solve the equation

$$\frac{\partial \psi_\epsilon}{\partial y}|_{X=0} = 7y^6 + 8y^3 + 3\epsilon y^2 + 1 = 0. \tag{5.5}$$

Note that this is the same equation as (4.10). For $\epsilon = 0.1$ this equation had six roots, and among them the three roots denoted by $A, D, \overline{D}$ (see (4.12)) were related to the new-born singular fibers of the splitting of F^ω.

For simplicity, we denote the values of ψ_ϵ on these tangent points $(0, A)$, $(0, D)$, and $(0, \overline{D})$ by $t(A)$, $t(D)$, and $t(\overline{D})$, respectively. The fibers $f^\epsilon_{t(A)}$, $f^\epsilon_{t(D)}$, and $f^\epsilon_{t(\overline{D})}$ of ψ_ϵ are non-singular, and are tangent to the y-axis at the above tangency points. As is easily seen, these tangency points are double tangencies, and they split into two points in corresponding nearby fibers, $f^\epsilon_{t(A)+\delta}$, $f^\epsilon_{t(D)+\delta}$, and $f^\epsilon_{t(\overline{D})+\delta}$. In other words, these nearby fibers intersect the y-axis in two points which are close to the corresponding tangency points.

The tangency points $(0, A)$, $(0, D)$, $(0, \overline{D})$ lift (under Π) to the nodes on the singular fibers $F^\epsilon_{t(A)}$, $F^\epsilon_{t(D)}$, $F^\epsilon_{t(\overline{D})}$. Short arcs α_A, α_D, $\alpha_{\overline{D}}$ in the nearby fibers, $f^\epsilon_{t(A)+\delta}$, $f^\epsilon_{t(D)+\delta}$, $f^\epsilon_{t(\overline{D})+\delta}$, respectively, joining the two "split" intersection points will lift to the vanishing cycles in $F^\epsilon_{t(A)+\delta}$, $F^\epsilon_{t(D)+\delta}$, $F^\epsilon_{t(\overline{D})+\delta}$, corresponding to the nodes.

Thus, our first task will be to draw these short arcs, α_A, α_D and $\alpha_{\overline{D}}$, on the fibers $f^\epsilon_{t(A)+\delta}$, $f^\epsilon_{t(D)+\delta}$ and $f^\epsilon_{t(\overline{D})+\delta}$, respectively, and our second task will be to move each of these fibers along a curve on the "base space" $\mathbb{C}_t$ to the common reference fiber $f^\epsilon_{t_0}$ over the base point t_0, and then to see the final positions of the translated arcs, α^0_A, α^0_D and $\alpha^0_{\overline{D}}$, on the reference fiber $f^\epsilon_{t_0}$. From this last piece of information, we will see the global monodromy associated with the new-born singular fibers, $F^\epsilon_{t(A)}$, $F^\epsilon_{t(D)}$, and $F^\epsilon_{t(\overline{D})}$, at the splitting of the original singular fiber F^ω.

Our first task encounters a little unexpected complication.

To explain this, we consider the projection $p_2 : \mathbb{C}^2_{(X,y)} \to \mathbb{C}_y$ of the Xy-plane onto the y-axis, defined by $p_2(X,y) = y$, and represent each fiber f^ϵ_t as a double branched covering over the y-axis. The projection of this double branched covering is the restriction of the projection p_2:

$$p_2|f^\epsilon_t : f^\epsilon_t \to \mathbb{C}_y \tag{5.6}$$

This makes sense, because the equation $\psi_\epsilon = t$ which gives the fiber f^ϵ_t is

$$(y(X-y^3-1)-\epsilon)(X-y^3-1) = t \tag{5.7}$$

(see (5.1)), or, equivalently

$$yX^2 - (2y^4+2y+\epsilon)X + (y^7+2y^4+\epsilon y^3+y+\epsilon-t) = 0 \tag{5.8}$$

This is a quadratic equation for X, provided that $y \neq 0$. Thus, given a generic point $y \neq 0$, there are two simple roots X_1 and X_2 of the equation, yielding two points (X_1, y) and (X_2, y) on f^ϵ_t projected to y under p_2. This means that the fiber f^ϵ_t spreads over the y-axis as a double branched covering.

The case when $y = 0$ seems to cause a problem, for then the equation (5.8) becomes linear. But what we really want to know is not f^ϵ_t itself but a fiber of $\phi_\epsilon : N_\epsilon \to \mathbb{C}_t$. And, near the point $y = 0$, the fiber $\phi_\epsilon = t$ is "absorbed" into the 2-handles attached to N', in such a way that the intersection of the fiber and each attached 2-handle is a 2-disk parallel to the core of the handle. Therefore, near the point $y = 0$, the topology of the fiber f^ϵ_t or that of the fiber of $\phi_\epsilon : N_\epsilon \to \mathbb{C}_t$ does not suffer so big a change. So we may think of f^ϵ_t as a "double branched cover" of the y-axis.

The branch locus of the branched covering $p_2|f^\epsilon_t : f^\epsilon_t \to \mathbb{C}_y$ is found by solving the equation $\Delta = 0$, where Δ is the discriminant of the quadratic equation (5.8).

In fact,

$$\begin{aligned} \Delta &= (2y^4+2y+\epsilon)^2 - 4y(y^7+2y^4+\epsilon y^3+y+\epsilon-t) \\ &= \epsilon^2 + 4ty. \end{aligned} \tag{5.9}$$

Thus, for $t \neq 0$, the branch locus is a single point on the y-axis

$$y = -\frac{\epsilon^2}{4t} \tag{5.10}$$

Since the fiber f^ϵ_t is a double branched covering of the y-axis branched at this point, f^ϵ_t is diffeomorphic to the 2-plane. If it is compactified by adding ∞, the compactified fiber $\hat{f}^\epsilon_t$ is a 2-sphere. The general fiber of $\phi_\epsilon : N_\epsilon \to \mathbb{C}_t$ is obtained by taking the double branched covering (under $\Pi : \mathbb{C}_{(x,y)} \to \mathbb{C}_{(X,y)}$) of the 2-sphere $\hat{f}_t^\epsilon$ branched at certain six points.

Remark. Recall that the double branched covering $\Pi : \mathbb{C}_{(x,y)} \to \mathbb{C}_{(X,y)}$ branches along the y-axis $X = 0$. Putting $X = 0$ in (5.8), we get a 7th degree equation for y, which has seven roots if t is generic. But one root is related to the part absorbed into the 2-handle, and may be neglected. Taking the double covering of the 2-sphere $\hat{f}_t^\epsilon$ branched at the remaining six roots, we get a closed surface of genus 2, that is F_t^ϵ. The details are left to the reader.

The unexpected phenomenon. The unexpected phenomenon alluded to above is the following: Let us for instance consider the fiber $f_{t(D)+\delta}^\epsilon$. This fiber has two intersection points with the y-axis, say $(0, D_1)$ and $(0, D_2)$, near $(0, D)$. ($(0, D)$ is the tangent point of the fiber $f_{t(D)}$ to the y-axis.) The unexpected phenomenon is that these split points $(0, D_1)$ and $(0, D_2)$ are on the *different* sheets of the double branched covering $p_2|f_{t(D)+\delta}^\epsilon : f_{t(D)+\delta}^\epsilon \to \mathbb{C}_y$. (Here we talk about the *sheets* of the double branched covering $p_2|f_{t(D)+\delta}^\epsilon$. In the concrete case below where $\epsilon = 0.1$ and $\delta = 0.0001$, the sheets are defined, for example, by making the "slit" along the lifted curve of the half-line whose terminal point is the branch locus (5.13) and which is parallel to the imaginary axis of the y-plane $\mathbb{C}_y$.) At first the author could not believe this, because the tangency point $(0, D)$ certainly lies on a sheet of the branched covering. If the deviation δ of the fiber $f_{t(D)+\delta}^\epsilon$ from $f_{t(D)}^\epsilon$ is very small, it seemed reasonable to expect that the split points $(0, D_1)$ and $(0, D_2)$, both of which were born from $(0, D)$, should be on the *same* sheet.

However, the points $(0, D_1)$ and $(0, D_2)$ are already on different sheets, even if δ takes a quite small value. To see this, let us make a numerical calculation, setting $\epsilon = 0.1$ and $\delta = 0.0001$.

For $\epsilon = 0.1$, we have already calculated

$$D = 0.516298 - 0.866582\sqrt{-1},$$

and

$$t(D) = \psi_\epsilon(0, D) = -0.00126867 - 0.00212911\sqrt{-1}.$$

To find D_1 and D_2 for $\delta = 0.0001$, we solve the following equation:

$$y^7 + 2y^4 + 0.1y^3 + y + 0.1 = t(D) + 0.0001 \tag{5.11}$$

Among the roots, there are two which are close to D, namely

$$D_1 = 0.513563 - 0.864834\sqrt{-1}, \quad D_2 = 0.519003 - 0.868327\sqrt{-1}.$$

The points $(0, D_1)$ and $(0, D_2)$ are two intersection points of $f_{t(D)+0.0001}^\epsilon$ and the y-axis. Let $c(s)$ be a line segment on the y-axis $\mathbb{C}_y$ joining D_1 and D_2 and parametrised by s $(0 \leq s \leq 1)$. More explicitly, $c(s)$ is given by

$$\begin{aligned} c(s) &= (1 - s)D_1 + sD_2 \\ &= (1 - s)(0.513563 - 0.864834\sqrt{-1}) + s(0.519003 - 0.868327\sqrt{-1}) \end{aligned} \tag{5.12}$$

The branch locus of the double branched covering $p_2|f_{t(D)+0.0001}^\epsilon : f_{t(D)+0.0001}^\epsilon \to \mathbb{C}_y$ is calculated from (5.10) by setting $\epsilon = 0.1$ and $t = t(D) + 0.0001$. The result is

$$0.495292 - 0.902334\sqrt{-1} \tag{5.13}$$

Note that this branch locus is not on the segment $c(s)$, $(0 \leq s \leq 1)$.

Now, in order to check whether or not $(0, D_1)$ and $(0, D_2)$ are on the same sheet of the double covering $p_2|f^{\epsilon}_{t(D)+0.0001}$, let us lift this segment $c(s)$ $(0 \leq s \leq 1)$ to $f^{\epsilon}_{t(D)+0.0001}$ under the double covering map $p_2|f^{\epsilon}_{t(D)+0.0001} : f_{t(D)+0.0001} \to \mathbb{C}_y$.

If $c(s)$ lifts to two arcs, and one of them joins $(0, D_1)$ and $(0, D_2)$, then $(0, D_1)$ and $(0, D_2)$ are on the same sheet. But if $(0, D_1)$ and $(0, D_2)$ belong to different components of the lifted arcs, then the two points $(0, D_1)$ and $(0, D_2)$ are on different sheets. To lift the segment $c(s)$, we successively solve the following quadratic equations for X, by putting $y = c(0.1\, i)$, $i = 0, 1, \ldots, 10$ (cf. equation (5.8)):

$$yX^2 - (2y^4 + 2y + 0.1)X + (y^7 + 2y^4 + 0.1y^3 + y + 0.1) = t(D) + 0.0001 \tag{5.14}$$

Let $X_1(s)$ and $X_2(s)$ denote the two roots of (5.14) with $y = c(s)$. In Figure (12), we plot the lines $\{(\Re(X_1(s)), s)\}_{0 \leq s \leq 1}$ and $\{(\Re(X_2(s)), s)\}_{0 \leq s \leq 1}$. These lines should be conceptually the same as the lifted arcs.

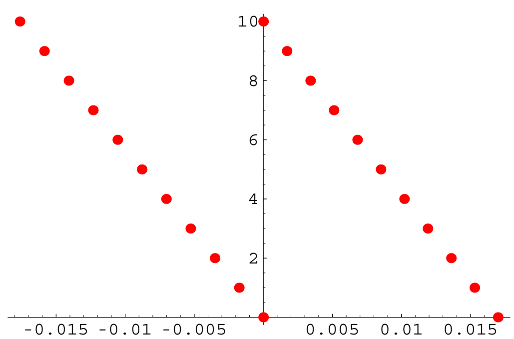


Figure 12. Lines $\{(\Re(X_1(s)), s)\}_{0 \leq s \leq 1}$ and $\{(\Re(X_2(s)), s)\}_{0 \leq s \leq 1}$

Each line of Figure (12) has a terminal point on $\Re(X) = 0$ which corresponds to the point $(0, D_1)$ or $(0, D_2)$. Thus we see that one of the lifted arcs of $c(s)$ has $(0, D_1)$ as a terminal point, and the other has $(0, D_2)$. This implies that $(0, D_1)$ and $(0, D_2)$ are on different sheets of the branched covering $p_2|f^{\epsilon}_{t(D)+0.0001}$.

The author has not yet clearly seen the geometric process by which the two points $(0, D_1)$ and $(0, D_2)$, which were simultaneously born from the same tangency point $(0, D)$, move so fast onto the different sheets of the double branched covering $p_2|f^{\epsilon}_{t(D)+0.0001} : f^{\epsilon}_{t(D)+0.0001} \to \mathbb{C}_y$. But this is surely related to the fact that the tangency point $(0, D)$ of $f^{\epsilon}_{t(D)}$ to the y-axis $\mathbb{C}_y$ and the branch point $(X_D, -\frac{\epsilon^2}{4t(D)})$ of the double branched covering $p_2|f^{\epsilon}_{t(D)} : f^{\epsilon}_{t(D)} \to \mathbb{C}_y$ are very close to each other.

In fact, we have that, for $\epsilon = 0.1$,

$$
\begin{aligned}
D &= 0.516298 - 0.866582\sqrt{-1} \qquad \text{and} \\
-\frac{\epsilon^2}{4t(D)} &= 0.516338 - 0.866530\sqrt{-1}.
\end{aligned}
\tag{5.15}
$$

K. Ahara personally communicated to the author that, if we take a very small value of δ such as $\delta = 0.00000001$, then the corresponding points $(0, D_1)$ and $(0, D_2)$ are on the same sheet of the branched covering $p_2|f^{\epsilon}_{t(D)+\delta}$ as we expected. Then the arguments below must be quite different, but the final conclusion on the splitting of the singular fiber F^{ω} should be the same. The author would like to see how the same conclusion is obtained through different geometric arguments, which will be left to the interested readers.

Arcs joining the split points. Now we work with the original value $\delta = 0.0001$, and want to connect the points $(0, D_1)$ and $(0, D_2)$ by an arc α_D on the fiber $f^{\epsilon}_{t(D)+0.0001}$. The projected image $p_2(\alpha_D)$ on $\mathbb{C}_y$ cannot be a segment, because as we saw above the two points are not on the same sheet of the double branched covering $p_2|f^{\epsilon}_{t(D)+0.0001} : f_{t(D)+0.0001} \to \mathbb{C}_y$. We can take instead the line $D_1 b_D D_2$ of Figure (13) having a bend at the point b_D, where b_D is the branch locus of the double branched covering $f^{\epsilon}_{t(D)+0.0001} \to \mathbb{C}_y$, that is, $b_D = 0.495292 - 0.902334\sqrt{-1}$ (see (5.13)) . Then we can construct a connecting arc α_D on $f^{\epsilon}_{t(D)+0.0001}$ as the union of the lifts of $D_1 b_D$ and $b_D D_2$, which contain $(0, D_1)$ and $(0, D_2)$, respectively. Thus obviously $p_2(\alpha_D) = D_1 b_D D_2$ holds.

Similarly, we can draw an arc α_A on $f^{\epsilon}_{t(A)-\delta}$ (resp. an arc $\alpha_{\overline{D}}$ on $f^{\epsilon}_{t(\overline{D})+\delta}$) connecting $(0, A_1)$ and $(0, A_2)$ (resp. $(0, \overline{D}_1)$ and $(0, \overline{D}_2)$). Here $(0, A_1)$ and $(0, A_2)$ are the two intersection points of $f^{\epsilon}_{t(A)-\delta}$ and the y-axis into which the tangency point $(0, A)$ splits. The explanation is similar for the points $(0, \overline{D}_1)$ and $(0, \overline{D}_2)$. (Note that we take $f^{\epsilon}_{t(A)-\delta}$ instead of $f^{\epsilon}_{t(A)+\delta}$. This is for later convenience.)

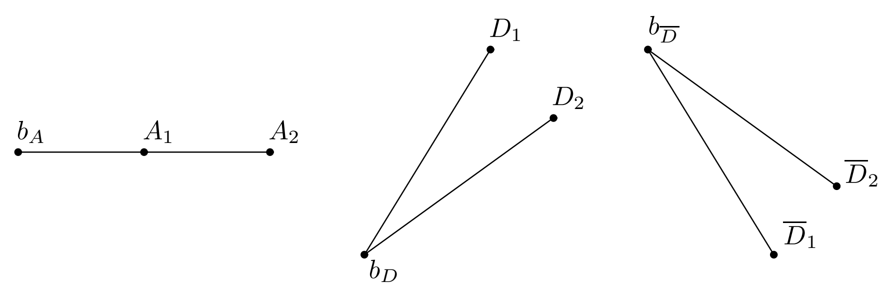


Figure 13. The images of the connecting arcs, $p_2(\alpha_A)$, $p_2(\alpha_D)$ and $p_2(\alpha_{\overline{D}})$.

Observation of local movements of the connecting arcs. We continue the explanation taking $f^{\epsilon}_{t(D)+0.0001}$ as a typical example. We drew an arc α_D on the fiber $f^{\epsilon}_{t(D)+0.0001}$ connecting $(0, D_1)$ and $(0, D_2)$. Next we will move this fiber $f^{\epsilon}_{t(D)+0.0001}$ around $f^{\epsilon}_{t(D)}$, and will see the movement of the arc α_D inside $f^{\epsilon}_{t(D)+0.0001}$. For this purpose, we put $\delta(s) = 0.0001\exp(2\pi\sqrt{-1}s)$, where $0 \le s \le 1$, and solve the equation

$$y^7 + 2y^4 + 0.1y^3 + y + 0.1 = t(D) + \delta(s). \tag{5.16}$$

Let $D_1(s)$ and $D_2(s)$ be the two solutions of (5.16) nearest to $D = 0.516298 - 0.866562\sqrt{-1}$. Changing the parameter s from 0 to 1, we observe the movements of $D_1(s)$ and $D_2(s)$ in $\mathbb{C}_y$.

Also, we observe the movement of the branch locus $b_D(s)$ of the double branched covering $p_2|f^{\epsilon}_{t(D)+\delta(s)} : f^{\epsilon}_{t(D)+\delta(s)} \to \mathbb{C}_y$. In fact, putting $\epsilon = 0.1$ and $t = t(D) + 0.0001\exp(2\pi\sqrt{-1}s)$ in (5.10), we have

$$b_D(s) = -\frac{0.01}{4(t(D) + 0.0001\exp(2\pi\sqrt{-1}s))} \tag{5.17}$$

We have calculated these movements with *Mathematica*. We will describe the results conceptually. The two points $D_1(s)$ and $D_2(s)$ are in opposite positions on a circle whose center is D and whose radius is about 0.003. As the parameter s changes from 0 to 1, the points $D_1(s)$ and $D_2(s)$ move on this circle starting from $D_1(0) = D_1$, and $D_2(0) = D_2$, through 180° counterclockwise until they exchange their positions, $D_1(1) = D_2$ and $D_2(1) = D_1$.

The movement of $b_D(s)$ is as follows: The point $b_D(s)$ starts from $b_D(0) = b_D = 0.495292 - 0.902334\sqrt{-1}$ (see (5.13)) and goes around once counterclockwise on a circle whose center is D and whose radius is about 0.045. Note that the radius of this circle is more than ten times that of the circle on which are D_1 and D_2 lie.

Figure (14) shows these movements conceptually, neglecting the precise proportion of the figures.

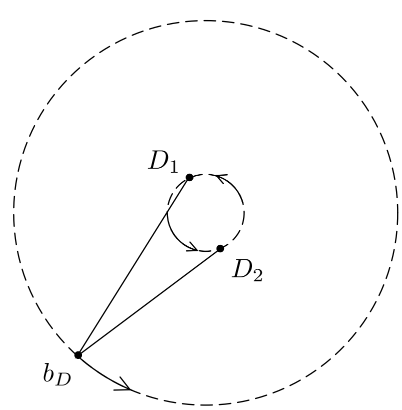


Figure 14. Movements of $D_1(s)$, $D_2(s)$ and $b_D(s)$

From Figure (14), we can also see the movement of the line with a bend $D_1(s)b_D(s)D_2(s)$, too. Tracing this movement from above in $f^\epsilon_{t(D)+\delta(s)}$ via the branched covering map $p_2|f^\epsilon_{t(D)+\delta(s)} : f^\epsilon_{t(D)+\delta(s)} \to \mathbb{C}_y$, we can see the movement of the connecting arc α_D: it rotates 180° counterclockwise inside a disk neighborhood of itself until it exchanges its terminal points $(0, D_1)$ and $(0, D_2)$ (see Figure (15)).

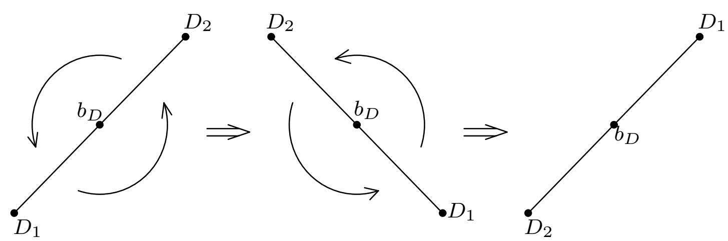


Figure 15. Movement of α_D

Lifting α_D further to the fiber $F^\epsilon_{t(D)+\delta}$ under the branched covering $\Pi|F^\epsilon_{t(D)+\delta} : F^\epsilon_{t(D)+\delta} \to f^\epsilon_{t(D)+\delta}$, we obtain the vanishing cycle corresponding to the node of the singular fiber $F^\epsilon_{t(D)}$. (Recall the commutative diagram at the beginning of §5.) And by lifting the movement of the connecting arc α_D to the fiber $F^\epsilon_{t(D)+\delta}$, we obtain the right-handed Dehn twist along the vanishing cycle.

By similar arguments we can see that, if we move the fiber $F^\epsilon_{t(A)-\delta}$ (or $F^\epsilon_{t(\overline{D})+\delta}$) around the singular fiber $F^\epsilon_{t(A)}$ (or $F^\epsilon_{t(\overline{D})}$), then the corresponding monodromy is the right-hand Dehn twist along the vanishing cycle which is obtained as the lift of α_A (or $\alpha_{\overline{D}}$) under the double branched covering $\Pi|F^\epsilon_{t(A)-\delta} : F^\epsilon_{t(A)-\delta} \to f^\epsilon_{t(A)-\delta}$ (or $\Pi|F^\epsilon_{t(\overline{D})+\delta} : F^\epsilon_{t(\overline{D})+\delta} \to f^\epsilon_{t(\overline{D})+\delta}$).

6. Positions of the vanishing cycles on the reference fiber

We come to the second task mentioned in §5. We choose a base point t_0 on $\mathbb{C}_t$ as

$$t_0 = -0.001 \tag{6.1}$$

and consider the fiber $F^\epsilon_{t_0}$ as the reference fiber. Take $\delta = 0.0001$ as before. We draw the following three paths on $\mathbb{C}_t$ which connect $t(A)-\delta$, $t(D)+\delta$, $t(\overline{D})+\delta$, and t_0, respectively:

$$\begin{aligned} l_A(s) &= (1-2s)(t(A)-\delta) + 0.002s && (0 \le s \le \tfrac{1}{2}) \\ &= 0.001\exp(\pi\sqrt{-1}(2s-1)) && (\tfrac{1}{2} \le s \le 1) \end{aligned} \tag{6.2}$$

$$l_D(s) = (1-s)(t(D)+\delta) - 0.001s \qquad (0 \le s \le 1) \tag{6.3}$$

$$l_{\overline{D}}(s) = (1-s)(t(\overline{D})+\delta) - 0.001s \qquad (0 \le s \le 1) \tag{6.4}$$

These curves are conceptually shown in Figure (16). Using this figure, it will become self-evident why we have chosen $t(A)-\delta$ instead of $t(A)+\delta$ as a generic locus near the singular locus $t(A)$.

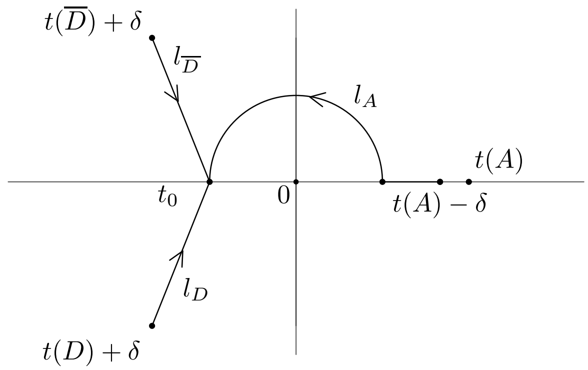


Figure 16. Paths l_A, l_D, and $l_{\overline{D}}$

Putting $X=0$, $\epsilon=0.1$ in (5.8) and moving the parameter t along the path l_D, from $t(D)+\delta$ to t_0, we successively solve the equation (5.8) for y with *Mathematica*, and observe the movements of D_1, D_2 in $\mathbb{C}_y$. Also we observe the movement of b_D using the formula (5.10). Then we see how the line $D_1b_DD_2$ with a bend moves and to what position it finally comes. Figure (17) (conceptually) shows the final position, which is nothing but the image $p_2(\alpha_D^0)$ of the arc α_D^0 under the projection $p_2|f_{t_0}^\epsilon : f_{t_0}^\epsilon \to \mathbb{C}_y$. Note that here α_D^0 denotes the final position in $f_{t_0}^\epsilon$ of the arc α_D. The branch locus b_D has come to the position $b_0 = 2.5$. This is the branch locus of the double branched covering $p_2|f_{t_0}^\epsilon : f_{t_0}^\epsilon \to \mathbb{C}_y$.

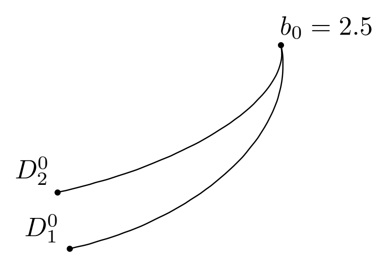


Figure 17. The final position of $D_1b_DD_2$ in $\mathbb{C}_y$

Similarly, using *Mathematica*, we can calculate how the arcs $\alpha_{\overline{D}}$ (in $f_{t(\overline{D})+\delta}$) and α_A (in $f_{t(A)-\delta}$) move as we change the parameter t along the paths $l_{\overline{D}}$ and l_A from $t(\overline{D})+\delta$ to t_0 and from $t(A)-\delta$ to t_0, respectively, and find their final positions $\alpha_{\overline{D}}^0$ and α_A^0 in $f_{t_0}^\epsilon$.

Figure (18) shows the projected images in $\mathbb{C}_y$ of the arcs α_D^0, $\alpha_{\overline{D}}^0$, and α_A^0 under the projection $p_2 : f_{t_0}^\epsilon \to \mathbb{C}_y$. These are three lines with a bend which meet at the point b_0.

Incidentally, the dotted half line in this figure (denoted by α_∞) corresponds to the vanishing cycle of the central fiber F_0^ϵ of type II. This is explained as follows. The point b_0 is the branch locus of the double branched covering $p_2 : f_{t_0}^\epsilon \to \mathbb{C}_y$. If we let t approach from t_0 to 0, from the negative side, the branch point $b(= -\frac{\epsilon^2}{4t})$ of the branched covering $p_2 : f_t^\epsilon \to \mathbb{C}_y$ moves along the dotted line α_∞ in the positive direction. And if we imagine the extreme case where $t = 0$, the branch locus b would disappear from our eyesight. Then the double covering $p_2 : f_0^\epsilon \to \mathbb{C}_y$ is no longer a branched covering but just a trivial covering consisting of two sheets of planes. The compactification $\hat{f}_0^\epsilon$ of the two planes is a bouquet of two 2-spheres. We obtain the fiber F_0^ϵ by further taking a double branched cover of these 2-spheres (each branched along 3 points together with the point ∞). The result is a bouquet of two tori, which is nothing but F_0^ϵ. The arc α_∞ connecting t_0 and ∞ (in the compactified $\hat{\mathbb{C}}_y$) lifts to a simple closed curve γ in $\hat{f}_{t_0}^\epsilon$, and is further doubly covered by a simple closed curve $\tilde{\gamma}$ in $F_{t_0}^\epsilon$. When t moves from t_0 to 0, the curve $\tilde{\gamma}$ shrinks to the node of F_0^ϵ. Thus $\tilde{\gamma}$ is the vanishing cycle for the singular point of F_0^ϵ. This explains the relationship of α_∞ and the vanishing cycle for F_0^ϵ.

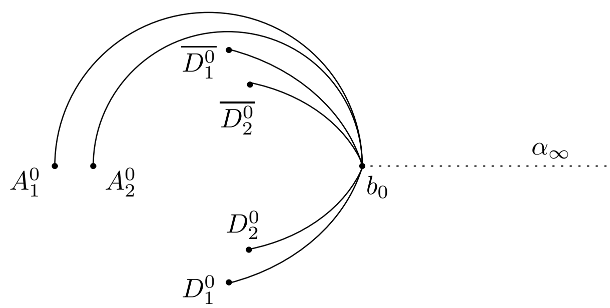


Figure 18. The three lines with a bend, $p_2(\alpha_A^0)$, $p_2(\alpha_D^0)$, $p_2(\alpha_{\overline{D}}^0)$ and α_∞

Recall that the compactification $\hat{f}_{t_0}^\epsilon$ of $f_{t_0}^\epsilon$ is a 2-sphere, and that the fiber $F_{t_0}^\epsilon$ of genus two is obtained by taking a double branched cover of $\hat{f}_{t_0}^\epsilon$ branched at certain six points. In fact, the six branch points are solutions to the equation (5.8) for y with $X = 0$, $\epsilon = 0.1$ and $t = t_0 = -0.001$. Using *Mathematica*, we calculate the following seven solutions:

$$\begin{aligned}
A_1^0 &= -1.00304 \\
A_2^0 &= -0.960718 \\
D_1^0 &= 0.501424 - 0.869172\sqrt{-1} \\
\overline{D}_1^0 &= 0.501424 + 0.869172\sqrt{-1} \\
D_2^0 &= 0.531010 - 0.864630\sqrt{-1} \\
\overline{D}_2^0 &= 0.531010 + 0.864630\sqrt{-1} \\
\text{extra} &= -0.101106 \risingdotseq -\epsilon
\end{aligned}$$

The solution denoted by "extra" is the "negligible" solution (see the *Remark* in §5). The remaining six solutions A_1^0, A_2^0, D_1^0, $\overline{D}_1^0$, D_2^0, $\overline{D}_2^0$ are the six branch points. They also coincide with (the p_2-image of) the terminal points of the three arcs α_A^0, α_D^0, $\alpha_{\overline{D}}^0$. It will be evident from the notation which points are the terminal points of which arc. In what follows, we will use the same notation for the terminal points (in $f_{t_0}^\epsilon$) and their p_2-images (in $\mathbb{C}_y$).

The 2-sphere $\hat{f}_{t_0}^\epsilon$ is divided into two hemispheres by the simple closed curve γ which is the lift of the dotted line α_∞ (of Figure (18)).

LEMMA (6.5). *The points A_1^0, D_1^0, $\overline{D}_1^0$ are on one hemisphere bounded by γ, and A_2^0, D_2^0, $\overline{D}_2^0$ are on the other hemisphere.*

Proof. If we move t from t_0 to 0 then, as we saw above, the branch point b moves from b_0 to ∞ on the line α_∞ and finally disappears. We can see, using *Mathematica*, that the points A_1^0, D_1^0, $\overline{D}_1^0$ meanwhile converge to the solutions

$$-1, \quad \frac{1}{2} \pm \frac{\sqrt{-3}}{2}$$

of the equation $y^3+1=0$, which is one factor of the equation (5.7) (with $X=0$ and $t=0$), and the points A_2^0, D_2^0, $\overline{D}_2^0$, and the "extra" one, converge to the solutions of the other factor $y^4+y+\epsilon=0$ of (5.7). During the movements, these points do not cross the dotted line α_∞. The two factors of (5.7) in the extreme case $t=0$ correspond to the two sheets of the trivial double covering $p_2 : f_0^\epsilon \to \mathbb{C}_y$, which is compactified to a bouquet of two 2-spheres $\hat{f}_0^\epsilon$. The three points $\{A_1^0, D_1^0, \overline{D}_1^0\}$ and the four points $\{A_2^0, D_2^0, \overline{D}_2^0, \text{"extra"}\}$ belong to the different sheets of f_0^ϵ, and thus lie on different components of the bouquet of two 2-spheres $\hat{f}_0^\epsilon$. On the other hand, as t moves from t_0 to 0, the simple closed curve γ is pinched to a point and $\hat{f}_{t_0}^\epsilon$ becomes the bouquet of two 2-spheres $\hat{f}_0^\epsilon$. Thus, in $\hat{f}_{t_0}^\epsilon$ the three points and the four points lie on different sides of γ. □

Regard Figure (18) as a picture drawn on $\hat{\mathbb{C}}_y$. Taking a double branched cover of $\hat{\mathbb{C}}_y$ branched at the two points $\{b_0, \infty\}$, we obtain $\hat{f}_{t_0}^\epsilon$. The picture on $\hat{\mathbb{C}}_y$ is doubly covered by the picture on $\hat{f}_{t_0}^\epsilon$ of the three arcs α_A, α_D, $\alpha_{\overline{D}}$ and the simple closed curve γ. By Lemma (6.5), the picture must be as shown in Figure (19).

The reference fiber $F_{t_0}^\epsilon$ is obtained by taking a double branched cover of $\hat{f}_{t_0}^\epsilon$ branched at the six points $\{A_1^0, A_2^0, D_1^0, D_2^0, \overline{D}_1^0, \overline{D}_2^0\}$. To see this, we continuously change the picture of the three arcs and γ as shown in Figure (20). Note that in the changed picture, we preserve the same notations A_1^0, A_2^0, etc., as before.

Cut open the sphere $\hat{f}_{t_0}^\epsilon$ along the three segments $A_1^0A_2^0$, $\overline{D}_1^0D_1^0$, $D_2^0\overline{D}_2^0$. The result is a 2-disk with two holes. See Figure (21).

We deform the disk with two holes into a surface as shown in Figure (22).

Take two copies of the deformed surface of Figure (22), and glue them together along their boundaries. We then obtain a closed surface of genus two, which may be considered as the reference fiber $F_{t_0}^\epsilon$. See Figure (23).

The deformed surface contains the arcs α_A, α_D, $\alpha_{\overline{D}}$ and an arc which is γ cut open to a segment (we denote this arc by γ again). When gluing the two copies of the deformed surface, we at the same time glue the copies of these arcs to obtain

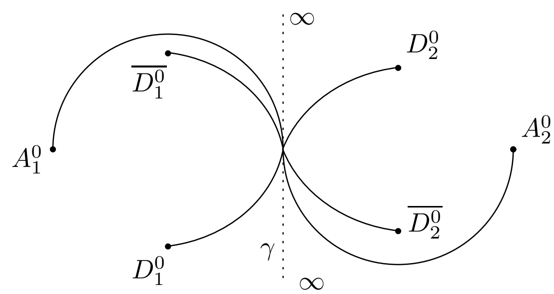


Figure 19. The three arcs and the simple closed curve γ on $\hat{f}^{\epsilon}_{t_0}$

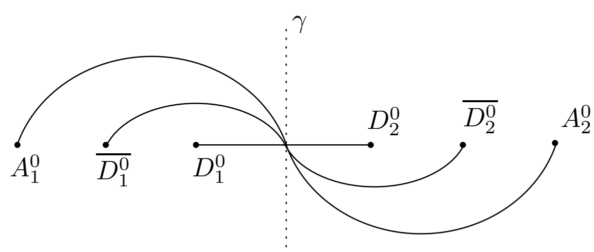


Figure 20. Continuously changed picture

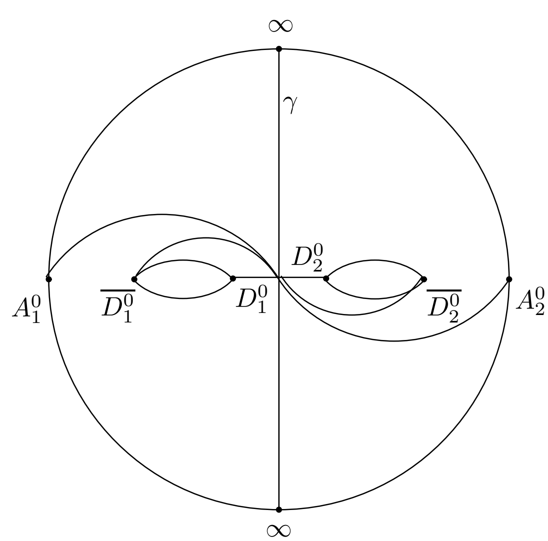


Figure 21. The disk with two holes

on $F^{\epsilon}_{t_0}$ four simple closed curves β_1, β_2, β_3, and β_4, which doubly cover α_D, γ, α_A and $\alpha_{\overline{D}}$, respectively. (The curve β_2 is nothing but $\tilde{\gamma}$.) See Figure (23). By

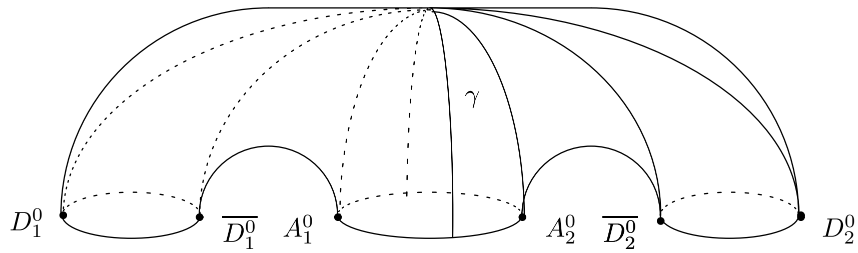


Figure 22. The deformed surface

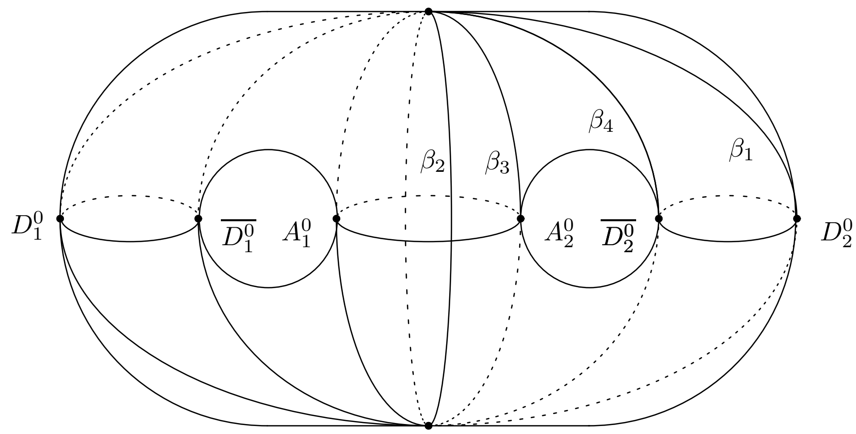


Figure 23. The closed surface $F_{t_0}^{\epsilon}$

construction, these curves are the vanishing cycles for the singular fibers $F_{t(D)}^{\epsilon}$, F_0^{ϵ}, $F_{t(A)}^{\epsilon}$, $F_{t(\overline{D})}^{\epsilon}$, respectively.

On the parameter plane $\mathbb{C}_t$, the paths l_D, $\overline{t_0 0}$, l_A and $l_{\overline{D}}$ arrive at the points t_0 in this cyclic order (counterclockwise). See Figure (16). Since the vanishing cycle near each singular fiber is carried to the reference fiber $F_{t_0}^{\epsilon}$ along these paths, composing the right-handed Dehn twists about the simple closed curves

$$\beta_1, \quad \beta_2, \quad \beta_3, \quad \beta_4$$

in this cyclic order, we obtain the monodromy ω around the singular fiber F^{ω}. This completes the proof of Theorem (2.1).

7. On the proof of Theorem (2.6)

Our original proof of Theorem (2.6) followed a line similar to the one used for Theorem (2.1), and it also used computer calculations. Ito [8] extended our proof to the general case of arbitrary genus. His proof does not depend on the use of computers. As an application, he constructed a Lefchetz fibration $\mathbb{C}P^2\#(4g+5)\overline{\mathbb{C}P}^2 \to S^2$ of genus g, for each $g \geq 1$, which extends our previous construction in the case of genus two [10]. His Lefschetz fibration has the total monodromy

$(\zeta_1\zeta_2\ldots\zeta_{2g+1}\zeta_{2g+1}\ldots\zeta_2\zeta_1)^2 = 1$. Since Ito's paper [8] is now available, the author would like to refer the reader to that paper for the detailed argument.

Received August 07, 2003

Final version received May 31, 2004

GRADUATE SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY OF TOKYO
KOMABA, MEGURO-KU
TOKYO 153-8914
JAPAN
yukiomat@ms.u-tokyo.ac.jp

REFERENCES

[1] T. ARAKAWA AND T. ASHIKAGA, *Local splitting families of hyperelliptic pencils*, I, Tohoku Math. J. **53** (2001), 369–394.
[2] L. BALKE, Notes (5 pages, unpublished), December, 2000.
[3] J. BIRMAN, *Braids, links, and mapping class groups*, Ann. Math. Studies No.82, Princeton university Press, 1974.
[4] C. CADAVID, *A remarkable set of words in the mapping class group*, Dissertation, U. of Texas, Austin, 1998.
[5] H. ENDO, *Meyer's signature cocycle and hyperelliptic fibrations*, Math. Ann. **316** (2000), 237–257.
[6] R. E. GOMPF AND A. I. STIPSICZ, *4-manifolds and Kirby calculus*, Grad. Stud. Math. **20**, Amer. Math. Soc., Providence, RI, 1999.
[7] E. HORIKAWA, *Local deformation of pencils of curves of genus two*, Proc. Japan Acad. Ser. A, Math.Sci., **64** (1988), 241–244.
[8] T. ITO, *Splitting of singular fibers in certain holomorphic fibrations*, J. Math. Sci. Univ. Tokyo, **9** (2002), 425–480.
[9] M. KORKMAZ, *Noncomplex smooth 4-manifolds with Lefschetz fibrations*, Internat. Math. Res. Notices (2001), no. 3, 115–128.
[10] Y. MATSUMOTO, *Lefschetz fibrations of genus two - a topological approach*, Proc. of the 37th Taniguchi Symposium (Kojima, et.al., eds.), World Scientific (1996), 123–148.
[11] Y. MATSUMOTO AND J.M. MONTESINOS-AMILIBIA, *Pseudoperiodic homeomorphisms and degeneration of Riemann surfaces*, Bull. Amer. Math. Soc. **30** (1994), 70–75.
[12] B. OZBAGCI AND A.I. STIPSICZ, *Noncomplex smooth 4-manifolds with genus-2 Lefschetz fibrations*, Proc. Amer. Math. Soc., **128** (2000), 3125–3128.
[13] U. PERSSON, *Genus two fibrations revisited*, Lecture Notes in Math. **1507**, Springer-Verlag, Berlin, (1992), 133–144.
[14] M. REID, *Problems on pencils of small genus*, (preprint)
[15] S. TAKAMURA, *Towards the Classification of Atoms of Degenerations*, I, Splitting Criteria via Configurations of Singular Fibers, J. Math. Soc. Japan **56** (2004), 115–145.
[16] S. TAKAMURA, *Towards the Classification of Atoms of Degenerations*, II, Linearization of Degenerations of Complex Curves, (preprint) RIMS (2001).
[17] S. TAKAMURA, *Towards the Classification of Atoms of Degenerations*, III, Barking Deformations of Degenerations, (preprint) RIMS (2002).
[18] G. XIAO, *Surfaces fibrées en courbes de genre deux*, Lecture Notes in Math. **1137**, Springer-Verlag, Berlin, 1985.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

ON HYPERBOLIC POLYHEDRA ARISING AS CONVEX CORES OF QUASI-FUCHSIAN PUNCTURED TORUS GROUPS

A.D. MEDNYKH, J.R. PARKER, AND A.YU. VESNIN

Dedicated to Francisco Javier González Acuña on the occasion of his 60th birthday

Abstract. We consider two families of hyperbolic polyhedra. With one set of face pairings, these polyhedra give the convex core of certain quasi-Fuchsian punctured torus groups. With additional face pairings, they are related to hyperbolic cone manifolds with singularities over certain links. For both families we derive formulae relating the dihedral angles, side lengths and the volume of the polyhedron.

1. Introduction

A *Kleinian* group G is a discrete subgroup of $\mathrm{PSL}(2,\mathbb{C})$, the isometry group of hyperbolic space $\mathbb{H}^3$. Such a group also acts by conformal automorphisms on the Riemann sphere $\widehat{\mathbb{C}} = \partial\mathbb{H}^3$. The action on $\widehat{\mathbb{C}}$ decomposes into the *regular set* $\Omega(G)$ on which the action is properly discontinuous, and the *limit set* $\Lambda(G)$ on which the action is minimal, that is every orbit is dense. The limit set $\Lambda(G)$ is the set of accumulation points of the fixed points of G. A Kleinian group G is *Fuchsian* if $\Lambda(G)$ is a round circle.

Let S be an oriented surface of negative Euler characteristic, homeomorphic to a closed surface with at most a finite number of punctures. A finitely generated Kleinian group G is *quasi-Fuchsian* if $\mathbb{H}^3/G$ is homeomorphic to the product of such a surface with the open interval $(0,1)$, and if $\Omega(G)$ has exactly two simply connected G-invariant components Ω^+ and Ω^-. Equivalently, $G = \pi_1(S)$ and $\Lambda(G)$ is topological circle. In this situation, the quotients Ω^+/G and Ω^-/G are Riemann surfaces, both homeomorphic to S.

Let $M = \mathbb{H}^3/G$ be the 3-manifold uniformized by the Kleinian group G. The *convex core* $\mathcal{C}/G$ of M is the smallest closed convex set containing all closed geodesics of M. This means that $\mathcal{C}$ can be defined in the universal cover $\mathbb{H}^3$ as the hyperbolic *convex hull* of the limit set $\Lambda(G)$, also called the *Nielsen region* of G. If G is quasi-Fuchsian, then $\partial\mathcal{C}$ has exactly two components $\partial\mathcal{C}^+$ and $\partial\mathcal{C}^-$ which "face" the components Ω^+ and Ω^- of Ω. The quotients $\partial\mathcal{C}^+/G$ and

---

2000 *Mathematics Subject Classification*: 30F40, 51M10, 57M50.

Keywords and phrases: Hyperbolic polyhedron, quasi-Fuchsian group, hyperbolic cone-manifold, bending, convex core.

This research was partially carried out during a visit of the first author to Durham. This visit was supported by a grant from the London Mathematical Society. The third author was partially supported by the Russian Foundation for Basic Research (grant number 02-01-01118).

$\partial\mathcal{C}^-/G$ are homeomorphic to Ω^+/G and Ω^-/G, respectively, and, hence, to S. In the case where G is Fuchsian, $\mathcal{C}$ is contained in a single flat plane.

The convex hull boundary $\partial\mathcal{C}$ is made up of convex pieces of flat hyperbolic planes which meet along a disjoint set of complete geodesics called *pleating* or *bending lines* (see [2] and [3] for more discussions).

It is well known that a Kleinian group is geometrically finite if and only if its convex core has finite volume. Moreover, it is also well known that finitely generated quasi-Fuchsian groups are geometrically finite.

In the present paper we are interested in the case where S is homeomorphic to a punctured torus. So, $G = \langle X, Y \mid [X,Y] \text{ is parabolic}\rangle$, where X and Y are isometries of $\mathbb{H}^3$. We will be interested in cases where certain elements of $\langle X, Y\rangle$ are purely hyperbolic. An isometry X of $\mathbb{H}^3$ is called purely hyperbolic if its associated matrix $\mathbf{X}$ in $\mathrm{SL}(2,\mathbb{C})$ has trace $\mathrm{tr}(\mathbf{X})$ that is real and either greater than 2 or less than -2. Geometrically such an isometry is a hyperbolic translation along a geodesic with no twisting.

We find hyperbolic polyhedra which are fundamental domains for the convex cores of certain quasi-Fuchsian punctured torus groups. In particular, we consider the two cases of punctured torus groups $\langle X, Y\rangle$ for which:

(i) the isometries X and Y are purely hyperbolic;
(ii) the isometries XY and XY^{-1} are purely hyperbolic.

These quasi-Fuchsian punctured torus groups are such that the pleating locus on each component of the convex hull boundary is a simple closed geodesic and either these geodesics are a pair of neighbours or else they are next-but-one neighbours. For each of these two types of group we find a polyhedron and face pairings so that identifying the faces of the polyhedron gives the convex core of the quasi-Fuchsian manifold (see Sections 2.1 and 3.1). These polyhedra will have all their dihedral angles equal to $\pi/2$ except for the dihedral angles along the pleating curves. We demonstrate two approaches to find relations between the lengths of these curves and the dihedral angles. In Sections 2.2 and 3.2 we use the bending formulae due to Parker and Series [12]. In Sections 2.3 and 3.3 we derive these and other formulae (which will be necessary to obtain expressions for volumes) from the Gram matrix of the polyhedra. We then go to use Schläfli's formula (see [1, 9, 14]) to obtain volumes of these polyhedra in Sections 2.4 and 3.4. In particular, we give expressions for volumes in terms of the Lobachevsky function $\Lambda(x)$, which is traditionally used to express volumes of hyperbolic 3-polyhedra and 3-manifolds. In Sections 2.5 and 3.5 we discuss links and cone-manifolds naturally associated with our polyhedra. For the first case the singular set of the cone manifold is the Borromean rings, a well known three component link, and for the second case it is a six-component link.

2. The case where X and Y are purely hyperbolic

(2.1) Constructing the polyhedron. Let matrices $\mathbf{X}, \mathbf{Y} \in \mathrm{SL}(2,\mathbb{C})$, $\mathrm{tr}[\mathbf{X}, \mathbf{Y}] = -2$, represent isometries X and Y of $\mathbb{H}^3$ which generate a punctured torus group. For the rest of this section we suppose that $\mathrm{tr}(\mathbf{X})$ and $\mathrm{tr}(\mathbf{Y})$ are both real and greater than 2. (We remark that one may choose the signs of the traces of $\mathbf{X}$ and $\mathbf{Y}$ when lifting from $\mathrm{PSL}(2,\mathbb{C})$ to $\mathrm{SL}(2,\mathbb{C})$.) We define the multiplier of a matrix $\mathbf{M}$, $\lambda(\mathbf{M})$ by $\mathrm{tr}(\mathbf{M}) = 2\cosh\lambda(\mathbf{M})$ (see [12] for details).

We denote $x = \cosh\lambda(\mathbf{X}) = \frac{1}{2}\mathrm{tr}(\mathbf{X})$ and $y = \cosh\lambda(\mathbf{Y}) = \frac{1}{2}\mathrm{tr}(\mathbf{Y})$. Thus, x and y are real and greater than 1 in our case. In Theorem 6.3 of [12] it is shown that either $\langle X, Y\rangle$ is Fuchsian or else the axes of X and Y are the pleating loci of the convex hull boundary of $\langle X, Y\rangle$. Specifically this theorem states that

Proposition (2.1.1) ([12], Theorem 6.3). *Suppose that $\langle X, Y\rangle$ is a punctured torus group with $x = \cosh\lambda(\mathbf{X}) > 1$ and $y = \cosh\lambda(\mathbf{Y}) > 1$.*

(i) *If $x^2 + y^2 \le x^2y^2$ then $\langle X, Y\rangle$ is Fuchsian.*

(ii) *If $x^2 + y^2 > x^2y^2$ then $\langle X, Y\rangle$ is quasi-Fuchsian and the axes of X and Y are the pleating loci.*

From now on we suppose that $x^2 + y^2 > x^2y^2$, that is the non-Fuchsian case.

We want to construct a fundamental polyhedron for the convex hull of the limit set (Nielsen region) of $\langle X, Y\rangle$. This will be $\mathcal{P} = \mathcal{P}(\alpha, \beta)$. Since $\mathbf{X}$, $\mathbf{Y}\mathbf{X}^{-1}\mathbf{Y}^{-1}$ and their product $\mathbf{Y}\mathbf{X}^{-1}\mathbf{Y}^{-1}\mathbf{X}$ all have real trace, the corresponding isometries X and $YX^{-1}Y^{-1}$ generate a Fuchsian group. Similarly, since $\mathbf{Y}$, $\mathbf{X}^{-1}\mathbf{Y}^{-1}\mathbf{X}$ and their product $\mathbf{Y}\mathbf{X}^{-1}\mathbf{Y}^{-1}\mathbf{X}$ all have real trace, the corresponding isometries Y and $X^{-1}Y^{-1}X$ also generate a Fuchsian group.

- Let Π_+ denote the plane preserved by the group $\langle X,\, YX^{-1}Y^{-1}\rangle$;
- Let Π_- denote the plane preserved by the group $\langle Y,\, X^{-1}Y^{-1}X\rangle$.

It will follow from our construction that Π_+ and Π_- are support planes for the convex hull boundary of $\langle X, Y\rangle$. In [12] this was shown using a different method.

We define geodesics γ_X, γ_Y and γ_0 by:

- γ_X is the axis of X and γ_Y is the axis of Y;
- γ_0 is the common perpendicular of γ_X and γ_Y.

A halfturn is an elliptic isometry of order 2 fixing a geodesic pointwise. We define halfturns I_0, I_1 and I_2 as follows.

- Let I_0 to be the halfturn fixing γ_0.
- Define I_1 by $I_1 = I_0X$. Then I_1 is a halfturn fixing a geodesic γ_1.
- Define I_2 by $I_2 = YI_0$. Then I_2 is a halfturn fixing a geodesic γ_2.

Then we have

$$\begin{array}{lll} I_0XI_0 = X^{-1}, & I_1XI_1 = X^{-1}, & I_2XI_2 = YX^{-1}Y^{-1}, \\ I_0YI_0 = Y^{-1}, & I_1YI_1 = X^{-1}Y^{-1}X, & I_2YI_2 = Y^{-1}. \end{array}$$

Lemma (2.1.2). *The halfturn I_2 preserves the plane Π_+ and the halfturn I_1 preserves the plane Π_-.*

Proof. Since $I_2XI_2 = YX^{-1}Y^{-1}$ it is clear that I_2 swaps the axes of X and YXY^{-1}. These geodesics span the plane Π_+ and so I_2 preserves this plane. Similarly, since I_1 swaps the axes of Y and $X^{-1}YX$, it preserves Π_-. □

We now define reflections R_0 and R_0' in planes Π_0 and Π_0' as follows:

- Let R_0 be reflection in the plane Π_0 containing γ_0 and γ_X.
- Let R_0' be reflection in the plane Π_0' containing γ_0 and γ_Y.

Then we have

$$R_0XR_0 = X, \quad R_0'YR_0' = Y.$$

Lemma (2.1.3). *The plane Π_0 is orthogonal to γ_Y and the plane Π_0' is orthogonal to γ_X.*

Proof. In order to show this, we calculate the complex distance $\delta(X,Y)$ between γ_X and γ_Y and show that $\cosh\delta(X,Y)$ is purely imaginary.

We find $\cosh\delta(X,Y)$ by constructing a right angled hexagon and using Fenchel's generalised cosine rule (see [4]). Doing this we obtain the following formula (which is (1.3) of [12]).

$$\cosh\delta(X,Y) = \frac{\cosh\lambda(\mathbf{XY}) - \cosh\lambda(\mathbf{X})\cosh\lambda(\mathbf{Y})}{\sinh\lambda(\mathbf{X})\sinh\lambda(\mathbf{Y})} \tag{2.1.4}$$

From the well known expression for the trace of the commutator

$$\mathrm{tr}[\mathbf{X},\mathbf{Y}] = \mathrm{tr}^2(\mathbf{X}) + \mathrm{tr}^2(\mathbf{Y}) + \mathrm{tr}^2(\mathbf{XY}) - \mathrm{tr}(\mathbf{X})\,\mathrm{tr}(\mathbf{Y})\,\mathrm{tr}(\mathbf{XY}) - 2, \tag{2.1.5}$$

we see that the traces of $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{XY}$ satisfy the Markov equation [12]:

$$\mathrm{tr}^2(\mathbf{X}) + \mathrm{tr}^2(\mathbf{Y}) + \mathrm{tr}^2(\mathbf{XY}) = \mathrm{tr}(\mathbf{X})\,\mathrm{tr}(\mathbf{Y})\,\mathrm{tr}(\mathbf{XY}). \tag{2.1.6}$$

Therefore

$$x^2 + y^2 + \cosh^2\lambda(\mathbf{XY}) = 2xy\cosh\lambda(\mathbf{XY}). \tag{2.1.7}$$

Hence

$$\cosh^2\delta(X,Y) = \frac{(\cosh\lambda(\mathbf{XY}) - xy)^2}{(x^2-1)(y^2-1)} = \frac{x^2y^2 - x^2 - y^2}{(x^2-1)(y^2-1)} < 0,$$

where we have used $x > 1$, $y > 1$ and $x^2 + y^2 > x^2y^2$. Thus the imaginary part of the complex distance between the axes of X and Y is $\pi/2$ (it also can be seen by the arguments of [7]). □

A consequence of this lemma is

$$R_0R_0' = I_0, \quad R_0YR_0 = Y^{-1}, \quad R_0'XR_0' = X^{-1}.$$

Moreover, define

- $R_1 = R_0'X$, a reflection fixing a plane Π_1 and
- $R_2 = YR_0$, a reflection fixing a plane Π_2.

Then γ_X is the common orthogonal of Π_1 and Π_0'. The distance between these planes is $\lambda(\mathbf{X})$, the multiplier of $\mathbf{X}$. Also Π_1 contains γ_0 and γ_1. Similarly, γ_Y is the common orthogonal of Π_2 and Π_0; the distance between them is $\lambda(\mathbf{Y})$; and Π_2 contains γ_0 and γ_2.

Lemma (2.1.8). *The planes Π_1 and Π_2 are each orthogonal to both of the planes Π_+ and Π_-.*

Proof. We have

$$R_1(X)R_1 = (R_0'X)X(X^{-1}R_0') = R_0'XR_0' = X^{-1}.$$

Also

$$\begin{aligned} R_1(YX^{-1}Y^{-1}X)R_1 &= (R_0'X)YX^{-1}Y^{-1}X(X^{-1}R_0') \\ &= (R_0'XR_0')(R_0'YR_0')(R_0'X^{-1}R_0')(R_0'Y^{-1}R_0') \\ &= XY^{-1}X^{-1}Y \\ &= (YX^{-1}Y^{-1}X)^{-1}. \end{aligned}$$

Therefore R_1 preserves the plane Π_+ preserved by X and $YX^{-1}Y^{-1}$.

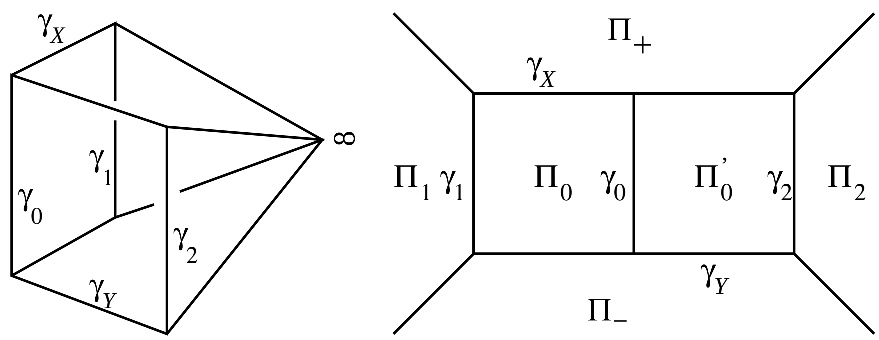


Figure 2.1. The polyhedron $\mathcal{P}'(\alpha, \beta)$.

Moreover

$$R_1(Y)R_1 = (R_0'X)Y(X^{-1}R_0') = X^{-1}YX = (X^{-1}Y^{-1}X)^{-1}.$$

Therefore R_1 swaps the axes of Y and $X^{-1}YX$, which both lie in Π_-. Therefore R_1 also preserves the plane Π_- preserved by Y and $X^{-1}Y^{-1}X$. Since Π_1 is distinct from Π_+ and Π_- we see that it must be orthogonal to them both.

A similar argument shows Π_2 is orthogonal to both Π_+ and Π_-. □

Summarising we have:

- The planes Π_0 and Π_0' meet at right angles along γ_0;
- the planes Π_0 and Π_1 meet at right angles along γ_1;
- the planes Π_0' and Π_2 meet at right angles along γ_2;
- the planes Π_+ and Π_0 meet along γ_X at dihedral angle say $\alpha/2$;
- the planes Π_- and Π_0' meet along γ_Y at dihedral angle say $\beta/2$;
- the planes Π_+ and Π_0' meet at right angles;
- the planes Π_- and Π_0 meet at right angles;
- the planes Π_+ and Π_1 meet at right angles;
- the planes Π_- and Π_1 meet at right angles;
- the planes Π_+ and Π_2 meet at right angles;
- the planes Π_- and Π_2 meet at right angles.

Therefore, the intersection of halfspaces bounded by Π_+, Π_-, Π_0, Π_0', Π_1 and Π_2 is a polyhedron, which we denote by $\mathcal{P}' = \mathcal{P}'(\alpha, \beta)$, with six faces and eleven edges, having one vertex at infinity (ideal vertex) (see Figure 2.1). We remark that the polyhedron $\mathcal{P}'(\alpha, \beta)$ presented in Figure 2.1 can be regarded as a degenerate Lambert cube $\mathcal{L}(\alpha/2, \beta/2, 0)$.

In this polyhedron, and all subsequent polyhedra we shall consider, the 3-valent vertices are interior points of $\mathbb{H}^3$ and the 4-valent vertices are ideal vertices on $\partial\mathbb{H}^3$. We denote these vertices by the symbol ∞ in the figures.

We are now in a position to construct the polyhedron $\mathcal{P} = \mathcal{P}(\alpha, \beta)$. The polyhedron $\mathcal{P}$ will be the common intersection of halfspaces bounded by Π_+, Π_-, Π_1, Π_2 and their images under I_0. This consists of four copies of $\mathcal{P}'(\alpha, \beta)$ glued together along the planes Π_0 and Π_0'. For $i = 1, 2$ let F_i, F_{i+2} be the faces of $\mathcal{P}$ contained in Π_i, $I_0(\Pi_i)$ respectively. We claim that $\mathcal{P}$ has the combinatorial structure shown in Figure 2.2. In particular:

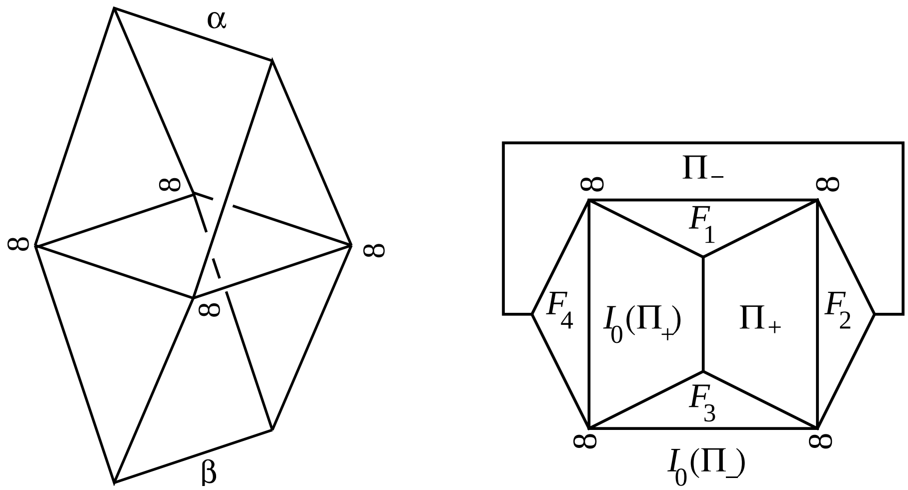


Figure 2.2. The polyhedron $\mathcal{P}(\alpha, \beta)$.

Proposition (2.1.9). *The polyhedron $\mathcal{P}$ has eight vertices. Four of these vertices are the fixed points of the parabolic maps $YX^{-1}Y^{-1}X$, $X^{-1}Y^{-1}XY$, $Y^{-1}XYX^{-1}$ and $XYX^{-1}Y^{-1}$. The other four are the intersection of the axes of the following pairs of transformations X, I_1; X, $I_0I_1I_0$; Y, I_2; Y, $I_0I_2I_0$. Every edge with (at least) one ideal endpoint has dihedral angle $\pi/2$.*

Proof. We will sketch the reason for this to be true. For example, Π_1 intersects Π_+ along the geodesic with one endpoint the fixed point of $YX^{-1}Y^{-1}X$ and passing through the intersection of the axes of X and I_2. We have already seen these two planes intersect orthogonally. Likewise, Π_1 intersects Π_- along the geodesic with endpoints the fixed points of $YX^{-1}Y^{-1}X$ and $X^{-1}Y^{-1}XY$. Again, we have seen that these planes intersect orthogonally. The other edges and vertices may be found similarly. □

Proposition (2.1.10). *The polyhedron $\mathcal{P}$ with the side pairings $X : F_1 \longrightarrow F_3$ and $Y : F_4 \longrightarrow F_2$ is a fundamental polyhedron for the convex core of the group $\langle X, Y\rangle$.*

Proof. Define

$$\mathcal{N} = \bigcup_{T \in \langle X,Y\rangle} T(\mathcal{P}).$$

We show that $\mathcal{N}$ is the smallest group invariant convex subset of $\mathbb{H}^3$ and so is the Nielsen region (convex hull of the limit set) of $\langle X, Y\rangle$. This means that $\mathcal{N}/\langle X, Y\rangle$ is the convex core.

It is clear that $\mathcal{P}$ is convex. Now consider $\mathcal{P}$ and $X(\mathcal{P})$. These two polyhedra share a face $F_3 = I_0(F_1) = X(F_1)$ (since F_1 is sent to itself by I_1 and $X = I_0I_1$). The dihedral angles along the three edges of $\mathcal{P}$ bounding F_3 are all $\pi/2$. Similarly, the dihedral angles along the three edges of $X(\mathcal{P})$ bounding $X(F_1)$ are all $\pi/2$. Thus gluing these two polyhedra along their common face gives another convex polyhedron. Proceeding by induction, we see that $\mathcal{N}$ itself is convex. Thus $\mathcal{N}$ contains the smallest $\langle X, Y\rangle$ invariant convex set, the Nielsen region.

The intersection of Π_+ with $\partial\mathcal{N}$ is formed by removing from Π_+ infinitely many hyperbolic halfspaces bounded by the axes of X, YXY^{-1} and all their images under $\langle X, YXY^{-1}\rangle$. This is the Nielsen region of this subgroup and so is contained in the Nielsen region of $\langle X, Y\rangle$. Similarly, every other face of $\mathcal{P}$ is contained in the Nielsen region of $\langle X, Y\rangle$. If the boundary of $\mathcal{N}$ is contained in the Nielsen region then, by convexity, the whole of $\mathcal{N}$ must be as well. Thus $\mathcal{N}$ both contains and is contained in the Nielsen region. This proves the result. □

(2.2) The trigonometry from bending formulae. In this section we use the bending formulae of [12] to show that the polyhedron $\mathcal{P}$ only depends on the dihedral angles across γ_X and γ_Y.

The only free parameters of $\mathcal{P}$ are the lengths and dihedral angles in the sides of $\mathcal{P}$ contained in the axes of X and Y. According to the above notation, α is the dihedral angle between Π_+ and $I_0(\Pi_+)$ along the axis of X and we define ℓ_α to be length of the corresponding side of $\mathcal{P}$. (We choose the convention that α is the interior angle of $\mathcal{P}$ and remark that this is the opposite convention to that used in [12].) Similarly, as above, β is the dihedral angle between Π_- and $I_0(\Pi_-)$ along the axis of Y and we define ℓ_β to be length of the corresponding side of $\mathcal{P}$.

In [12] formulae were developed that relate the lengths and complex shear along the pleating locus of convex hull boundaries. As indicated above, the bending angles of [12] are related to our angles by $\theta = \pi - \alpha$, $\phi = \pi - \beta$. Similarly, the length ℓ_α is the length of the geodesic represented by X and so is twice $\lambda(\mathbf{X})$. Similarly for ℓ_β. That is $\lambda(\mathbf{X}) = \ell_\alpha/2$ and $\lambda(\mathbf{Y}) = \ell_\beta/2$. In the proof of Theorem 6.1 of [12], it was shown that

$$\sinh\lambda(\mathbf{X}) = \sin(\phi/2)\cot(\theta/2), \qquad \sinh\lambda(\mathbf{Y}) = \sin(\theta/2)\cot(\phi/2).$$

In our notation, these formulae give us

Proposition (2.2.1). *The (essential) angles α, β and edge lengths ℓ_α, ℓ_β of $\mathcal{P}(\alpha, \beta)$ are related by*

$$\sinh(\ell_\alpha/2) = \cos(\beta/2)\tan(\alpha/2),\ \sinh(\ell_\beta/2) = \cos(\alpha/2)\tan(\beta/2). \tag{2.2.2}$$

These formulae indicate that the polyhedron $\mathcal{P}$ only depends on the angles α and β, where $\alpha, \beta \in (0, \pi)$. This justifies our notation $\mathcal{P} = \mathcal{P}(\alpha, \beta)$.

It is easy to see that formulae (2.2.2) imply the following:

Proposition (2.2.3) (Tangent Rule). *The (essential) angles α, β and the edge lengths ℓ_α, ℓ_β of the polyhedron $\mathcal{P}(\alpha, \beta)$ are related by*

$$\frac{\tan(\alpha/2)}{\tanh(\ell_\alpha/2)} = \frac{\tan(\beta/2)}{\tanh(\ell_\beta/2)} = T, \tag{2.2.4}$$

where T is a positive number given by

$$T^2 = \tan^2(\alpha/2) + \tan^2(\beta/2) + 1. \tag{2.2.5}$$

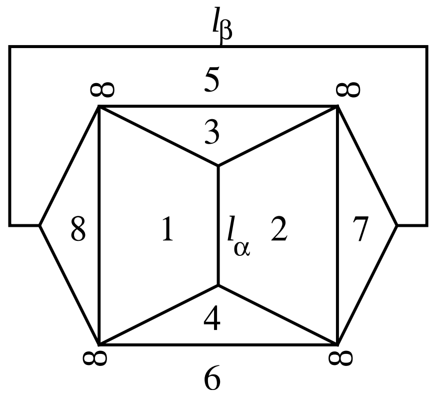


Figure 2.3. The projection of $\mathcal{P}(\alpha,\beta)$.

(2.3) The trigonometry from the Gram matrix. In this section we use the Gram matrix of the polyhedron to re-derive the formulae of the previous section.

Consider the numbering of faces of $\mathcal{P}(\alpha,\beta)$ as shown in its projection in Figure 2.3. Let $\rho(j,k)$ be the hyperbolic distance between the faces j and k. Then we write $A=\cosh\ell_\alpha=\cosh\rho(3,4)$, $B=\cosh\ell_\beta=\cosh\rho(7,8)$, $u=\cosh\rho(1,7)=\cosh\rho(2,8)$, $v=\cosh\rho(3,6)=\cosh\rho(4,5)$.

Denote by $G_{\alpha,\beta}$ the Gram matrix of the polyhedron $\mathcal{P}(\alpha,\beta)$:

$$G_{\alpha,\beta}=\begin{pmatrix} 1 & -\cos\alpha & 0 & 0 & -1 & -1 & -u & 0\\ -\cos\alpha & 1 & 0 & 0 & -1 & -1 & 0 & -u\\ 0 & 0 & 1 & A & 0 & -v & -1 & -1\\ 0 & 0 & -A & 1 & -v & 0 & -1 & -1\\ -1 & -1 & 0 & -v & 1 & -\cos\beta & 0 & 0\\ -1 & -1 & -v & 0 & -\cos\beta & 1 & 0 & 0\\ -u & 0 & -1 & -1 & 0 & 0 & 1 & -B\\ 0 & -u & -1 & -1 & 0 & 0 & -B & 1 \end{pmatrix}$$

Denote by $G(i_1,i_2,\ldots,i_k)$, $k\le 8$, the diagonal minor of $G_{\alpha,\beta}$, formed by rows and columns with numbers $i_1,i_2,\ldots,i_k$. Since the rank of $G_{\alpha,\beta}$ is equal to 4 the determinants of each of its 5×5–minors $\det G(i_1,i_2,i_3,i_4,i_5)$ vanishes. This gives equations relating the entries of $G_{\alpha,\beta}$. More precisely, taking (i_1,i_2,i_3,i_4,i_5) to be $(1,2,3,4,5)$, $(1,2,3,4,8)$, $(2,5,6,7,8)$, $(4,5,6,7,8)$, respectively, we will get following four equations.

$$v^2 = (A^2-1)\,\frac{1+\cos\alpha}{1-\cos\alpha}, \tag{2.3.1}$$

$$u^2 = (1-\cos^2\alpha)\,\frac{A+1}{A-1}, \tag{2.3.2}$$

$$u^2 = (B^2-1)\,\frac{1+\cos\beta}{1-\cos\beta}, \tag{2.3.3}$$

$$v^2 = (1-\cos^2\beta)\,\frac{B+1}{B-1}. \tag{2.3.4}$$

Recall that values A, B, u, v are greater than 1 in these equations. Taking $t = uv$ and calculating it in two ways using (2.3.1), (2.3.2) and using (2.3.3), (2.3.4) we obtain:

$$t = (1 + \cos\alpha)(A + 1) = (1 + \cos\beta)(B + 1). \tag{2.3.5}$$

Therefore

$$A = \frac{t}{1 + \cos\alpha} - 1, \qquad B = \frac{t}{1 + \cos\beta} - 1. \tag{2.3.6}$$

It is easy to see from (2.3.1), (2.3.4) and (2.3.6) that t satisfies the equation

$$(t - 2 - 2\cos\alpha)(t - 2 - 2\cos\beta) = (1 - \cos^2\alpha)(1 - \cos^2\beta).$$

This is equivalent to:

$$(t - 2 - \cos\alpha - \cos\beta)^2 = (1 - \cos\alpha\cos\beta)^2.$$

Therefore there are two possibilities. Either

$$t - 2 - \cos\alpha - \cos\beta = -1 + \cos\alpha\cos\beta$$

or

$$t - 2 - \cos\alpha - \cos\beta = 1 - \cos\alpha\cos\beta.$$

In the first case

$$t = (1 + \cos\alpha)(1 + \cos\beta),$$

which contradicts (2.3.5) since $A > 1$ and $B > 1$. In the second case

$$t = 4 - (1 - \cos\alpha)(1 - \cos\beta).$$

Hence

$$\cosh^2(\ell_\alpha/2) = \frac{A+1}{2} = \frac{t}{2 + 2\cos\alpha} = \frac{1 - \sin^2(\alpha/2)\sin^2(\beta/2)}{\cos^2(\alpha/2)}$$

and

$$\cosh^2(\ell_\beta/2) = \frac{B+1}{2} = \frac{t}{2 + 2\cos\beta} = \frac{1 - \sin^2(\alpha/2)\sin^2(\beta/2)}{\cos^2(\beta/2)}.$$

It easy to see that simplifying and taking square roots we will get the formulae (2.2.2) obtained earlier using the methods of [12]. Also, Proposition (2.2.3) follows immediately.

(2.4) Volume formulae. In this section we use the Schläfli formula and the computations of the previous sections to find the volume of $\mathcal{P}(\alpha, \beta)$.

Define $V = V(\alpha, \beta) = \textbf{Vol}\ \mathcal{P}(\alpha, \beta)$ to be the hyperbolic volume of $\mathcal{P}(\alpha, \beta)$. To find V we use the Schläfli formula (see [9] and [14] for details):

$$dV = -\frac{\ell_\alpha}{2}d\alpha - \frac{\ell_\beta}{2}d\beta.$$

Set $M = \tan(\alpha/2)$, $N = \tan(\beta/2)$ for $0 < \alpha, \beta < \pi$. Then $d\alpha = \dfrac{2dM}{1 + M^2}$ and $d\beta = \dfrac{2dN}{1 + N^2}$. Using equation (2.2.4), we obtain $\ell_\alpha = 2\operatorname{arctanh}(M/T)$ and $\ell_\beta = 2\operatorname{arctanh}(N/T)$. We have to integrate the differential form

$$\omega = -\frac{dV}{2} = \operatorname{arctanh}(M/T)\,\frac{dM}{1 + M^2} + \operatorname{arctanh}(N/T)\,\frac{dN}{1 + N^2},$$

where $T^2 = M^2 + N^2 + 1$. In order to do so, consider the extended differential form $\Omega = \Omega(M, N, T)$ of three independent variables M, N, T:

$$\begin{aligned}\Omega \quad = \quad & \operatorname{arctanh}(M/T)\,\frac{dM}{1+M^2} + \operatorname{arctanh}(N/T)\,\frac{dN}{1+N^2} \\ & + \log\left[\frac{(T^2-M^2)(T^2-N^2)}{(1+M^2)(1+N^2)}\right]\frac{dT}{1+T^2}.\end{aligned}$$

Note that Ω satisfies the following properties:

- Ω is smooth and exact in the region

$$G = \{(M, N, T) \in \mathbb{R}^3 : M > 0, N > 0, T > 0\};$$

- $\Omega = \omega$ for all $(M, N, T) \in G$ satisfying equation $T^2 = M^2 + N^2 + 1$.

Let us consider

$$W = W(M, N) = \int_T^{+\infty} \log\left[\frac{(t^2-M^2)(t^2-N^2)}{(1+M^2)(1+N^2)}\right]\frac{dt}{1+t^2}$$

where T is a positive root of the equation $T^2 = M^2 + N^2 + 1$. Straightforward calculations give

$$\frac{\partial W}{\partial M} = -\frac{2\operatorname{arctanh}(M/T)}{1+T^2}, \qquad \frac{\partial W}{\partial N} = -\frac{2\operatorname{arctanh}(N/T)}{1+T^2}$$

and $W(M, N) \to 0$ as $M, N \to \infty$.

Using $M = \tan(\alpha/2)$ and $N = \tan(\beta/2)$, we see that the volume function $V = V(\alpha, \beta) = V(M, N)$ satisfies the following conditions:

$$\begin{aligned}\frac{\partial V}{\partial M} \quad &= \quad \frac{\partial V}{\partial \alpha}\cdot\frac{\partial \alpha}{\partial M} = -\frac{\ell_\alpha}{2}\cdot\frac{2}{1+M^2} = -\frac{2\ \operatorname{arctanh}\ (M/T)}{1+M^2}, \\ \frac{\partial V}{\partial N} \quad &= \quad \frac{\partial V}{\partial \beta}\cdot\frac{\partial \beta}{\partial N} = -\frac{\ell_\beta}{2}\cdot\frac{2}{1+N^2} = -\frac{2\ \operatorname{arctanh}\ (N/T)}{1+N^2},\end{aligned}$$

and $V(M, N) \to 0$ as $M, N \to \infty$. The last follows from the fact that $\mathcal{P}(\alpha, \beta)$ collapses to be flat as α (or β) tends to π. Hence, we conclude that $V(M, N) = W(M, N)$ for all $M, N > 0$.

Theorem (2.4.1). *Let α and β be angles in the interval $(0, \pi)$. The volume of the polyhedron $\mathcal{P}(\alpha, \beta)$ is given by the formula*

$$\text{(2.4.2)} \qquad \mathbf{Vol}\ \mathcal{P}(\alpha, \beta) = \int_T^{\infty} \log\left[\frac{(t^2-M^2)(t^2-N^2)}{(1+M^2)(1+N^2)}\right]\frac{dt}{1+t^2},$$

where $M = \tan(\alpha/2)$, $N = \tan(\beta/2)$ and T is a positive root of the equation $T^2 = M^2 + N^2 + 1$.

Recall that the Lobachevsky function $\Lambda(x)$ is defined by the formula (see [9] and [14]):

$$\Lambda(x) = -\int_0^x \log \mid 2\sin\zeta \mid d\zeta.$$

To represent the volume of $\mathcal{P}(\alpha, \beta)$ in terms of the Lobachevsky function, we will use the following observation.

Lemma (2.4.3). *Consider*

$$\mathbf{I}(L,S)=\int_S^{+\infty}\log\left|\frac{\zeta^2-L^2}{1+L^2}\right|\frac{d\zeta}{1+\zeta^2},$$

where $L=\tan\mu$, $S=\tan\sigma$, *and* $0<\mu,\,\sigma<\pi$. *Then*

$$\mathbf{I}(L,S)=\Delta(\mu,\,\sigma)-\Delta(\pi/2,\,\sigma),$$

where $\Delta(\mu,\,\sigma)=\Lambda(\mu+\sigma)-\Lambda(\mu-\sigma)$.

Proof. Set $\zeta=\tan\tau$, $0\le\tau\le\pi/2$. We have

$$\begin{aligned}
\mathbf{I}(L,S) &= \int_S^{+\infty}\log\left|\frac{\zeta^2-L^2}{1+L^2}\right|\frac{d\zeta}{1+\zeta^2}\\
&= \int_\sigma^{\pi/2}\log\left|\frac{\sin(\tau-\mu)\sin(\tau+\mu)}{\sin(\tau-\pi/2)\sin(\tau+\pi/2)}\right|d\tau\\
&= \int_\sigma^{\pi/2}\log|2\sin(\tau-\mu)|\,d\tau+\int_\sigma^{\pi/2}\log|2\sin(\tau+\mu)|\,d\tau\\
&\quad -\int_\sigma^{\pi/2}\log|2\sin(\tau-\pi/2)|\,d\tau-\int_\sigma^{\pi/2}\log|2\sin(\tau+\pi/2)|\,d\tau\\
&= \int_{\sigma+\mu}^{\pi/2+\mu}\log|2\sin\eta|\;d\eta+\int_{\sigma-\mu}^{\pi/2-\mu}\log|2\sin\eta|\;d\eta\\
&\quad -\int_{\sigma+\pi/2}^{\pi}\log|2\sin\eta|\;d\eta-\int_{\sigma-\pi/2}^{0}\log 2\,|\sin\eta|\;d\eta\\
&= -\Lambda(\pi/2+\mu)+\Lambda(\sigma+\mu)-\Lambda(\pi/2-\mu)+\Lambda(\sigma-\mu)\\
&\quad +\Lambda(\pi)-\Lambda(\sigma+\pi/2)+\Lambda(0)-\Lambda(\sigma-\pi/2)\\
&= \Lambda(\mu+\sigma)-\Lambda(\mu-\sigma)-(\Lambda(\pi/2+\sigma)-\Lambda(\pi/2-\sigma))\\
&= \Delta(\mu,\,\sigma)-\Delta(\pi/2,\,\sigma),
\end{aligned}$$

where we used well-known properties of the Lobachevsky function (see [14] for details). □

From Theorem (2.4.1) and Lemma (2.4.3) we immediately get the following expression for the volume.

Corollary (2.4.4). *The volume of a convex hull* $\mathcal{P}(\alpha,\beta)$, *where* $0<\alpha,\beta<\pi$, *is given by the formula*

$$\mathbf{Vol}\,\mathcal{P}(\alpha,\beta)=\Delta(\alpha/2,\,\theta)+\Delta(\beta/2,\,\theta)-2\Delta(\pi/2,\,\theta),\tag{2.4.5}$$

where $\Delta(\mu,\,\sigma)=\Lambda(\mu+\sigma)-\Lambda(\mu-\sigma)$, *and* θ, *with* $0<\theta<\pi/2$, *is the principal parameter defined by* $\tan^2\theta=\tan^2(\alpha/2)+\tan^2(\beta/2)+1$.

As observed above, the polyhedron $\mathcal{P}(\alpha,\beta)$ is four copies of the degenerate Lambert cube $\mathcal{L}(\alpha/2,\beta/2,0)$. Moreover, the parameter θ, $0<\theta<\pi/2$, such that $T=\tan\theta$ for T defined by (2.2.5), is the principal parameter of the Lambert cube $\mathcal{L}(\alpha/2,\beta/2,0)$ introduced in [6]. Thus, the expression for the volume from (2.4.5) is, naturally, four times more than the expression for the volume of the Lambert cube $\mathcal{L}(\alpha/2,\beta/2,0)$ given by R. Kellerhals in [6].

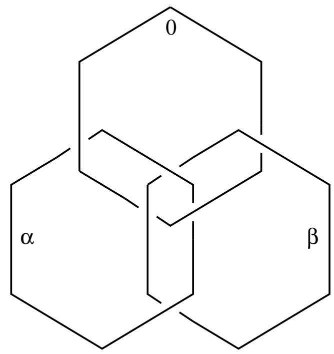


Figure 2.4. The Borromean rings.

(2.5) The associated cone manifolds. It is interesting remark that volumes of convex hulls coincide or are commensurable with volumes of well-known cone-manifolds.

For the case $\beta = \alpha$ we have

Corollary (2.5.1). *The volume of a convex hull $\mathcal{P}(\alpha, \alpha)$, $0 < \alpha < \pi$ is given by the formula*

$$\textbf{Vol}\ \mathcal{P}(\alpha,\alpha) = 2\int_{\alpha}^{\pi} \operatorname{arcsinh}\,(\sin(\zeta/2))\,d\zeta. \tag{2.5.2}$$

Proof. We have $\frac{d}{d\alpha}V(\alpha,\alpha) = 2\frac{\partial V}{\partial \alpha} = -\ell_\alpha$, $\tanh(\ell_\alpha/2) = \frac{M}{T}$, and $T^2 = 2\tan^2(\alpha/2) + 1 = 2M^2 + 1$. Hence

$$\sinh^2(\ell_\alpha/2) = \frac{\tanh^2(\ell_\alpha/2)}{1-\tanh^2(\ell_\alpha/2)} = \frac{M^2}{T^2 - M^2} = \frac{M^2}{M^2+1} = \sin^2(\alpha/2),$$

that is $\sinh(\ell_\alpha/2) = \sin(\alpha/2)$. Since $V(\pi,\pi) = 0$ the result follows. □

The formula we have obtained coincides with the volume formula for the Whitehead cone-manifold $\mathcal{W}(\alpha, 0)$ whose singular set is the Whitehead link with the cone angle α on one cusp and the complete hyperbolic structure on the other (see [11]).

Denote by $\mathcal{B}(\alpha, \beta, 0)$ a Borromean cone-manifold whose singular set are Borromean rings with cone angles α and β on two components and a complete hyperbolic cusp on the third one (see Figure 2.4).

Recall that the fundamental set of $\mathcal{B}(\alpha, \beta, 0)$ consists of eight copies of the Lambert cube $\mathcal{L}(\alpha/2, \beta/2, 0)$ (see, for example [5]). Hence we immediately get the following

Proposition (2.5.3). *The volume of the convex hull $\mathcal{P}(\alpha, \alpha)$ coincides with the volume of the Whitehead link cone-manifold $\mathcal{W}(\alpha, 0)$. The volume of the convex hull $\mathcal{P}(\alpha, \beta)$ is equal to one half of the volume of the Borromean cone-manifold $\mathcal{B}(\alpha, \beta, 0)$.*

3. The case where XY and XY^{-1} are purely hyperbolic

(3.1) Constructing the polyhedron. Let matrices $\mathbf{X}, \mathbf{Y} \in \mathrm{SL}(2,\mathbb{C})$ with $\mathrm{tr}[\mathbf{X},\mathbf{Y}] = -2$ represent isometries X and Y of $\mathbf{H}^3$ which generate a punctured torus group. For the rest of this section we suppose that $\mathrm{tr}(\mathbf{XY})$ and $\mathrm{tr}(\mathbf{XY}^{-1})$ are both real and greater than 2. Thus both XY and XY^{-1} are purely hyperbolic. We will show that either $\langle X, Y\rangle$ is Fuchsian or else the axes of XY and XY^{-1} are the pleating loci of the convex hull boundary.

From the expression for the trace of the commutator given above (2.1.5), we see that the traces of $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{XY}$ and the traces of $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{XY}^{-1}$ satisfy the Markov equations (see [12]):

$$\begin{aligned} \mathrm{tr}^2(\mathbf{X}) + \mathrm{tr}^2(\mathbf{Y}) + \mathrm{tr}^2(\mathbf{XY}) &= \mathrm{tr}(\mathbf{X})\,\mathrm{tr}(\mathbf{Y})\,\mathrm{tr}(\mathbf{XY}) \\ \mathrm{tr}^2(\mathbf{X}) + \mathrm{tr}^2(\mathbf{Y}) + \mathrm{tr}^2(\mathbf{XY}^{-1}) &= \mathrm{tr}(\mathbf{X})\,\mathrm{tr}(\mathbf{Y})\,\mathrm{tr}(\mathbf{XY}^{-1}) \end{aligned}$$

As above, to simplify the notation, we define

$$\begin{aligned} x &= \cosh\lambda(\mathbf{X}) = \frac{1}{2}\,\mathrm{tr}(\mathbf{X}), \\ y &= \cosh\lambda(\mathbf{Y}) = \frac{1}{2}\,\mathrm{tr}(\mathbf{Y}), \\ A &= \cosh\lambda(\mathbf{XY}) = \frac{1}{2}\,\mathrm{tr}(\mathbf{XY}), \\ B &= \cosh\lambda(\mathbf{XY}^{-1}) = \frac{1}{2}\,\mathrm{tr}(\mathbf{XY}^{-1}). \end{aligned}$$

From the Markov equations we see that A and B are the two roots of the equation

$$t^2 - 2xy\,t + x^2 + y^2 = 0.$$

Therefore, by the Vietta theorem,

$$\begin{aligned} 2xy &= A + B > 2 \\ x^2 + y^2 &= AB > 1. \end{aligned}$$

In particular, both of these quantities are real. We obtain the following analogue of Proposition (2.1.1).

Proposition (3.1.1). *Suppose that $\langle X, Y\rangle$ is a punctured torus group for which $A = \cosh\lambda(\mathbf{XY}) > 1$ and $B = \cosh\lambda(\mathbf{XY}^{-1}) > 1$.*

(i) *If $(A + B) \le AB$ then $\langle X, Y\rangle$ is Fuchsian.*

(ii) *If $(A + B) > AB$ then $\cosh\lambda(\mathbf{Y}) = \overline{\cosh\lambda(\mathbf{X})}$, which is not real.*

Proof. (i) In this case we have, by hypothesis, that

$$\begin{aligned} 0 \le AB - A - B = x^2 + y^2 - 2xy = (x - y)^2, \\ 0 < AB + A + B = x^2 + y^2 + 2xy = (x + y)^2. \end{aligned}$$

Therefore $x - y$ and $x + y$ are real, and so x and y are both real. Thus we have $\mathrm{tr}(\mathbf{X}) = 2x$, $\mathrm{tr}(\mathbf{Y}) = 2y$ and $\mathrm{tr}(\mathbf{XY}) = 2A$ all being real. Therefore $\langle X, Y\rangle$ maps a hyperplane in $\mathbb{H}^3$ to itself. Thus $\langle X, Y\rangle$ is a two generator group of isometrics of the hyperbolic plane for which the commutator of the generators is parabolic. Hence this group is discrete [8].

(ii) In this case, by hypothesis, we have

$$0 > AB - A - B = x^2 + y^2 - 2xy = (x - y)^2.$$

Therefore $x - y$ is purely imaginary. Together with the fact that $x + y$ is real we see that $\cosh \lambda(\mathbf{X}) = x$ and $\cosh \lambda(\mathbf{Y}) = y$ are (non-real) complex conjugates of one another. □

In what follows we will be interested in the case where $AB < A + B$, that is the non-Fuchsian case. Unless we indicate otherwise we always assume that we are in this case.

LEMMA (3.1.2). *Let $\langle X, Y\rangle$ be a punctured torus group where $A = \cosh \lambda(\mathbf{XY})$ and $B = \cosh \lambda(\mathbf{XY}^{-1})$ are both real and greater than 1. Then the axes of X and Y intersect with angle $\theta(X, Y)$ where*

$$\cos^2 \theta(X, Y) = \frac{(A - B)^2}{(A - B)^2 + 4}.$$

Proof. Calculating the complex distance $\delta(X, Y)$ as before, we obtain equation (2.1.4). Squaring this expression and substituting for $2xy = A+B$ and $x^2+y^2 = AB$ we find that

$$\cosh^2 \delta(X, Y) = \frac{(A - B)^2}{(A - B)^2 + 4}.$$

As this is real and less than 1 we see that $\Re\, \delta(X, Y) = 0$. In other words, the axes of X and Y intersect with angle $\Im\, \delta(X, Y) = \theta(X, Y)$. This gives the result. □

Now we will construct the fundamental polyhedron for the convex core of $\langle X, Y\rangle$. This will be $\widetilde{\mathfrak{Q}} = \widetilde{\mathfrak{Q}}(\alpha, \beta)$.

Since $\mathbf{XY}$, $(\mathbf{YX})^{-1}$ and their product $\mathbf{XYX}^{-1}\mathbf{Y}^{-1}$ all have real trace, the corresponding isometries XY, $(YX)^{-1}$ and $XYX^{-1}Y^{-1}$ generate a Fuchsian group. Likewise, since $\mathbf{XY}^{-1}$, $(\mathbf{Y}^{-1}\mathbf{X})^{-1}$ and $\mathbf{XY}^{-1}\mathbf{X}^{-1}\mathbf{Y}$ all have real trace, the corresponding isometries XY^{-1}, $(Y^{-1}X)^{-1}$ and $XY^{-1}X^{-1}Y$ too generate a Fuchsian group.

- Let Π_+ be the plane preserved by the group $\langle XY, YX\rangle$;
- Let Π_- be the plane preserved by the group $\langle XY^{-1}, Y^{-1}X\rangle$.

Following the construction in Section 2.1, we define geodesics γ_X, γ_Y and γ_0 by

- γ_X is the axis of X;
- γ_Y is the axis of Y;
- γ_0 is the common perpendicular of γ_X and γ_Y.

We define the following halfturns:

- Let I_0 denote the halfturn fixing γ_0.
- Define I_1 by $I_1 = I_0X$. Then I_1 is a halfturn fixing a geodesic γ_1.
- Define I_2 by $I_2 = YI_0$. Then I_2 is a halfturn fixing a geodesic γ_2.

Thus γ_1 is orthogonal to γ_X and the complex distance along γ_X between γ_0 and γ_1 is $\lambda(\mathbf{X})$. Similarly, γ_2 is orthogonal to γ_Y and the complex distance between γ_0 and γ_1 is $\lambda(\mathbf{Y}) = \overline{\lambda(\mathbf{X})}$. Moreover, we have

$$\begin{aligned} I_0XI_0 &= X^{-1}, \quad I_1XI_1 = X^{-1}, \quad I_2XI_2 = YX^{-1}Y^{-1}, \\ I_0YI_0 &= Y^{-1}, \quad I_1YI_1 = X^{-1}Y^{-1}X, \quad I_2YI_2 = Y^{-1}. \end{aligned}$$

We claim that

LEMMA (3.1.3). *The geodesic γ_0 is orthogonal to Π_+ and Π_-.*

Proof. We have

$$I_0\,(XY)\,I_0 = (YX)^{-1}, \qquad I_0\,(XY^{-1})\,I_0 = (Y^{-1}X)^{-1}.$$

In other words I_0 interchanges the axes of XY and YX and so preserves Π_+. Likewise, I_0 interchanges the axes of XY^{-1} and $Y^{-1}X$ and so preserves Π_-. Moreover,

$$I_0\,(XYX^{-1}Y^{-1})\,I_0 = (YXY^{-1}X^{-1})^{-1},$$
$$I_0\,(XY^{-1}X^{-1}Y)\,I_0 = (YX^{-1}Y^{-1}X)^{-1}.$$

Thus I_0 swaps the fixed points of $XYX^{-1}Y^{-1}$ and $YXY^{-1}X^{-1}$ which lie on the boundary of Π_+. Since these two fixed points are not separated by the axes of XY and YX, elementary plane hyperbolic geometry shows that I_0 acts on Π_+ as a rotation. Similarly, I_0 swaps the fixed points of $XY^{-1}X^{-1}Y$ and $YX^{-1}Y^{-1}X$ and so acts on Π_- as a rotation. This gives the result. □

Consider a plane Π_0 containing γ_0 so that the angle between Π_0 and γ_X is the same as the angle between Π_0 and γ_Y. There are two planes with this property. Let Π_0 be the plane separating $\gamma_X \cap \gamma_1$ and $\gamma_Y \cap \gamma_2$. Let Π_1 be the other such plane.

- Let R_0 be reflection in Π_0;
- Let R_1 be reflection in Π_1;
- Then $R_0(\gamma_X) = R_1(\gamma_X) = \gamma_Y$;
- $R_0(\gamma_1) = \gamma_2$;
- $R_0R_1 = I_0$.

For the penultimate line we used $\cosh\lambda(\mathbf{Y}) = \overline{\cosh\lambda(\mathbf{X})}$. Furthermore, we have

$$R_0I_0R_0 = I_0, \quad R_0I_1R_0 = I_2, \quad R_0I_2R_0 = I_1.$$

Hence

$$R_0XR_0 = R_0I_0I_1R_0 = I_0I_2 = Y^{-1}, \quad R_0YR_0 = R_0I_2I_0R_0 = I_1I_0 = X^{-1}.$$

Because $I_0 = R_0R_1$, we see that Π_1 contains γ_0 and that Π_0 and Π_1 are orthogonal.

$$R_1XR_1 = R_0I_0XI_0R_0 = Y, \quad R_1YR_1 = R_0I_0YI_0R_0 = X.$$

Lemma (3.1.4). *The planes Π_0 and Π_1 satisfy:*

(i) Π_0 *is orthogonal to the axes of XY and YX, and hence to Π_+;*
(ii) Π_0 *is orthogonal to Π_- and contains the fixed points of parabolic isometries $XY^{-1}X^{-1}Y$ and $YX^{-1}Y^{-1}X$;*
(iii) Π_1 *is orthogonal to the axes of XY^{-1} and $Y^{-1}X$, and hence to Π_-;*
(iv) Π_1 *is orthogonal to Π_+ and contains the fixed points of parabolic isometries $XYX^{-1}Y^{-1}$ and $YXY^{-1}X^{-1}$.*

Proof. We prove (i) and (ii). Parts (iii) and (iv) will follow similarly.

$$R_0(XY)R_0 = (XY)^{-1}, \qquad R_0(YX)R_0 = (YX)^{-1}$$

and so R_0 preserves the axes of XY and YX. Hence Π_0 is orthogonal to Π_+. Similarly,

$$R_0(XY^{-1})R_0 = Y^{-1}X, \qquad R_0(Y^{-1}X)R_0 = XY^{-1}$$

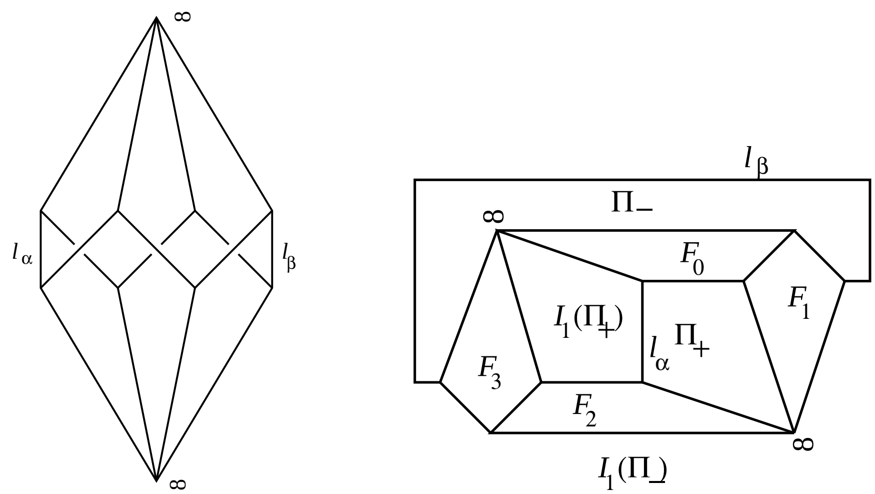


Figure 3.1. The polyhedron $\mathcal{Q}(\alpha, \beta)$.

and so R_0 swaps the axes of XY^{-1} and $Y^{-1}X$. Hence it preserves Π_- and so Π_0 is orthogonal to Π_-. Furthermore,

$$\begin{aligned} R_0(XY^{-1}X^{-1}Y)R_0 &= (XY^{-1}X^{-1}Y)^{-1}, \\ R_0(YX^{-1}Y^{-1}X)R_0 &= (YX^{-1}Y^{-1}X)^{-1}. \end{aligned}$$

Thus R_0 fixes their fixed points, which must lie in Π_0. □

Let $\mathcal{Q}$ be the hyperbolic polyhedron formed by the common intersection of halfspaces bounded by Π_+, Π_-, Π_0, Π_1 and their images under I_1. For $i = 0, 1$ let F_i, F_{i+2} be the face of $\mathcal{Q}$ contained in Π_i, $I_1(\Pi_i)$ respectively (see Figure 3.1).

The intersection of the faces Π_+ and $I_1(\Pi_+)$ is the segment of the axis of YX with length $\ell_\alpha = \lambda(\mathbf{XY})$. Let us denote the dihedral angle at this edge by α. (This is twice the angle between the axis of I_1 and the plane Π_+.) Similarly, the intersection of the faces Π_- and $I_1(\Pi_-)$ is the segment of the axis of $Y^{-1}X$ with length $\ell_\beta = \lambda(\mathbf{XY}^{-1})$. We denote the dihedral angle at this edge by β. (This is twice the angle between the axis of I_1 and the plane Π_-.) By the construction, all other dihedral angles of $\mathcal{Q}$ are right angles.

We see that planes Π_1 and $I_1(\Pi_0)$ meet at the fixed point of the parabolic isometry $X^{-1}Y^{-1}XY$. This point is also on Π_+ and $I_1(\Pi_-)$. Therefore faces F_1 and F_2 have a common point at infinity. Likewise, Π_0 and $I_1(\Pi_1)$ meet at the fixed point of the isometry $YX^{-1}Y^{-1}X$ which also lies on Π_- and $I_1(\Pi_+)$. Similarly, F_0 and F_3 have a common point at infinity too. All other vertices of $\mathcal{Q}$ are ordinary. To summarise:

Proposition (3.1.5). *The polyhedron $\mathcal{Q}$ has ten vertices. Two of these are ideal vertices and are the fixed points of $X^{-1}Y^{-1}XY$ and $YX^{-1}Y^{-1}X$. The other eight vertices are finite and correspond to the intersection of the axes of YX, $Y^{-1}X$, I_0 and $I_1I_0I_1$ with the common perpendiculars of the axes of I_0, YX; I_0, $Y^{-1}X$; $I_1I_0I_1$, YX; $I_1I_0I_1$, $Y^{-1}X$.*

Let $\mathcal{Q}'$ be the hyperbolic polyhedron formed by the common intersection of halfspaces bounded by Π_+, Π_-, Π_0, Π_1 and their images under I_2. For $i = 0, 1$ let F_i', F_{i+2}' be the face of $\mathcal{Q}'$ contained in Π_i, $I_2(\Pi_i)$ respectively. Clearly R_0 swaps $\mathcal{Q}$ and $\mathcal{Q}'$. Denote $\widetilde{\mathcal{Q}} = \mathcal{Q} \cup \mathcal{Q}'$.

Proposition (3.1.6). *The polyhedron $\widetilde{\mathcal{Q}} = \mathcal{Q} \cup \mathcal{Q}'$ with the side pairings*

$$Id : F_0 \longrightarrow F_0', \quad Y : F_1 \longrightarrow F_3', \quad YX : F_2 \longrightarrow F_2', \quad X : F_3 \longrightarrow F_1'.$$

is a fundamental domain for the convex core of the group $\langle X, Y \rangle$.

Proof. The proof is similar to the proof of Proposition (2.1.10). Again, it is clear that

$$\mathcal{N} = \bigcup_{T \in \langle X, Y \rangle} T(\mathcal{Q} \cup \mathcal{Q}')$$

is invariant under $\langle X, Y \rangle$ and is convex.

The fact that the boundary of $\mathcal{N}$ consists of the orbit of Nielsen regions of the Fuchsian subgroups $\langle XY, YX \rangle$ and $\langle XY^{-1}, Y^{-1}X \rangle$ means that it is contained in the Nielsen region of $\langle X, Y \rangle$. This gives the result. $\square$

(3.2) The trigonometry from bending formulae. In this section we use the bending formulae of [12] to show that $\mathcal{Q}$ only depends on the dihedral angles across the axes of XY and XY^{-1}.

The only free parameters for $\mathcal{Q}$ are the lengths and dihedral angles in the sides of $\mathcal{Q}$ contained in the axes of XY and $Y^{-1}X$. According to the above notation, α is the dihedral angle between Π_+ and $I_1(\Pi_+)$ along the axis of YX and we define ℓ_α to be length of the corresponding side of $\mathcal{Q}$. According to the above notation, β is the dihedral angle between Π_- and $I_1(\Pi_-)$ along the axis of $Y^{-1}X$ and we define ℓ_β to be length of the corresponding side of $\mathcal{Q}$.

We now show how to relate α, β, ℓ_α, ℓ_β using the formulae of Parker and Series [12]. Now the pleating loci are next-but-one neighbours with common neighbour X. It is easy to see that the real part of the translation along XY is half the length of this curve, that is $\lambda(\mathbf{XY})$. Also, from the way the polyhedron is constructed, we see that ℓ_α is $\lambda(\mathbf{XY})$. Likewise for the other face. Therefore, using the formula (I) of [12] (with $U = X$) first with $\lambda(W) = \ell_\alpha$, $\tau = \ell_\alpha + i(\pi - \alpha)$ and then with $\lambda(W) = \ell_\beta$, $\tau = \ell_\beta + i(\pi - \beta)$ we obtain

$$\cosh^2 \lambda(\mathbf{X}) = \frac{\cosh^2(\ell_\alpha/2 + i(\pi - \alpha)/2)}{\tanh^2 \ell_\alpha} = \frac{\cosh^2(\ell_\beta/2 + i(\pi - \beta)/2)}{\tanh^2 \ell_\beta}.$$

Taking square roots and equating the real and imaginary parts we obtain

$$\begin{aligned} \frac{\cosh(\ell_\alpha/2)\sin(\alpha/2)}{\tanh(\ell_\alpha)} &= \pm\frac{\cosh(\ell_\beta/2)\sin(\beta/2)}{\tanh(\ell_\beta)}, \\ \frac{\sinh(\ell_\alpha/2)\cos(\alpha/2)}{\tanh(\ell_\alpha)} &= \pm\frac{\sinh(\ell_\beta/2)\cos(\beta/2)}{\tanh(\ell_\beta)}. \end{aligned}$$

Squaring and using the duplication formula for cos and cosh we obtain

Proposition (3.2.1). *The (essential) angles α, β and the edge lengths ℓ_α, ℓ_β of $\mathcal{Q} = \mathcal{Q}(\alpha, \beta)$ are related by*

$$\frac{A^2(1 - \cos\alpha)}{A - 1} = \frac{B^2(1 - \cos\beta)}{B - 1}, \qquad \frac{A^2(1 + \cos\alpha)}{A + 1} = \frac{B^2(1 + \cos\beta)}{B + 1},$$

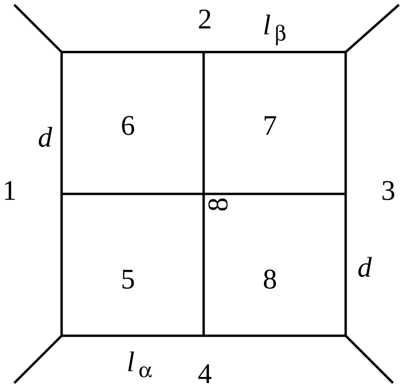


Figure 3.2. The polyhedron $\mathcal{Q}(\alpha, \beta)$.

where $A = \cosh \ell_\alpha$ *and* $B = \cosh \ell_\beta$.

These formulae indicate that the polyhedron $\mathcal{Q}$ only depends on the angles α and β, where $\alpha, \beta \in (0, \pi)$. This justifies our notation $\mathcal{Q} = \mathcal{Q}(\alpha, \beta)$ and $\widetilde{\mathcal{Q}} = \widetilde{\mathcal{Q}}(\alpha, \beta)$. In the next section we will see how to write A and B in terms of $\cos \alpha$ and $\cos \beta$.

It easy to see from proposition (3.2.1) that the following relation follows:

$$\frac{\tan(\alpha/2)}{\tanh(\ell_\alpha/2)} = \frac{\tan(\beta/2)}{\tanh(\ell_\beta/2)} = T$$

for some parameter T, but to find it we need to know more relations between essential angles and lengths. The effective way to obtain these relations is to consider the Gram matrix.

(3.3) The trigonometry from the Gram matrix. In this section we consider the Gram matrix of $\mathcal{Q}$ and re-derive the formulae from the previous section.

Consider the numbering of faces of $\mathcal{Q}(\alpha, \beta)$, for $0 < \alpha, \beta < \pi$, as shown on its projection in Figure 3.2. Let $\rho(i, j)$ be the distance between faces of $\mathcal{Q}(\alpha, \beta)$ with numbers i and j. Denote $A = \cosh \ell_\alpha$, $B = \cosh \ell_\beta$, $u = \cosh d = \cosh \rho(2, 5) = \cosh \rho(4, 7)$, $v = \cosh \rho(2, 8) = \cosh \rho(1, 7)$, $w = \cosh \rho(3, 5) = \cosh \rho(4, 6)$.

Remark that the edges marked by d (which also denotes their lengths) are common perpendiculars to faces 2 and 5, and faces 4 and 7. As we see from the construction, d is distance between planes Π_+ and Π_-.

Let $G_{\alpha,\beta}$ be the Gram matrix of the polyhedron $\mathcal{Q}(\alpha, \beta)$:

$$G_{\alpha,\beta} = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 0 & -v & -A \\ 0 & 1 & 0 & -1 & -x & 0 & -\cos\beta & -v \\ -1 & 0 & 1 & 0 & -w & -B & 0 & 0 \\ 0 & -1 & 0 & 1 & -\cos\alpha & -w & -u & 0 \\ 0 & -u & -w & -\cos\alpha & 1 & 0 & -1 & 0 \\ 0 & 0 & -B & -w & 0 & 1 & 0 & -1 \\ -v & -\cos\beta & 0 & -u & -1 & 0 & 1 & 0 \\ -A & -v & 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}$$

Denote by $G(i_1, i_2, \dots, i_k)$, $k \le 8$, the diagonal minor of $G_{\alpha,\beta}$, formed by rows and columns with numbers $i_1, i_2, \dots, i_k$. Since the rank of $G_{\alpha,\beta}$ is equal to 4 the determinants of each of its 5×5–minors $\det G(i_1, i_2, i_3, i_4, i_5)$ will vanish. Again, this gives relations between the entries of $G_{\alpha,\beta}$. Taking the minors corresponding to the columns $(1,2,4,5,6)$, $(1,2,5,6,7)$, $(1,2,4,5,8)$, $(2,3,4,6,8)$, $(1,2,5,6,8)$, $(1,2,3,5,6)$ respectively, we obtain following six equations.

$$w^2(u^2-1) = (u+\cos\alpha)^2, \tag{3.3.1}$$

$$v^2(u^2-1) = (u+\cos\beta)^2, \tag{3.3.2}$$

$$(A^2-1)(u+\cos\alpha)^2 = v^2(1-\cos^2\alpha), \tag{3.3.3}$$

$$(B^2-1)(u+\cos\beta)^2 = w^2(1-\cos^2\beta), \tag{3.3.4}$$

$$v^2 = A^2(u^2-1), \tag{3.3.5}$$

$$w^2 = B^2(u^2-1). \tag{3.3.6}$$

Recall that quantities A, B, u, v and w are greater than 1 in these equations. From equations (3.3.1) and (3.3.6) we get

$$B(u^2-1) = (u+\cos\alpha). \tag{3.3.7}$$

Substituting (3.3.6) into (3.3.4) we have

$$B^2(1+u\cos\beta)^2 - (u+\cos\beta)^2 = 0.$$

Factorising this equation and substituting for B from (3.3.7) we obtain

$$f_{\alpha,\beta}(u)\, g_{\alpha,\beta}(u) = 0,$$

where

$$f_{\alpha,\beta}(u) = u^3 - u(2+\cos\alpha\cos\beta) - \cos\alpha - \cos\beta \tag{3.3.8}$$

and

$$g_{\alpha,\beta}(u) = u^3 + 2u^2\cos\beta + u\cos\alpha\cos\beta + \cos\alpha - \cos\beta.$$

Analogously, from (3.3.2) and (3.3.5) we get

$$A(u^2-1) = (u+\cos\beta), \tag{3.3.9}$$

and substituting (3.3.5) into (3.3.3), we obtain

$$A^2(1+u\cos\alpha)^2 - (u+\cos\alpha)^2 = 0,$$

which gives

$$f_{\alpha,\beta}(u)\, g_{\beta,\alpha}(u) = 0.$$

Therefore, u is a root of the equation

$$f_{\alpha,\beta}(u)\, h_{\alpha,\beta}(u) = 0,$$

where

$$h_{\alpha,\beta}(u) = g_{\alpha,\beta}(u) - g_{\beta,\alpha}(u) \;=\; 2(u^2-1)(\cos\beta - \cos\alpha).$$

We remark that equations $h_{\alpha,\beta}(u) = 0$ and $g_{\alpha,\alpha}(u) = 0$ have no roots with $u > 1$. Therefore, $f_{\alpha,\beta}(u) = 0$. It is easy to see that if $\alpha \ne \pi$ and $\beta \ne \pi$ then $f_{\alpha,\beta}(1) < 0$. Furthermore, $f_{\alpha,\beta}(u)$ is strictly increasing on the interval $(1,\infty)$, and so has only one root u with $u > 1$. Using (3.3.7) and (3.3.9) to substitute for $\cos(\alpha)$ and $\cos(\beta)$ in (3.3.8) we find that for such a root u we have

$$u = \frac{A+B}{AB}. \tag{3.3.10}$$

Therefore we obtain

$$\cos\alpha = \frac{A+B-A^2B}{A^2}, \qquad \cos\beta = \frac{A+B-AB^2}{B^2}. \tag{3.3.11}$$

These equations are equivalent to the formulae in Proposition (3.2.1).

From this it is easy to see that the following four conditions are equivalent: (i) $u=1$; (ii) $\cos\alpha=-1$; (iii) $\cos\beta=-1$; (iv) $A+B=AB$. They correspond to the case when the polyhedron $\mathcal{Q}(\alpha,\beta)$ has collapsed. As we saw in Proposition (3.1.1), the polyhedron $\mathcal{Q}(\alpha,\beta)$ is non-degenerate if and only if $A+B>AB$.

Substituting for $B=A/(Au-1)$ and $A=B/(Bu-1)$ into the expressions for $\cos\alpha$ and $\cos\beta$ in (3.3.11) gives

$$\cos\alpha = \frac{u-A}{Au-1}, \qquad \cos\beta = \frac{u-B}{Bu-1}.$$

Rearranging gives

$$A = \frac{u+\cos\alpha}{1+u\cos\alpha}, \qquad B = \frac{u+\cos\beta}{1+u\cos\beta}. \tag{3.3.12}$$

Combining these with the expression for u given in (3.3.8) we obtain:

Proposition (3.3.13). *For a non-degenerate polyhedron $\mathcal{Q}(\alpha,\beta)$, the parameters $A=\cosh\ell_\alpha$ and $B=\cosh\ell_\beta$ can be found by*

$$A = \frac{u+\cos\alpha}{1+u\cos\alpha} \qquad B = \frac{u+\cos\beta}{1+u\cos\beta},$$

where $u>1$ is the root of the equation

$$u^3 - u(2+\cos\alpha\cos\beta) - \cos\alpha - \cos\beta = 0. \tag{3.3.14}$$

Recall that by definition $u=\cosh d$, where d is distance between planes Π_+ and Π_-. Set $T=\coth(d/2)$, and note that $T^2=(u+1)/(u-1)$. Using standard relations

$$\cos\nu = \frac{1-\tan^2(\nu/2)}{1+\tan^2(\nu/2)}, \qquad \cosh\mu = \frac{1+\tanh^2(\mu/2)}{1-\tanh^2(\mu/2)},$$

we are able to rewrite the above proposition in the following way:

Proposition (3.3.15) (Tangent Rule). *The (essential) angles α, β and the edge lengths ℓ_α, ℓ_β of the polyhedron $\mathcal{Q}(\alpha,\beta)$ are related by*

$$\frac{\tan(\alpha/2)}{\tanh(\ell_\alpha/2)} = \frac{\tan(\beta/2)}{\tanh(\ell_\beta/2)} = T, \tag{3.3.16}$$

where T is a positive number given by $T^2=(u+1)/(u-1)$, and u is a root of the equation (3.3.14).

Remark that $u=(T^2+1)/(T^2-1)$, and it follows from (3.3.14) that u satisfies the equation

$$(u^2-1)^2 = (u\cos\alpha+1)(u\cos\beta+1).$$

By direct computations, we see that T satisfies the equation

$$\frac{T^2-M^2}{1+M^2}\,\frac{T^2-N^2}{1+N^2}\left[\frac{T^2-1}{2T^2}\right]^2 = 1,$$

where $M=\tan(\alpha/2)$ and $N=\tan(\beta/2)$.

(3.4) Volume formulae. In this section we use Schläfli's formula to find the volume of $\mathcal{Q}(\alpha,\beta)$.

By the construction, the volume of the convex hull $\widetilde{\mathcal{Q}}(\alpha,\beta)$ is twice the volume of the polyhedron $\mathcal{Q}(\alpha,\beta)$. To find the volume of the latter polyhedron, we will use the method of the extended Schläfli differential form. Consider Schläfli's differential form

$$\omega = d\,\mathbf{Vol}\,\mathcal{Q}(\alpha,\beta) = -\frac{1}{2}\left(\ell_\alpha d\alpha + \ell_\beta d\beta\right)$$

defined for $0<\alpha,\beta<\pi$. Let us extend it to a differential form $\Omega=\Omega(\alpha,\beta,u)$ of three independent variables α,β,u:

$$\Omega = -\frac{1}{2}\left(\ell_\alpha d\alpha + \ell_\beta d\beta + \ell_u du\right),$$

where u plays a role of the principal parameter. We have to choose Ω in such a way that following properties are satisfied:

- Ω is smooth and exact in the region
$$G = \{(\alpha,\beta,u)\in\mathbb{R}^3 : 0<\alpha<\pi, 0<\beta<\pi, u>1\};$$
- $\Omega=\omega$ for all $(\alpha,\beta,u)\in G$ satisfying equation (3.3.14).

Since Ω is supposed to be exact, we have

$$\frac{\partial\ell_u}{\partial\alpha} = \frac{\partial\ell_\alpha}{\partial u} = \frac{\partial}{\partial u}\left(\operatorname{arccosh}\frac{u+\cos\alpha}{1+u\cos\alpha}\right) = \frac{\sin\alpha}{(1+u\cos\alpha)\sqrt{u^2-1}}.$$

So

$$\begin{aligned}
\ell_u &= \int\frac{\sin\alpha\,d\alpha}{(1+u\cos\alpha)\sqrt{u^2-1}}\\
&= -\frac{1}{u\sqrt{u^2-1}}\log(1+u\cos\alpha)+C(u,\beta)\\
&= -\frac{1}{u\sqrt{u^2-1}}\log\frac{(1+u\cos\alpha)(1+u\cos\beta)}{(u^2-1)^2}.
\end{aligned}$$

We note that for $u>1$ the equation (3.3.14) is equivalent to

$$\frac{(1+u\cos\alpha)(1+u\cos\beta)}{(u^2-1)^2}=1.$$

If this condition is satisfied, we have $\ell_u=0$ and consequently $\Omega=\omega$.

Applying the same arguments as in Theorem (2.4.1), we find

Theorem (3.4.1). *The volume of the convex hull $\widetilde{\mathcal{Q}}(\alpha,\beta)$ is given by*

$$\mathbf{Vol}\,\widetilde{\mathcal{Q}}(\alpha,\beta) = \int_1^u \log\left[\frac{(1+\zeta\cos\alpha)(1+\zeta\cos\beta)}{(\zeta^2-1)^2}\right]\cdot\frac{d\zeta}{\zeta\sqrt{\zeta^2-1}},$$

where $u>1$ is the root of the equation (3.3.14).

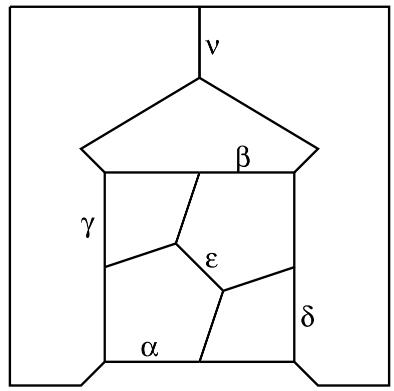


Figure 3.3. The polyhedron $\mathcal{O}(\alpha, \beta, \gamma, \delta, \varepsilon, \nu)$.

If $\alpha = \beta$, then $u = \frac{1}{2}(\cos\alpha + \sqrt{8 + \cos^2\alpha})$, and

$$\begin{aligned}\mathbf{Vol}\,\widetilde{\mathcal{Q}}(\alpha, \alpha) &= 2\int_1^u \log\frac{1 + \zeta\cos\alpha}{\zeta^2 - 1} \cdot \frac{d\zeta}{\zeta\sqrt{\zeta^2 - 1}} \\ &= 2\int_\alpha^\pi \operatorname{arccosh}\frac{\sqrt{8 + \cos^2\alpha} - \cos\alpha}{2}\, d\alpha.\end{aligned}$$

Now we want to express **Vol** $\widetilde{\mathcal{Q}}(\alpha, \beta)$ in term of the Lobachevsky function. To do this we write $M = \tan(\alpha/2)$ and $N = \tan(\beta/2)$ and make the following substitutions in the integral of Theorem (3.4.1): $\zeta = (t^2 + 1)/(t^2 - 1)$, $\cos\alpha = (1 - M^2)/(1 + M^2)$, and $\cos\beta = (1 - N^2)(1 + N^2)$. As a result we obtain:

Corollary (3.4.2). *The volume of the convex hull* $\widetilde{\mathcal{Q}}(\alpha, \beta)$ *is given by*

$$\mathbf{Vol}\,\widetilde{\mathcal{Q}}(\alpha, \beta) = 2\int_T^{+\infty} \log\left|\frac{(t^2 - M^2)(t^2 - N^2)}{(1 + M^2)(1 + N^2)}\left(\frac{t^2 - 1}{2t^2}\right)^2\right| \frac{dt}{1 + t^2},$$

where $M = \tan(\alpha/2)$, $N = \tan(\beta/2)$ *and* $T = \coth(d/2)$ *is the variable from the tangent rule.*

Using this result and Lemma (2.4.3) we have

Corollary (3.4.3). *The volume of the convex hull* $\widetilde{\mathcal{Q}}(\alpha, \beta)$ *is given by*

$$\begin{aligned}\mathbf{Vol}\,\widetilde{\mathcal{Q}}(\alpha, \beta) \quad = \quad & 2\Delta(\alpha/2, \theta) + 2\Delta(\beta/2, \theta) + \\ & 4\Delta(\pi/4, \theta) - 4\Delta(0, \theta) - 4\Delta(\pi/2, \theta),\end{aligned}$$

where $\Delta(\mu, \sigma) = \Lambda(\mu + \sigma) - \Lambda(\mu - \sigma)$, *and* θ *is a principal parameter such that* $T = \tan\theta$.

In particular, $\mathbf{Vol}\,\mathcal{Q}(0, 0) = 2.53735\ldots$, which is the maximal volume for the family $\mathcal{Q}(\alpha, \beta)$. Moreover, $\mathbf{Vol}\,\mathcal{Q}(\pi/2, \pi/2) = 1.83193\ldots$ which is one-half of the volume of the ideal right-angled octahedron.

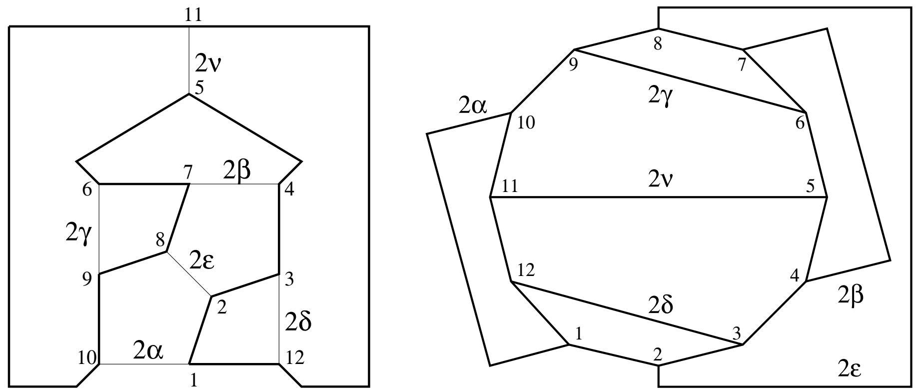


Figure 3.4. Two representations of the singular set of the orbifold Ω^+.

(3.5) The associated cone manifolds. In this section we will determine a link which is naturally related with the polyhedron $\mathcal{Q}(\alpha,\beta)$ in the same manner as the Lambert cube is related with the Borromean rings.

In order to do this we consider $\mathcal{Q}(\alpha,\beta)$ as a particular case of a more general polyhedron $\mathcal{O} = \mathcal{O}(\alpha,\beta,\gamma,\delta,\varepsilon,\nu)$ (see Figure 3.3). The dihedral angles of $\mathcal{O}$ are equal to α, β, γ, δ, ε, ν on edges labelled by these letters, and are $\pi/2$ on the other edges. We allow for angles α, β, γ, δ, ε, ν to be zero. In this case the corresponding edges become ideal vertices of polyhedra with a complete hyperbolic structure. Note that for $\alpha=\beta=\gamma=\delta=\varepsilon=\nu=0$ the polyhedron $\mathcal{O}$ is a regular right angled octahedron. The existence of $\mathcal{O}$ in the hyperbolic space $\mathbb{H}^3$ for all $0 \le \alpha,\beta,\gamma,\delta,\varepsilon,\nu < \pi$ follows from Rivin's theorem [13]. We remark that $\mathcal{Q}(\alpha,\beta) = \mathcal{O}(\alpha,\beta,\pi/2,\pi/2,0,0)$.

Consider a hyperbolic cone-manifold Ω whose underlying space is the polyhedron $\mathcal{O}$ and whose singular set consists of faces, edges and vertices of $\mathcal{O}$. Let Ω^+ be an orientable double of Ω. Then Ω^+ can be obtained by gluing together $\mathcal{O}$ and its mirror image along their common boundary. As a result, Ω^+ can be recognised as a hyperbolic cone-manifold with the 3-sphere as its underlying space and whose singular set is formed by the edges of $\mathcal{O}$ with cone angles twice the dihedral ones (see Figure 3.4, where unlabelled edges correspond to cone angles π).

To construct the two-fold covering we will use the approach from [10] based on the properties of the Hamiltonian circuit. Note that unbranched edges form a Hamiltonian circuit λ passing through all vertices of the singular set of Ω^+. Consider a two-fold covering $\Sigma \to \Omega^+$ of Ω^+ branched over the cycle λ. Since λ is unknotted in Ω^+, the underlying space of Σ is the 3-sphere again. The singular set of Σ is a six component link L formed by lifting the labelled edges. To recognise this link we represent λ as a circle with 12 vertices as in the right hand figure of Figure 3.4. After taking the two-fold covering branched along λ we obtain the link L (see Figure 3.5).

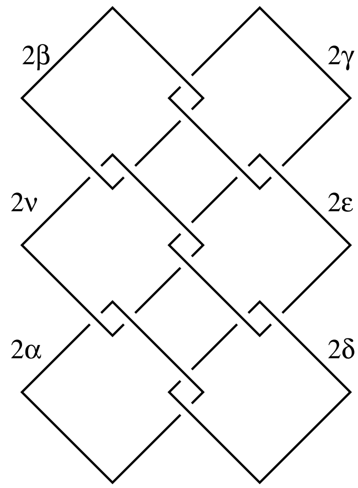


Figure 3.5. The link L.

Hence $\Sigma = \Sigma(2\alpha, 2\beta, 2\gamma, 2\delta, 2\epsilon, 2\nu)$ is a hyperbolic cone-manifold with singular set illustrated in Figure 3.5. By the construction we have

$$\mathbf{Vol}\, \mathcal{O}(\alpha, \beta, \gamma, \delta, \varepsilon, \nu) = \frac{1}{4}\, \mathbf{Vol}\, \Sigma(2\alpha, 2\beta, 2\gamma, 2\delta, 2\varepsilon, 2\nu). \tag{3.5.1}$$

In particular, we obtain

Proposition (3.5.2). *The volume of the convex hull* $\widetilde{\mathcal{Q}}(\alpha, \beta)$ *is equal to one half of the volume of cone-manifold* $\Sigma(2\alpha, 2\beta, \pi, \pi, 0, 0)$.

This statement gives us a convenient way to calculate the volume of $\widetilde{\mathcal{Q}}(\alpha, \beta)$ using J. Weeks' computer program *SnapPea* [15].

Received June 25, 2003

Final version received April 14, 2004

A.D. Mednykh
A.Yu. Vesnin
Sobolev Institute of Mathematics
Novosibirsk 630090
Russia
mednykh@math.nsc.ru
vesnin@math.nsc.ru

J.R. Parker
Department of Mathematical Sciences
University of Durham
Durham, DH1 3LE
England
j.r.parker@durham.ac.uk

References

[1] F. Bonahon, *A Schläfli-type formula for convex cores of hyperbolic 3-manifolds*, J. Diff. Geom. **50** (1998), 25–58.

[2] R.D. Canary, D.B.A. Epstein, P. Green, *Notes on notes of Thurston*, Analytical and geometric aspects of hyperbolic space (Coventry/Durham, 1984), 3–92, London Math. Soc. Lecture Note Ser. **111**, Cambridge Univ. Press, Cambridge, 1987.

[3] D.B.A. Epstein, A. Marden, *Convex hulls in hyperbolic space, a theorem of Sullivan, and measured pleated surfaces*, Analytical and geometric aspects of hyperbolic space (Coventry/Durham, 1984), 113–253, London Math. Soc. Lecture Note Ser. **111**, Cambridge Univ. Press, Cambridge, 1987.

[4] W. Fenchel, Elementary Geometry in Hyperbolic Space, de Gruyter, 1989.

[5] H. Hilden, M.T. Lozano, J.M. Montesinos-Amilibia, *On the Borromean orbifolds: Geometry and arithmetic*, Topology '90, Contrib. Res. Semester Low Dimensional Topol., Columbus/OH (USA) 1990, Ohio State Univ. Math. Res. Inst. Publ. **1** (1992), 133–167.

[6] R. Kellerhals, *On the volume of hyperbolic polyhedra,* Math. Ann. **285** (1989), 541–569.

[7] E. Klimenko, *Some examples of discrete groups and hyperbolic orbifolds of infinite volume,* J. Lie Theory **11** (2001), 491–503.

[8] P.J. Matelski, *The classification of discrete* 2*-generator subgroups of* $\mathrm{PSL}(2,\mathbb{R})$*,* Israel J. Math. **42** (1982), 309–317.

[9] J. Milnor, *The Schläfli differential equality,* in: Collected papers. Vol. 1. Geometry, Publish or Perish, Inc., Houston, TX, 1994, 281–295.

[10] A.D. Mednykh, A.Yu. Vesnin, *Three-dimensional hyperelliptic manifolds and Hamiltonian graphs,* Siberian Math. J. **40** (1999), 628–643.

[11] A.D. Mednykh, A.Yu. Vesnin, *On the volume of hyperbolic Whitehead link cone-manifolds,* SCIENTIA, Series A: Mathematical Sciences, **8** (2002), 1–11.

[12] J.R. Parker, C. Series, *Bending formulae for convex hull boundaries,* J. Anal. Math. **67** (1995), 165–198.

[13] I. Rivin, *A characterization of ideal polyhedra in hyperbolic* 3*-space,* Ann. of Math. (2) **143** (1996), no. 1, 51–70.

[14] E.B. Vinberg, Geometry II: spaces of constant curvature, Encyclopaedia of Mathematical Sciences. Berlin: Springer-Verlag. **29** (1993).

[15] J. Weeks, *SnapPea*, Software for hyperbolic 3-manifolds, available at `http://ftp.geometrygames.org/SnapPea/index.html`

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

HOPF CONSTRUCTION MAP IN HIGHER DIMENSIONS

GUILLERMO MORENO

Abstract. We study the zero set of the Hopf construction map $F_n : \mathbb{A}_n \times \mathbb{A}_n \to \mathbb{A}_n \times \mathbb{A}_0$ given by $F_n(x, y) = (2xy, ||y||^2 - ||x||^2)$ for $n \geq 4$, where $\mathbb{A}_n$ is the Cayley-Dickson algebra of dimension 2^n on $\mathbb{R}$.

0. Introduction

Let $f_1 : S^3 \to S^2, f_2 : S^7 \to S^4$ and $f_3 : S^{15} \to S^8$ be the classical *Hopf maps*; these can be defined using *the Hopf construction*. Let $\mathbb{A}_1 = \mathbb{C}, \mathbb{A}_2 = \mathbb{H}$ and $\mathbb{A}_3 = \mathbb{O}$ be the complex, quaternion and octonion numbers respectively, and let $F_n : \mathbb{A}_n \times \mathbb{A}_n \to \mathbb{A}_n \times \mathbb{R}$ be given by

$$F_n(x, y) = (2xy, ||y||^2 - ||x||^2)$$

for $n = 1, 2, 3$. Write $S^{2^{n+1}-1} = \{(x, y) \in \mathbb{A}_n \times \mathbb{A}_n : ||x||^2 + ||y||^2 = 1\}$. By definition,

$$F_n|S^{2^{n+1}-1} = f_n$$

are the Hopf maps. Since $\mathbb{A}_n$ is a normed real algebra of dimension 2^n, for $n = 1, 2, 3$ we have

$$\begin{aligned} ||(2xy, ||y||^2 - ||x||^2)||^2 &= 4||xy||^2 + (||y||^2 - ||x||^2)^2 \\ &= 4||x||^2||y||^2 + ||y||^4 + ||x||^4 - 2||x||^2||y||^2 \\ &= (||x||^2 + ||y||^2)^2, \end{aligned}$$

so if $||x||^2 + ||y||^2 = 1$, then $||F_n(x, y)|| = ||(2xy, ||y||^2 - ||x||^2)|| = 1$.

Using the Cayley-Dickson doubling process ([D]) define

$$\mathbb{A}_{n+1} = \mathbb{A}_n \times \mathbb{A}_n$$

with

$$(a, b)(x, y) = (ax - \overline{y}b, ya + b\overline{x}) \text{ for } a, b, x \text{ and } y \text{ in } \mathbb{A}_n$$

and

$$\overline{x} = (\overline{x}_1, -x_2) \text{ if } x = (x_1, x_2) \text{ is in } \mathbb{A}_{n-1} \times \mathbb{A}_{n-1}.$$

Thus if $\mathbb{A}_0 = \mathbb{R}$ with $\overline{x} = x$ for x a real number, then $\mathbb{A}_1 = \mathbb{C}, \mathbb{A}_2 = \mathbb{H}$ and $\mathbb{A}_3 = \mathbb{O}$, which are normed algebras; i.e., $||xy|| = ||x||||y||$ for all x, y in $\mathbb{A}_n$.

For $n \geq 4$, $\mathbb{A}_n$ is no longer normed and has zero divisors as well (see [K-Y] and [Mo1]).

Let us define $X_n^\infty = \{(x, y) \in \mathbb{A}_n \times \mathbb{A}_n | F_n(x, y) = (0, 0)\}$ and, for any non-negative real number r, $(x, y) \in X_n^r$ if and only if $xy = 0$ and $||x|| = ||y|| = r$. It

2000 *Mathematics Subject Classification*: 17A99, 55Q25.

Keywords and phrases: Cayley-Dickson algebras, alternative algebras, zero divisors, flexible algebra, normed algebra.

is clear that for real numbers $r > 0$ and $s > 0$, X_n^r is homeomorphic to X_n^s. Let us further define $X_n := X_n^1$.

The set X_n appears in some important problems in algebraic topology:
(1) Cohen's approach to the Arf invariant one problem (see [C1] and [C2]).
(2) The Adem-Lam construction of normed and non-singular bilinear maps (see [A] and [L]).

In this paper we will show that for $n \geq 4$, X_n is related to certain Stiefel manifolds; using the algebra structure in $\mathbb{A}_{n+1}$ we will construct a chain of inclusions

$$X_n \subset W_{2^{n-1}-1,2} \subset V_{2^n-2,2} \subset V_{2^n-1,2}$$

(see section 2) where $V_{m,2}$ and $W_{m,2}$ denote the real and complex Stiefel manifolds of 2-frames in $\mathbb{R}^m$ and $\mathbb{C}^m$, respectively.

In section 3 we show that we can attach to every element in $W_{2^{n-1},2}$, in a canonical way, an eight dimensional vector subspace of $\mathbb{A}_{n+1}$, and that, only for the elements in X_n, this vector subspace is isomorphic, as an algebra, to $\mathbb{A}_3 = \mathbb{O}$.

In section 4 we describe X_n as a certain type of algebra monomorphisms from $\mathbb{A}_3 = \mathbb{O}$ to $\mathbb{A}_{n+1}$ for $n \geq 4$.

This paper is a sequel to [Mo1] and we use freely the results of [Sch]. We acknowledge with gratitude the hard work done by the reviewer.

1. Pure and doubly pure elements in $\mathbb{A}_{n+1}$

Throughout this paper we use the following notational conventions:
(1) Elements in $\mathbb{A}_n$ will be denoted by Latin characters $a, b, c, \ldots, x, y, z.$ and elements in $\mathbb{A}_{n+1}$ will be denoted by Greek characters $\alpha, \beta, \gamma, \ldots$ For example,

$$\alpha = (a, b) \in \mathbb{A}_n \times \mathbb{A}_n.$$

(2) When we need to represent elements in $\mathbb{A}_n$ as elements in $\mathbb{A}_{n-1} \times \mathbb{A}_{n-1}$ we use subscripts, for instance, $a = (a_1, a_2)$, $\quad b = (b_1, b_2)$, and so on, with a_1, a_2, b_1, b_2 in $\mathbb{A}_{n-1}$.

Let $\{e_0, e_1, \ldots, e_{2^n-1}\}$ denote the canonical basis in $\mathbb{A}_n$. Then by the doubling process,

$$\{(e_0, 0), (e_1, 0), \ldots, (e_{2^n-1}, 0), (0, e_0), \ldots, (0, e_{2^n-1})\}$$

is the canonical basis in $\mathbb{A}_{n+1} = \mathbb{A}_n \times \mathbb{A}_n$. By standard abuse of notation we also denote $e_0 = (e_0, 0), e_1 = (e_1, 0), \ldots, e_{2^n-1} = (e_{2^n-1}, 0), e_{2^n} = (0, e_0), \ldots, e_{2^{n+1}-1} = (0, e_{2^n-1})$ in $\mathbb{A}_{n+1}$.

For $\alpha = (a, b) \in \mathbb{A}_n \times \mathbb{A}_n = \mathbb{A}_{n+1}$ we write $\widetilde{\alpha} = (-b, a)$ (the complexification of α), so $\widetilde{e}_0 = (0, e_0)$ and $\alpha\widetilde{e}_0 = (a, b)(0, e_0) = (-b, a) = \widetilde{\alpha}$. Notice that $\widetilde{\widetilde{\alpha}} = -\alpha$.

The *trace* on $\mathbb{A}_{n+1}$ is the linear map $t_{n+1} : \mathbb{A}_{n+1} \to \mathbb{R}$ given by $t_{n+1}(\alpha) = \alpha + \overline{\alpha} = 2$ (real part of α), so $t_{n+1}(\alpha) = t_n(a)$ when $\alpha = (a, b) \in \mathbb{A}_n \times \mathbb{A}_n$.

Definition (1.1). $\alpha = (a, b)$ in $\mathbb{A}_{n+1}$ *is pure* if

$$t_{n+1}(\alpha) = t_n(a) = 0.$$

Furthermore, $\alpha = (a, b)$ in $\mathbb{A}_{n+1}$ *is doubly pure* if it is pure and also $t_n(b) = 0$; i.e., $\widetilde{\alpha}$ is pure in $\mathbb{A}_{n+1}$.

Note that $2\langle a,b\rangle = t_n(a\overline{b})$ when $\langle -,-\rangle$ is the inner product in $\mathbb{R}^{2^n}$ (see [A]). Also, for pure elements a and b, $a\perp b$ if and only if $ab=-ba$.

Notation (1.2). ${}_o\mathbb{A}_n = \{e_o\}^\perp \subset \mathbb{A}_n$ is the vector subspace consisting of pure elements in $\mathbb{A}_n$; i.e., ${}_o\mathbb{A}_n = \mathrm{Ker}(t_n) = \mathbb{R}^{2^n-1}$. $\widetilde{\mathbb{A}}_{n+1} = {}_o\mathbb{A}_n \times {}_o\mathbb{A}_n = \{e_0, \widetilde{e}_0\}^\perp = \mathbb{R}^{2^{n+1}-2}$ is the vector subspace consisting of doubly pure elements in $\mathbb{A}_{n+1}$.

LEMMA (1.3). *For a and b in $\widetilde{\mathbb{A}_n}$ we have that*
(1) $a\widetilde{e}_0 = \widetilde{a}$ *and* $\widetilde{e}_0 a = -\widetilde{a}$;
(2) $a\widetilde{a} = -||a||^2\widetilde{e}_0$ *and* $\widetilde{a}a = ||a||^2\widetilde{e}_0$ *so* $a\perp\widetilde{a}$;
(3) $\widetilde{a}b = -\widetilde{ab}$ *with a a pure element;*
(4) $a\perp b$ *if and only if* $\widetilde{a}b + \widetilde{b}a = 0$;
(5) $\widetilde{a}\perp b$ *if and only if* $ab = \widetilde{b}\widetilde{a}$;
(6) $a\perp b$ *and* $\widetilde{a}\perp b$ *if and only if* $\widetilde{a}b = a\widetilde{b}$.

Proof. Note that a is pure if and only if $\overline{a}=-a$; and if $a=(a_1,a_2)$ is doubly pure, then $\overline{a}_1 = -a_1$ and $\overline{a}_2=-a_2$. Then
(1) $\widetilde{e}_0 a = (0,e_0)(a_1,a_2) = (-\overline{a}_2, \overline{a}_1) = (a_2,-a_1) = -(-a_2,a_1) = -\widetilde{a}$.
(2) $a\widetilde{a} = (a_1,a_2)(-a_2,a_1) = (-a_1a_2 + a_1a_2, a_1^2 + a_2^2) = (0, -||a||^2 e_0) = -||a||^2\widetilde{e}_0$.
Similarly $\widetilde{a}a = (-a_2,a_1)(a_1,a_2) = (-a_2a_1 + a_2a_1, -a_2^2 - a_1^2) = ||a||^2\widetilde{e}_0$.
Now, since $-2\langle\widetilde{a},a\rangle = a\widetilde{a} + \widetilde{a}a = 0$ we have $a\perp\widetilde{a}$.
(3) $\widetilde{a}b = (-a_2,a_1)(b_1,b_2) = (-a_2b_1 + b_2a_1, -b_2a_2 - a_1b_1)$.
So $\widetilde{\widetilde{a}b} = (a_1b_1 + b_2a_2, b_2a_1 - a_2b_1) = (a_1,a_2)(b_1,b_2) = ab$ *and then* $-\widetilde{a}b = \widetilde{ab}$.
Note that in this proof we only use that $\overline{a}_1 = -a_1$; i.e., a is pure and b doubly pure.
(4) $a\perp b \Leftrightarrow ab + ba = 0 \Leftrightarrow ab = -ba \Leftrightarrow \widetilde{ab} = -\widetilde{ba}$.
$\Leftrightarrow -\widetilde{a}b = \widetilde{b}a \Leftrightarrow \widetilde{a}b + \widetilde{b}a = 0$ by (3).
(5) $\widetilde{a}\perp b \Leftrightarrow \widetilde{\widetilde{a}}b + \widetilde{b}\widetilde{a} = 0$ (by (4)) $\Leftrightarrow -ab + \widetilde{b}\widetilde{a} = 0$.
(6) If $\widetilde{a}\perp b$ and $a\perp b$, then by (3) and (4) $\widetilde{a}b = -\widetilde{ab} = \widetilde{ba} = -\widetilde{b}a = a\widetilde{b}$.
Conversely, put $a=(a_1,a_2)$ and $b=(b_1,b_2)$ in $\mathbb{A}_{n-1}\times\mathbb{A}_{n-1}$ and define
$c := (a_1b_1 + b_2a_2)$ and $d := (b_2a_1 - a_2b_1)$ in $\mathbb{A}_{n-1}$.
Then $a\widetilde{b} = (a_1,a_2)(-b_2,b_1) = (-a_1b_2 + b_1a_2, b_1a_1 + a_2b_2)$ so $a\widetilde{b} = (-\overline{d},\overline{c})$.
Now $ab = (a_1,a_2)(b_1,b_2) = (a_1b_1 + b_2a_2, b_2a_1 - a_2b_1) = (c,d)$, so $\widetilde{ab} = (-d,c)$ and then $\widetilde{a}b = (d,-c)$. Thus if $a\widetilde{b} = \widetilde{a}b$, then $\overline{c}=-c$ and $d = -\overline{d}$, and then

$$\begin{aligned} t_n(ab) &= t_{n-1}(c) = c + \overline{c} = 0 \text{ and } a\perp b \\ t_n(\widetilde{a}b) &= t_{n-1}(d) = d + \overline{d} = 0 \text{ and } \widetilde{a}\perp b. \end{aligned}$$

□

COROLLARY (1.4). *For each $a\neq 0$ in $\widetilde{\mathbb{A}}_n$, the four dimensional vector subspace generated by $\{e_0,\widetilde{a},a,\widetilde{e}_0\}$ is a copy of $\mathbb{A}_2 = \mathbb{H}$. We denote it by $\mathbb{H}_a$.*

Proof. We suppose that $||a||=1$, otherwise we consider $\frac{a}{||a||}$. Construct the following multiplication table.

	e_0	$\widetilde{a}$	a	$\widetilde{e}_0$
e_0	e_0	$\widetilde{a}$	a	$\widetilde{e}_0$
$\widetilde{a}$	$\widetilde{a}$	$-e_0$	$+\widetilde{e}_0$	$-a$
a	a	$-\widetilde{e}_0$	$-e_0$	$\widetilde{a}$
$\widetilde{e}_0$	$\widetilde{e}_0$	a	$-\widetilde{a}$	$-e_0$

By lemma (1.3), $a\widetilde{e}_0 = \widetilde{a}$; $\widetilde{e}_0 a = -\widetilde{a}$; $\widetilde{a}\widetilde{e}_0 = \widetilde{\widetilde{a}} = -a$; $\widetilde{e}_0\widetilde{a} = -\widetilde{\widetilde{a}} = a$; $a\widetilde{a} = -\widetilde{e}_0$ and $\widetilde{a}a = \widetilde{e}_0$.

Identifying $e_0 \leftrightarrow e_0$, $\widetilde{a} \leftrightarrow e_1$, $a \leftrightarrow e_2$ and $\widetilde{e}_0 \leftrightarrow e_3$ we have the multiplication table for $\mathbb{A}_2 = \mathbb{H}$. □

2. The Stiefel manifold $V_{2^n-1,2}$ in $\mathbb{A}_{n+1}$ and a T^2-action

Let $\langle a, b\rangle_n$ denote the standard inner product of a and b in $\mathbb{A}_n = \mathbb{R}^{2^n}$. Now by [A] and [Mo$_1$],

$$2\langle a, b\rangle_n = (a\overline{b} + b\overline{a}) = t_n(a\overline{b})\,.$$

It is also well known that for $\alpha = (a, b)$ and $\chi = (x, y)$ in $\mathbb{A}_n \times \mathbb{A}_n = \mathbb{A}_{n+1}$ we have

$$\langle \alpha, \chi\rangle_{n+1} = \langle a, x\rangle_n + \langle \overline{b}, \overline{y}\rangle_n\,.$$

In particular, if α and χ are doubly pure elements in $\mathbb{A}_{n+1}$, then y and b are pure elements in $\mathbb{A}_n$, and therefore

$$\langle \alpha, \chi\rangle_{n+1} = \langle a, x\rangle_n + \langle b, y\rangle_n\,.$$

Lemma (2.1). *For $\alpha = (a, b)$ in $\mathbb{A}_{n+1}$ define $\hat{\alpha} := (b, a)$. For $\alpha \in \widetilde{\mathbb{A}}_{n+1}$ we have that*

(i) *$\langle \alpha, \hat{\alpha}\rangle_{n+1} = 0$ i.e. $\alpha \perp \hat{\alpha}$ in $\mathbb{A}_{n+1}$ if and only if $\langle a, b\rangle_n = 0$, i.e. $a \perp b$ in $\mathbb{A}_n$;*

(ii) *$\langle \widetilde{\alpha}, \hat{\alpha}\rangle_{n+1} = 0$, i.e. $\widetilde{\alpha} \perp \hat{\alpha}$ in $\mathbb{A}_{n+1}$ if and only if $||a|| = ||b||$ in $\mathbb{A}_n$.*

Proof. (i) $\langle \alpha, \hat{\alpha}\rangle_{n+1} = \langle (a, b), (b, a)\rangle_{n+1} = 2\langle a, b\rangle_n$.

(ii) $\langle \widetilde{\alpha}, \hat{\alpha}\rangle_{n+1} = \langle (-b, a), (b, a)\rangle_{n+1} = -\langle b, b\rangle_n + \langle a, a\rangle_n = -||b||^2 + ||a||^2$. □

By §1 we know that, for each $\alpha \neq 0$ in $\widetilde{\mathbb{A}}_{n+1}$, $\mathbb{H}_\alpha = \text{Span}\,\{e_0, \widetilde{\alpha}, \alpha, \widetilde{e}_0\}$ is a copy of $\mathbb{A}_2$, and that, if $\mathbb{H}_\alpha^\perp$ denotes the orthogonal complement of $\mathbb{H}_\alpha$, then $\mathbb{A}_{n+1} = \mathbb{H}_\alpha \oplus \mathbb{H}_\alpha^\perp$.

Since α is doubly pure, $\hat{\alpha}$ is also doubly pure; i.e., $\hat{\alpha} \in \{e_0, \widetilde{e}_0\}^\perp$. If $\hat{\alpha} \perp \alpha$ and $\hat{\alpha} \perp \widetilde{\alpha}$, then $\hat{\alpha} \in \mathbb{H}_\alpha^\perp$. Let $S^{\sqrt{2}}(\widetilde{\mathbb{A}}_{n+1}) = S^{2^{n+1}-3}$ denote the sphere of radius $\sqrt{2}$ in of $\widetilde{\mathbb{A}}_{n+1}$. Thus we have a description of the real Stiefel manifold of 2-orthonormal frames in $\mathbb{R}^{2^n-1}$ as follows:

$$V_{2^n-1,2} = \{(a, b) \in {}_o\mathbb{A}_n \times {}_o\mathbb{A}_n = \widetilde{\mathbb{A}}_{n+1} : ||a|| = ||b|| = 1, a \perp b\}$$

and

$$V_{2^n-1,2} = \{\alpha \in S^{\sqrt{2}}(\widetilde{\mathbb{A}}_{n+1}) : \hat{\alpha} \in \mathbb{H}_\alpha^\perp\}\,.$$

Lemma (2.2). *If r and s are in $\mathbb{R}$ with $r^2 + s^2 = 1$ and $(a, b) \in V_{2^n-1,2}$ then $(ra - sb, sa + rb) \in V_{2^n-1,2}$.*

Proof. Suppose that $||a|| = ||b|| = 1$ and $a \perp b$ in $\mathbb{A}_n$. Then $||ra - sb||^2 = r^2||a||^2 + s^2||b||^2 - 2rs\langle a, b\rangle_n$ and $||sa + rb||^2 = s^2||a||^2 + r^2||b||^2 + 2rs\langle a, b\rangle_n$, so $||ra - sb||^2 = ||sa + rb||^2 = r^2 + s^2 = 1$. Hence,

$$\begin{aligned} \langle ra - sb, sa + rb\rangle_n &= rs\langle a, a\rangle_n - sr\langle b, b\rangle_n - s^2\langle b, a\rangle_n + r^2\langle a, b\rangle_n \\ &= rs||a||^2 - rs||b||^2 + 0 = rs - sr = 0. \end{aligned}$$

□

COROLLARY (2.3). *The map $S^1 \times V_{2^n-1,2} \to V_{2^n-1,2}$ given by*

$$((r, s), \alpha) \mapsto r\alpha + s\widetilde{\alpha} = (ra - sb, sa + rb)$$

defines a smooth, free S^1-action on $V_{2^n-1,2}$.

Proof. Clearly $(1, 0) \cdot \alpha = \alpha$ and $(r, s)[(q, t) \cdot \alpha] = ((r, s)(q, t)) \cdot \alpha = (rq - st, rt + sq) \cdot \alpha$, so the map defines an action. It is a smooth action because it is a restriction of a linear action of $GL_2(\mathbb{R})$ on $\widetilde{\mathbb{A}}_{n+1} = \mathbb{R}^{2^{n+1}-2}$. Finally, the map is a free action: if $r\alpha + s\widetilde{\alpha} = \alpha$ then $r = 1$ and $s = 0$, because $\alpha \perp \widetilde{\alpha}$ in $\widetilde{\mathbb{A}}_{n+1}$. □

Next we identify $V_{2^n-2,2}$, the real Stiefel manifold of 2-orthonormal frames on $\mathbb{R}^{2^n-2}$, as a submanifold of $V_{2^n-1,2}$ as follows:

$$V_{2^n-2,2} = \{(a, b) \in V_{2^n-1,2} | (a, b) \in \widetilde{\mathbb{A}}_n \times \widetilde{\mathbb{A}}_n\};$$

i.e., $(a, b) \in V_{2^n-1,2}$ belongs to $V_{2^n-2,2}$ whenever a and b are doubly pure elements in $\mathbb{A}_n$. We have the known fibration [Wh]

$$\begin{array}{ccccc} S^{2^n-4} & \to & V_{2^n-2,2} & \to & S(\widetilde{\mathbb{A}}_n) = S^{2^n-3} \\ & & (a, b) & \mapsto & b \end{array}$$

Thus $V_{2^n-2,2}$ has dimension $2^n - 3 + 2^n - 4 = 2^{n+1} - 7$.

Since $(ra - sb)$ and $(sa + rb)$ are doubly pure elements in $\mathbb{A}_n$ when a and b are doubly pure elements, we have that $V_{2^n-2,2}$ *is a S^1-invariant submanifold of* $V_{2^n-1,2}$; i.e., if $\alpha \in V_{2^n-2,2}$ then $(r, s) \cdot \alpha \in V_{2^n-2,2}$ for all $(r, s) \in S^1$.

We note that $\mathbb{A}_{n+1}$ becomes a complex vector space by defining $i\alpha = \widetilde{\alpha}$; thus, as a complex vector space,

$$\widetilde{\mathbb{A}}_{n+1} = {}_0\mathbb{A}_n \times {}_0\mathbb{A}_n \cong \mathbb{C} \otimes_{\mathbb{R}} {}_0\mathbb{A}_n.$$

The isomorphism takes $1 \otimes x$ to $(x, 0)$ and $i \otimes y$ to $(0, y)$, and S^1 (the set of modulo 1 complex numbers) acts naturally by multiplication on $\mathbb{C}$, hence on $\widetilde{\mathbb{A}}_{n+1}$.

Now we identify the complex Stiefel manifold $W_{2^{n-1}-1,2}$ of 2-orthonormal frames in $\mathbb{C}^{2^{n-1}-1}$ as a submanifold of $V_{2^n-2,2}$ in terms of the Cayley-Dickson algebra $\mathbb{A}_{n+1}$ for $n \geq 3$.

It is known that for a, b, and x in $\mathbb{A}_n$, $\langle ax, b\rangle_n = \langle a, b\overline{x}\rangle_n$ (see [A]). Thus if x is a pure element, i.e., $\overline{x} = -x$, then $\langle ax, b\rangle_n = -\langle a, bx\rangle_n$. In other words, right multiplication by a pure non-zero element is a skew-symmetric linear map. In particular $\langle \widetilde{a}, b\rangle_n = -\langle a, \widetilde{b}\rangle_n$.

Proposition (2.4). *For $n \geq 3$, the map $\mathcal{H}_n : \widetilde{\mathbb{A}}_n \times \widetilde{\mathbb{A}}_n \to \mathbb{C}$ given by*

$$\mathcal{H}_n(a,b) = 2\langle a,b\rangle_n - 2i\langle \widetilde{a}, b\rangle_n$$

defines a Hermitian inner product in $\widetilde{\mathbb{A}}_n$.

Proof. Clearly $\mathcal{H}_n$ is $\mathbb{R}$-linear and

$$\begin{aligned} \overline{\mathcal{H}_n(a,b)} &= 2\langle a,b\rangle_n + 2i\langle \widetilde{a}, b\rangle_n \\ &= 2\langle a,b\rangle_n - 2i\langle a, \widetilde{b}\rangle_n \\ &= \mathcal{H}_n(b,a). \end{aligned}$$

On the other hand,

$$\begin{aligned} \mathcal{H}_n(\widetilde{a},b) &= 2\langle \widetilde{a},b\rangle_n - 2i\langle \widetilde{\widetilde{a}}, b\rangle_n \\ &= 2\langle \widetilde{a},b\rangle_n + 2i\langle a, b\rangle_n \\ &= 2i\langle a,b\rangle_n - 2i^2\langle \widetilde{a}, b\rangle_n \\ &= i\mathcal{H}_n(a,b). \end{aligned}$$

$\square$

Proposition (2.5). *For $n \geq 3$,*

$$W_{2^{n-1}-1,2} = \{(a,b) \in V_{2^n-2,2} | b \in \mathbb{H}_a^{\perp}\}.$$

Proof. First we observe that $b \in \mathbb{H}_a^{\perp}$ for a and b in $\widetilde{\mathbb{A}}_n$ if and only if $b \perp a$ and $b \perp \widetilde{a}$, i.e. $\mathcal{H}_n(a,b) = 0$. If $||a|| = ||b|| = 1$ and $\mathcal{H}_n(a,b) = 0$, then $(a,b) \in W_{m,2}$, where $m = \frac{1}{2}(2^n - 2) = 2^{n-1} - 1$. $\square$

Proposition (2.6). *$W_{2^{n-1}-1,2}$ is S^1-invariant.*

Proof. Suppose $(a,b) \in \widetilde{\mathbb{A}}_n \times \widetilde{\mathbb{A}}_n$ with $||a|| = ||b|| = 1$ and $b \in \mathbb{H}_a^{\perp}$. From this we have $b \perp a$, and $\widetilde{b} \perp a$ (equivalently $\widetilde{a} \perp b$).

Now $r(a,b)+s(-b,a) = (ra-sb, rb+sa)$ and we know that $(ra-sb) \perp (rb+sa)$. To finish, we need to show that $\widetilde{(ra - sb)} \perp (rb + sa)$. But

$$\begin{aligned} \langle \widetilde{ra - sb}, rb + sa\rangle_n &= \langle r\widetilde{a} - s\widetilde{b}, rb + sa\rangle_n \\ &= r^2\langle \widetilde{a}, b\rangle_n - s^2\langle \widetilde{b}, a\rangle_n + rs\langle \widetilde{a}, a\rangle_n - rs\langle \widetilde{b}, a\rangle_n \\ &= 0, \end{aligned}$$

and therefore $(ra - sb) \in \mathbb{H}_{rb+sa}$. $\square$

Note that we have a fibration

$$\begin{array}{ccccc} S^{2^n-5} & \to & W_{2^{n-1}-1,2} & \xrightarrow{\pi} & S(\widetilde{\mathbb{A}}_n) = S^{2^n-3} \\ & & (a,b) & \mapsto & b \end{array}$$

and $\pi^{-1}(b) = S(\mathbb{H}_b^{\perp}) = S^{2^n-5}$ since dim $\mathbb{H}_b^{\perp} = 2^n - 4$. Thus dim $W_{2^{n-1}-1,2} = 2^n - 5 + 2^n - 3 = 2^{n+1} - 8$.

In [Mo$_1$] it is shown that for a and b in $\mathbb{A}_n$ with $n \geq 4$ and $||a|| = ||b|| = 1$, if $ab = 0$ then (i) $(a,b) \in \widetilde{\mathbb{A}}_n \times \widetilde{\mathbb{A}}_n$ and (ii) $b \in \mathbb{H}_a^{\perp}$ (or equivalently $a \in \mathbb{H}_b^{\perp}$). Thus

$$X_n := \{(a,b) \in \mathbb{A}_n \times \mathbb{A}_n : ||a|| = ||b|| = 1 \quad \text{and} \quad ab = 0\}$$

is a subset of $W_{2^{n-1}-1,2}$. This gives a chain of inclusions for $n \geq 3$,

$$X_n \subset W_{2^{n-1}-1,2} \subset V_{2^n-2,2} \subset V_{2^n-1,2}\,.$$

We show that X_n and $W_{2^{n-1}-1,2}$ admit a $T := S^1 \times S^1$ action.

LEMMA (2.7). *For $(a,b) \in V_{2^n-2,2}$ and r,s,q,p in $\mathbb{R}$ with $r^2+s^2=1$ and $p^2+q^2=1$, define*

$$(a,b) \overset{\tau}{\mapsto} (ra+s\tilde{a}, pb+q\tilde{b})\,.$$

Then

(i) *if $(a,b) \in W_{2^{n-1}-1,2}$ then $(ra+s\tilde{a}, pb+q\tilde{b}) \in W_{2^{n-1}-1,2}$;*
(ii) *if $(a,b) \in X_n$ then $(ra+s\tilde{a}, pb+q\tilde{b}) \in X_n$;*
(iii) *τ defines a free T-action on $W_{2^{n-1}-1,2}$ and X_n respectively.*

Proof. By direct calculations. If $a \perp b$ and $\tilde{a} \perp b$, then

$$\begin{aligned}
\langle ra+s\tilde{a}, pb+q\tilde{b}\rangle_n &= rp\langle a,b\rangle_n + sq\langle \tilde{a},\tilde{b}\rangle_n + rq\langle a,\tilde{b}\rangle_n + sp\langle \tilde{a},b\rangle_n\\
&= 0+0+0+0\\
&= 0.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\langle ra+s\tilde{b}, (pb+q\tilde{b})\tilde{e}_0\rangle_n &= \langle ra+s\tilde{a}, p\tilde{b}-qb\rangle_n\\
&= rp\langle a,\tilde{b}\rangle_n + sp\langle \tilde{a},\tilde{b}\rangle_n - rq\langle a,b\rangle_n - sq\langle \tilde{a},b\rangle_n\\
&= 0+0+0+0\\
&= 0.
\end{aligned}$$

If $ab=0$, then

$$\begin{aligned}
(ra+s\tilde{a})(pb+q\tilde{b}) &= rp(ab) + sq\tilde{a}\tilde{b} + sp\tilde{a}b + rqa\tilde{b}\\
&= 0.
\end{aligned}$$

Also

$$\begin{aligned}
||ra+s\tilde{a}||^2 &= r^2||a||^2 + s^2||\tilde{a}||^2 = (r^2+s^2)||a||^2 = 1 \quad \text{and}\\
||pb+q\tilde{b}||^2 &= p^2||b||^2 + q^2||\tilde{b}||^2 = (p^2+q^2)||b||^2 = 1
\end{aligned}$$

Thus we have proved (i) and (ii). Finally $(ra+s\tilde{a}, pb+q\tilde{b}) = (a,b)$ if and only if $r=1, s=0, p=1$ and $q=0$. Clearly this action is smooth and free (see Corollary (2.3)). ☐

3. X_n, Octonions and an S^3 action

In this section we show that we can attach to every element in X_n a copy of $\mathbb{A}_3 = \mathbb{O}$, the octonions inside of $\mathbb{A}_{n+1}$ for $n > 3$. This allows us to identify X_n with a subset of algebra monomorphisms of $\mathbb{A}_3$ into $\mathbb{A}_{n+1}$, which will be our main goal in §4. We recall some notation from section 1. Let $e_0 \in \mathbb{A}_{n-1}$ be the unit, so $(e_0,0) = e_0$ is the unit in $\mathbb{A}_n$ and $\widetilde{e}_0 = (0,e_0)$ in $\mathbb{A}_n$. For $\widetilde{e}_0$ in $\mathbb{A}_n$ we denote $\varepsilon = (\widetilde{e}_0, 0)$ in $\mathbb{A}_{n+1}$. For example, for $n=4$, $\widetilde{e}_0 = e_8$ in $\mathbb{A}_4$, and $\varepsilon = (e_8, 0)$ in $\mathbb{A}_5$. In general $\widetilde{e}_0 = e_{2^n}$ in $\mathbb{A}_{n+1}$ and $\varepsilon = e_{2^{n-1}}$ in $\mathbb{A}_{n+1}$. Since ε is a doubly pure element of norm one, we have that $\mathbb{H}_\varepsilon = \mathrm{Span}\{e_0, \widetilde{\varepsilon}, \varepsilon, \widetilde{e}_0\} \subset \mathbb{A}_{n+1}$ is a copy of $\mathbb{A}_2$ and a direct sum decomposition $\mathbb{A}_{n+1} = \mathbb{H}_\varepsilon \oplus \mathbb{H}_\varepsilon^\perp$. By definition

$\alpha = (a, b) \in \mathbb{A}_n \times \mathbb{A}_n = \mathbb{A}_{n+1}$ is doubly pure in $\mathbb{A}_{n+1}$ with doubly pure entries in $\mathbb{A}_n$ if and only if $\alpha \in \mathbb{H}_\varepsilon^\perp$.

In section 2 we constructed the chain

$$X_n \subset W_{2^{n-1}-1,2} \subset V_{2^n-2,2} \subset V_{2^n-1,2} \subset \mathbb{A}_{n+1}$$

for $n \geq 3$ with $X_3 = \Phi$, the empty set. Therefore by definition, $V_{2^n-2,2} = V_{2^n-1,2} \cap \mathbb{H}_\varepsilon^\perp$.

Lemma (3.1). *For $\alpha \in \mathbb{H}_\varepsilon^\perp \subset \mathbb{A}_{n+1}$ with $\alpha = (a, b) \in \widetilde{\mathbb{A}}_n \times \widetilde{\mathbb{A}}_n$,*
1) $(\alpha\varepsilon) \in \mathbb{H}_\varepsilon^\perp$ *and* $\alpha\varepsilon = (\widetilde{a}, -\widetilde{b})$.
2) $\alpha\widetilde{\varepsilon} \in \mathbb{H}_\varepsilon^\perp$ *and* $\alpha\widetilde{\varepsilon} = \widetilde{\alpha}\varepsilon = -\widetilde{\alpha\varepsilon} = (-\tilde{b}, -\tilde{a})$.

Proof. By direct calculation,

$$\alpha\varepsilon = (a, b)(\widetilde{e}_0, 0) = (a\tilde{e}_0, -b\tilde{e}_0) = (\tilde{a}, -\tilde{b}) \in \tilde{\mathbb{A}}_n \times \tilde{\mathbb{A}}_n = \mathbb{H}_\varepsilon^\perp$$

and

$$\alpha\tilde{\varepsilon} = (a, b)(0, \widetilde{e}_0) = (\tilde{e}_0 b, \tilde{e}_0 a) = (-\tilde{b}, -\tilde{a})$$

by Lemma (1.3) (1). Finally, using Lemma (1.3) (6) and (3), respectively, we get $\alpha\tilde{\varepsilon} = \tilde{\alpha}\varepsilon = -\widetilde{\alpha\varepsilon} \in \mathbb{H}_\varepsilon^\perp$. □

Corollary (3.2). *For a non-zero α in $\mathbb{H}_\varepsilon^\perp \subset \mathbb{A}_{n+1}$ and $n \geq 3$,*

$$\mathbb{O}_\alpha := Span\{e_0, \tilde{\varepsilon}, \varepsilon, \widetilde{e}_0, \tilde{\alpha}, \alpha\varepsilon, \tilde{\varepsilon}\alpha, \alpha\} \subset \mathbb{A}_{n+1}$$

is an 8-dimensional vector subspace of $\mathbb{A}_{n+1} = \mathbb{R}^{2^{n+1}}$.

Proof. By definition $\{e_0, \tilde{\varepsilon}, \varepsilon, \tilde{e}_0\}, \{e_0, \tilde{\alpha}, \alpha, \tilde{e}_0\}, \{\varepsilon, \alpha, \tilde{\varepsilon}, \tilde{\alpha}\}$ are an orthogonal set of vectors and $\alpha\varepsilon \in \mathbb{H}_\varepsilon^\perp \cap \mathbb{H}_\alpha^\perp$. Also by Lemma (3.1). $\tilde{\varepsilon}\alpha = -\alpha\tilde{\varepsilon} \in \mathbb{H}_\varepsilon^\perp \cap \mathbb{H}_\alpha^\perp$. Thus $\{e_0, \tilde{\varepsilon}, \varepsilon, \tilde{e}_0, \tilde{\alpha}, \alpha\varepsilon, \tilde{\varepsilon}\alpha, \alpha\}$ is an orthogonal set of vectors in $\mathbb{A}_{n+1}$. □

Remark (3.3). In particular, for $\alpha \in V_{2^n-2,2}$, we have that $\mathbb{O}_\alpha \cong \mathbb{R}^8 \subset \mathbb{A}_{n+1}$ and $\mathbb{O}_\alpha \oplus \mathbb{O}_\alpha^\perp = \mathbb{A}_{n+1}$.

Lemma (3.4). *For $\alpha \in V_{2^n-2,2}$, $\alpha \in W_{2^{n-1}-1,2}$ if and only if $\hat{\alpha} \in \mathbb{O}_\alpha^\perp$.*

Proof. Recall that by definition $\hat{\alpha} = (b, a)$ if $\alpha = (a, b)$, so

$$\hat{\alpha} \in (\text{Span}((\{e_0, \tilde{\varepsilon}, \varepsilon, \tilde{e}_0, \alpha, \tilde{\alpha}\}))^\perp$$

(see Lemma (2.1). above). Now

$$\begin{aligned}
\langle\hat{\alpha}, \tilde{\varepsilon}\alpha\rangle_{n+1} &= \langle(b, a), (\tilde{b}, \tilde{a})\rangle_{n+1} = \langle b, \tilde{b}\rangle_n + \langle a, \tilde{a}\rangle_n = 0 \\
\langle\hat{\alpha}, \alpha\varepsilon\rangle_{n+1} &= \langle(b, a), (\tilde{a}, -\tilde{b})\rangle_{n+1} = \langle b, \tilde{a}\rangle_n - \langle a, \tilde{b}\rangle_n = 2\langle b, \tilde{a}\rangle_n,
\end{aligned}$$

so $\hat{\alpha} \perp (\alpha\varepsilon)$ in $\mathbb{A}_{n+1}$ if and only if $\tilde{a} \perp b$ in $\mathbb{A}_n$, i.e. $b \in \mathbb{H}_a^\perp$. Therefore $\hat{\alpha} \in \mathbb{O}_\alpha^\perp$ if and only if $b \in \mathbb{H}_a^\perp$. □

From this we see that $W_{2^{n-1}-1,2} = \{\alpha \in V_{2^n-2,2} | \hat{\alpha} \in \mathbb{O}_\alpha^\perp\}$.

Theorem (3.5). *For $\alpha \in W_{2^{n-1}-1,2}$ and $n \geq 4$, the following statements are equivalent.*
(i) $\alpha \in X_n$;
(ii) α *alternates with* ε *i.e.,* $(\alpha, \alpha, \varepsilon) = 0$;
(iii) *the vector subspace of* $\mathbb{A}_{n+1}$

$$V(\alpha; \varepsilon) := \text{Span}\{e_0, \alpha, \varepsilon, \alpha\varepsilon\}.$$

is multiplicatively closed and isomorphic to $\mathbb{A}_2 = \mathbb{H}$*;*

(iv) $\mathbb{O}_\alpha$ *is multiplicatively closed and isomorphic to* $\mathbb{A}_3 = \mathbb{O}$*;*

(v) $\hat{\alpha} \in \mathrm{Ker} L_\alpha \subset \mathbb{O}_\alpha^\perp$, *where* L_α *is left multiplication by* α.

Proof. First of all we calculate

$$\begin{aligned} \alpha(\alpha\varepsilon) &= (a,b)[(a,b)(\tilde{e},0)] = (a,b)(\tilde{a},-\tilde{b}) = (a\tilde{a} - \tilde{b}b, -\tilde{b}a - b\tilde{a}) \\ &= (-||a||^2\tilde{e}_0 - ||b||^2\tilde{e}_0, -\tilde{b}a - \tilde{b}a) \text{ (by Lemma 1.1 (2) and (5))} \\ &= -||\alpha||^2\varepsilon + 2(0, \tilde{ba}). \end{aligned}$$

Therefore $\alpha(\alpha\varepsilon) = \alpha^2\varepsilon = -||\alpha||^2\varepsilon$ if and only if $ba = 0$; i.e., $\alpha \in X_n$ and we have (i)$\Leftrightarrow$ (ii).

Clearly if $\alpha \in W_{2^{n-1}-1,2}$ then $\{e_0, \alpha, \varepsilon, \alpha\varepsilon\}$ form an orthonormal set of vectors, $\dim_{\mathbb{R}}(V(\alpha;\varepsilon)) = 4$ so $-||\alpha||^2 = \alpha^2 = (\alpha\varepsilon)^2$ and $\alpha(\alpha\varepsilon) = -||\alpha||^2\varepsilon$ if and only if $V(\alpha;\varepsilon) = \mathbb{H}$, and we have proved (ii)$\Leftrightarrow$(iii).

To prove (iii)$\Leftrightarrow$(iv) we establish the following correspondence between the canonical basis in $\mathbb{A}_3$ and the orthonormal basis of $\mathbb{O}_\alpha$.

$$e_1 \to \tilde{\varepsilon}; e_2 \mapsto \varepsilon; e_3 \to \tilde{e}_0; ||\alpha||e_4 \to \tilde{\alpha}; ||\alpha||e_5 \to \alpha\varepsilon; ||\alpha||e_6 \to \tilde{\varepsilon}\alpha; ||\alpha||e_7 \to \alpha$$

Using ii) it is a routine calculation to see that this correspondence defines an algebra isomorphism. (See also Lemma (4.8) (1) below). Finally, from Lemma (3.4) we know that $\hat{\alpha} \in \mathbb{O}_\alpha^\perp$ and

$$\alpha\hat{\alpha} = (a,b)(b,a) = (ab + ab, a^2 - b^2) = (2ab, ||b||^2 - ||a||^2)$$

is the Hopf construction. Thus $\alpha\hat{\alpha} = 0$ in $\mathbb{A}_{n+1}$ if and only if $\alpha \in X_n$. Recall that $\mathbb{A}_3 \cong \mathbb{O}_\alpha$ does not admit zero divisors. □

THEOREM (3.6). $\mathbb{H}_\varepsilon^\perp$ *admits a left* $\mathbb{H}_\varepsilon$*-module structure for* $n \geq 3$.

Proof. For $\alpha = (a,b)$ in $\mathbb{H}_\varepsilon^\perp = \widetilde{\mathbb{A}}_n \times \widetilde{\mathbb{A}}_n$ and $u = re_0 + s\widetilde{\varepsilon} + q\varepsilon + p\widetilde{e}_0$ with r, s, q and p in $\mathbb{R}$. Define

$$u \cdot \alpha = \alpha u = r\alpha + s\alpha\widetilde{\varepsilon} + q\alpha\varepsilon + p\widetilde{\alpha}\,.$$

Trivially $\widetilde{\alpha} \in \mathbb{H}_\varepsilon^\perp$ and $(\alpha\widetilde{\varepsilon})$ and $(\alpha\varepsilon)$ are in $\mathbb{H}_\varepsilon^\perp$ by Lemma (3.1) (2) and (1), respectively. Since $\widetilde{\varepsilon}, \varepsilon$ and $\widetilde{e}_0$ are alternative elements in $\mathbb{A}_{n+1}$ (actually they belong to the canonical basis) we have that $\widetilde{\varepsilon} \cdot \alpha = (\alpha\widetilde{\varepsilon})\widetilde{\varepsilon} = \alpha(\widetilde{\varepsilon})^2 = -||\alpha||^2 e_0 = \widetilde{\varepsilon} \cdot \alpha$ and similarly $\varepsilon \cdot (\varepsilon \cdot \alpha) = \varepsilon^2 \cdot \alpha$ and $\widetilde{e}_0 \cdot (\widetilde{e}_0 \cdot \alpha) = \widetilde{e}_0^2 \cdot \alpha$.

Now $\varepsilon \cdot (\widetilde{e}_0 \cdot \alpha) = \varepsilon \cdot (\alpha\widetilde{e}_0) = \varepsilon \cdot \widetilde{\alpha} = \widetilde{\alpha}\varepsilon$ and $(\varepsilon\widetilde{e}_0) \cdot \alpha = \widetilde{\varepsilon} \cdot \alpha = \alpha\widetilde{\varepsilon} = \widetilde{\alpha}\varepsilon$ by Lemma (3.1) (2). Similarly,

$$\begin{aligned} \widetilde{\varepsilon} \cdot (\widetilde{e}_0 \cdot \alpha) &= (\widetilde{\varepsilon}\widetilde{e}_0) \cdot \alpha = \varepsilon\alpha \\ \widetilde{e}_0 \cdot (\widetilde{\varepsilon} \cdot \alpha) &= (\widetilde{e}_0\widetilde{\varepsilon}) \cdot \alpha = \alpha\varepsilon \\ \varepsilon \cdot (\widetilde{e}_0 \cdot \alpha) &= (\varepsilon\widetilde{e}_0) \cdot \alpha = \widetilde{\alpha}\varepsilon \\ \widetilde{e}_0 \cdot (\varepsilon \cdot \alpha) &= (\widetilde{e}_0\varepsilon) \cdot \alpha = -\alpha\widetilde{\varepsilon}. \end{aligned}$$

Finally $\widetilde{\varepsilon} \cdot (\varepsilon \cdot \alpha) = \widetilde{\varepsilon} \cdot (\alpha\varepsilon) = (\alpha\varepsilon)\widetilde{\varepsilon} = (\widetilde{\alpha\varepsilon})\varepsilon = -\widetilde{(\alpha\varepsilon)}\varepsilon = \widetilde{-\alpha} = (\widetilde{\varepsilon}\varepsilon) \cdot \alpha$ and $\varepsilon \cdot (\widetilde{\varepsilon} \cdot \alpha) = \varepsilon \cdot (\alpha\widetilde{\varepsilon}) = (\alpha\widetilde{\varepsilon})\varepsilon = (\widetilde{\alpha}\varepsilon)\varepsilon = \widetilde{\alpha} = \varepsilon \cdot (\widetilde{\varepsilon}\alpha)$. Apply Lemma (3.1) and Lemma (1.3). and we are done. □

We now define a S^3 action on X_n. Consider the unit sphere in $\mathbb{H}_\varepsilon \subset \mathbb{A}_{n+1}$,

$$S^3 = S(\mathbb{H}_\varepsilon) = \{re_0 + s\tilde{\varepsilon} + q\varepsilon + p\tilde{e}_0 | r^2 + s^2 + q^2 + p^2 = 1\}\,.$$

For $\alpha \in \mathbb{H}_\varepsilon^\perp$ with $\alpha = (a,b) \in \tilde{\mathbb{A}}_n \times \tilde{\mathbb{A}}_n$ define $\mathbb{H}_\varepsilon^\perp \times S^3 \to \mathbb{H}_\varepsilon^\perp$ by

$$\begin{aligned} \alpha(re_0 + s\tilde{\varepsilon} + q\varepsilon + p\tilde{e}_0) &= r\alpha + s\alpha\tilde{\varepsilon} + q\alpha\varepsilon + p\alpha\tilde{e}_0 = r\alpha + s\tilde{\alpha}\varepsilon + q\alpha\varepsilon + p\tilde{\alpha} \\ &= r(a,b) + s(-\tilde{b}, -\tilde{a}) + q(\tilde{a}, -\tilde{b}) + p(-b, a) \\ &= (ra - s\tilde{b} + q\tilde{a} - pb, rb - s\tilde{a} - q\tilde{b} + pa). \end{aligned}$$

By definition this is a group action which is smooth and free of fixed points.

Corollary (3.7). *The above action of $S^3 = S(\mathbb{H}_\varepsilon)$ on $\mathbb{H}_\varepsilon^\perp$ is a group action which is smooth, orthogonal, and free of fixed points.*

Proof. By Theorem (3.6) this is a smooth group action because it is a restriction of a linear action. Since right multiplication by $\widetilde{e}_0, \varepsilon$ and $\widetilde{\varepsilon}$ are orthogonal linear transformations, we have that the action is orthogonal. Finally this action is free of fixed points because $\{e_0, \widetilde{\varepsilon}, \varepsilon, \varepsilon\}$ is an orthonormal basis, so $\alpha(re_0 + s\widetilde{\varepsilon} + q\varepsilon + p\widetilde{e}_0) = \alpha$ if and only if $r = 1, s = q = p = 0$. □

Theorem (3.8). i) *The subsets X_n and $W_{2^{n-1}-1,2}$ of $\mathbb{H}_\varepsilon^\perp$ are S^3-equivariant;*
ii) *for α and β in $W_{2^{n-1}-1,2,}$ $\mathbb{O}_\alpha = \mathbb{O}_\beta$ as vector spaces if and only if α and β lie in the same S^3-orbit.*

Proof. For $\alpha \in \mathbb{H}_\varepsilon^\perp$ with $\alpha = (a,b) \in \tilde{\mathbb{A}}_n \times \tilde{\mathbb{A}}_n$ and r, s, q, p in $\mathbb{R}$ with $r^2 + s^2 + q^2 + p^2 = 1$ we have that

$$\alpha(re_0 + s\tilde{\varepsilon} + q\varepsilon + s\tilde{e}_0) = (ra - s\tilde{b} + q\tilde{a} - pb, rb - s\tilde{a} - q\tilde{b} + pa)\,.$$

Suppose that $\alpha \in W_{2^{n-1}-1,2}$; then $\langle \tilde{a}, b\rangle_n = -\langle a, \tilde{b}\rangle_n = 0; \langle a, b\rangle_n = \langle \tilde{a}, \tilde{b}\rangle_n = 0$ and, by definition, $\langle a, \tilde{a}\rangle_n = \langle b, \tilde{b}\rangle_n = 0$ with $||a|| = ||\tilde{a}|| = ||\tilde{b}|| = ||b|| = 1$, so $\langle ra - s\tilde{b} + q\tilde{a} - pb, rb - s\tilde{a} - q\tilde{b} + pa\rangle = p||a||^2 + sq||b||^2 - qs||\tilde{a}||^2 - pr||b||^2 = 0$ and $(\alpha(re_0 + s\tilde{\varepsilon} + q\varepsilon + s\tilde{e}_0)) \in V_{2^n-2,2}$.

Similarly, $\langle ra - s\tilde{b} + q\tilde{a} - pb, r\tilde{b} + sa + qb + p\tilde{a}\rangle = 0$ and

$$\alpha(re_0 + s\tilde{\varepsilon} + q\varepsilon + s\tilde{e}_0) \in W_{2^{n-1}-1,2}\,.$$

A direct calculation shows that if $ab = 0$ then

$$ab = \tilde{a}b = a\tilde{b} = \tilde{a}\tilde{b} = 0$$

$$(ra - s\tilde{b} + q\tilde{a} - pb)(rb - s\tilde{a} - q\tilde{b} + pa) =$$

$$-rsa\tilde{a} + rpa^2 - sr\tilde{b}b + sq\tilde{b}^2 - qs\tilde{a}^2 + qpaa - prb^2 + pqb\tilde{b} =$$

$$-rs(-||a||^2\tilde{e}_0 + ||b||^2\tilde{e}_0) + pq(-||a||^2\tilde{e}_0 + ||b||^2\tilde{e}_0) + rp(a^2 - b^2) + sq(\tilde{b}^2 - \tilde{a}^2) = 0$$

since $||a||^2 = ||b||^2 = 1$ and $a^2 = b^2 = -e_0$, so we have (i).

To prove (ii) we note that $\mathbb{O}_\alpha = \mathbb{H}_\varepsilon \oplus \operatorname{Span}\{\tilde{\alpha}, \alpha\varepsilon, \tilde{\varepsilon}\alpha, \alpha\}$. Thus, if $\beta = r\alpha + s\tilde{\alpha}\varepsilon + q\alpha\varepsilon + p\tilde{\alpha}$ (recall that $\tilde{\alpha}\varepsilon = -\tilde{\varepsilon}\alpha$ by Lemma (3.1)) and $||\beta|| = 1$, then $r^2 + s^2 + q^2 + p^2 = 1$ and $\alpha \equiv \beta \bmod S^3$ if and only if $\mathbb{O}_\beta \subset \mathbb{O}_\alpha$, but

$$\dim \mathbb{O}_\beta = \dim \mathbb{O}_\alpha = 8 \text{ and } \mathbb{O}_\beta = \mathbb{O}_\alpha\,.$$

□

Remark (3.9). Note that $T = S^1 \times S^1$ as in Lemma (2.7) and $S^3 = S(\mathbb{H}_\varepsilon)$ intersect in a copy of S^1. Suppose that $r^2 + s^2 + p^2 + q^2 = 1$, $u^2 + v^2 = 1$ and $t^2 + m^2 = 1$ in $\mathbb{R}$. If $(ra - s\tilde{b} + q\tilde{a} - pb, rb - s\tilde{a} - q\tilde{b} + pa) = (ua + v\tilde{a}, tb + m\tilde{b})$ then $r = u$, $q = v, s = 0, p = 0, r = t, -q = m$, so

$$S(\mathbb{H}_\varepsilon) \cap T = S^1 = \{(r, -q) | r^2 + q^2 = 1\}\,.$$

4. X_n and monomorphisms from $\mathbb{A}_3$ to $\mathbb{A}_{n+1}$

In this section we will assume that $1 \leq m \leq n$.

Definition (4.1). An algebra monomorphism from $\mathbb{A}_m$ to $\mathbb{A}_n$ is a linear monomorphism $\varphi : \mathbb{A}_m \to \mathbb{A}_n$ such that

i) $\varphi(e_0) = e_0$ (the first e_0 is in $\mathbb{A}_m$ and the second e_0 in $\mathbb{A}_n$);

ii) $\varphi(xy) = \varphi(x)\varphi(y)$ for all x and y in $\mathbb{A}_m$.

By definition we have that $\varphi(re_0) = r\varphi(e_0)$ for all r in $\mathbb{R}$, so $\varphi({}_0\mathbb{A}_m) \subset \varphi({}_0\mathbb{A}_n)$ and $\varphi(\overline{x}) = \overline{\varphi(x)}$. Therefore $||\varphi(x)||^2 = \varphi(x)\overline{\varphi(x)} = \varphi(x)\varphi(\overline{x}) = \varphi(x\overline{x}) = \varphi(||x||^2) = ||x||^2$ for all $x \in \mathbb{A}_m$, hence $||\varphi(x)|| = ||x||$ and φ is an orthogonal linear transformation from $\mathbb{R}^{2^m-1}$ to $\mathbb{R}^{2^n-1}$.

The *trivial* monomorphism is the one given by $\varphi(x) = (x, 0, 0, \ldots, 0)$ for $x \in \mathbb{A}_m$ and 0 in $\mathbb{A}_m$. $\mathcal{M}(\mathbb{A}_m; \mathbb{A}_n)$ denotes the set of algebra monomorphisms from $\mathbb{A}_m$ to $\mathbb{A}_n$. For $m = n$, $\mathcal{M}(\mathbb{A}_m; \mathbb{A}_n) = \mathrm{Aut}(\mathbb{A}_n)$ is the group of algebra automorphisms of $\mathbb{A}_n$.

PROPOSITION (4.2). $\mathcal{M}(\mathbb{A}_1; \mathbb{A}_n) = S({}_o\mathbb{A}_n) = S^{2^n-2}$.

Proof. $\mathbb{A}_1 = \mathbb{C} = \mathrm{Span}\{e_0, e_1\}$. If $x \in \mathbb{A}_1$, then $x = re_0 + se_1$ and for $w \in_o \mathbb{A}_n$ with $||w|| = 1$ we have that $\varphi_w(x) = re_0 + sw$ defines an algebra monomorphism from $\mathbb{A}_1$ to $\mathbb{A}_n$. This can be seen by direct calculation, recalling that $\mathrm{Center}(\mathbb{A}_n) = \mathbb{R}$ for all n and that every associator with one real entry vanishes.

Conversely, for $\varphi \in \mathcal{M}(\mathbb{A}_1; \mathbb{A}_n)$ set $w = \varphi(e_1)$, so $||w|| = 1$ and $\varphi_w = \varphi$. □

Remark (4.3). In particular, we have that

$$\mathrm{Aut}(\mathbb{A}_1) = S^0 = \mathbb{Z}/2 = \{\text{Identity, conjugation}\} = \{\varphi_{e_1}, \varphi_{-e_1}\}\,.$$

To calculate $\mathcal{M}(\mathbb{A}_2; \mathbb{A}_n)$ for $n \geq 2$ we need to recall the following (see [Mo$_2$]).

Definition (4.4). For a and b in $\mathbb{A}_n$. We say that *a alternates with b* and write $a \rightsquigarrow b$, if $(a, a, b) = 0$. We say that *a alternates strongly with b* and write $a \leftrightsquigarrow b$, if $(a, a, b) = 0$ and $(a, b, b) = 0$.

Clearly a alternates strongly with e_0 for all a in $\mathbb{A}_n$, and if a and b are linearly dependent then $a \leftrightsquigarrow b$ (by flexibility). Also, by definition a is an alternative element if and only if $a \rightsquigarrow x$ for all x in $\mathbb{A}_n$.

By Lemma (1.3) (1) and (2) we have that for any doubly pure element a in $\mathbb{A}_n$ $(a, a, \tilde{e}_0) = 0$ and (by the above remarks) $\tilde{e}_0$ alternate strongly with any a in $\mathbb{A}_n$.

For a and b pure elements in $\mathbb{A}_n$, we define the vector subspace of $\mathbb{A}_n$

$$V(a; b) = \mathrm{Span}\{e_0, a, b, ab\}\,.$$

LEMMA (4.5). *If* $(a, b) \in V_{2^n-1,2}$ *and* $a \leftrightsquigarrow b$, *then* $V(a; b) = \mathbb{A}_2 = \mathbb{H}$.

Proof. Suppose that $(a,b)\in V_{2^n-1,2}$ and that $(a,a,b)=0$. Then we have

$$\begin{aligned}\langle ab,a\rangle &= \langle b,\overline{a}a\rangle=\langle b,||a||^2e_0\rangle=||a||^2\langle b,e_0\rangle=0\\ \langle ab,a\rangle &= \langle a,b\overline{b}\rangle=\langle a,||b||^2e_0\rangle=||b||^2\langle a,e_0\rangle=0\\ ||ab||^2=\langle ab,ab\rangle &= \langle\overline{a}(ab),b\rangle=\langle -a(ab),b\rangle=\langle -a^2b,b\rangle\\ &= -a^2\langle b,b\rangle=||a||^2||b||^2=1,\end{aligned}$$

so $\{e_0,a,b,ab\}$ is an orthonormal set of vectors in $\mathbb{A}_n$.

Finally, using also that $(a,b,b)=0$ and $ab=-ba$, we may check by direct calculation that the multiplication table of $\{e_0,a,b,ab\}$ coincides with that of the quaternions, and by the identification $e_0\mapsto e_0, a\mapsto e_1, b\mapsto e_2$ and $ab\mapsto e_3$ we have an algebra isomorphism between $\mathbb{A}_2=\mathbb{H}$ and $V(a;b)$. ☐

Proposition (4.6). *$\mathcal{M}(\mathbb{A}_2;\mathbb{A}_n)=\{(a,b)\in V_{2^n-1,2}|a\leftrightsquigarrow b\}$ for $n\geq 2$. In particular,*

$$\mathrm{Aut}(\mathbb{A}_2)=\mathcal{M}(\mathbb{A}_2;\mathbb{A}_2)=V_{3,2}=SO(3)$$

and

$$\mathcal{M}(\mathbb{A}_2,\mathbb{A}_3)=V_{7,2}.$$

Proof. The inclusion "$\supset$" follows from Lemma (4.5). Conversely, suppose that $\varphi\in\mathcal{M}(\mathbb{A}_2,\mathbb{A}_n)$; then $\varphi(e_0)=e_0, (\varphi(e_1),\varphi(e_2))\in V_{2^n-1,2}$ and $V(\varphi(e_1),\varphi(e_2))=\mathrm{Im}\varphi=\mathbb{H}\subset\mathbb{A}_n$.

Since $\mathbb{A}_2$ is an asociative algebra and $\mathbb{A}_3$ is an alternative algebra we have that $a\leftrightsquigarrow b$ for any two elements in $\mathbb{A}_n$ when $n=2$ or $n=3$. ☐

Definition (4.7). Recall that $\tilde{\mathbb{A}}_n=\{e_0,\tilde{e}_0\}^{\perp}=\mathbb{R}^{2^n-2}$ denotes the vector subspace of doubly pure elements. Since $a\leftrightsquigarrow\tilde{e}_0$ for any element in $\tilde{\mathbb{A}}_n$, we have that, if $a\in S(\tilde{\mathbb{A}}_n)$ i.e., $||a||=1$ then $(a,\tilde{e}_0)\in V_{2^n-1,2}$ and the assignment $a\mapsto(a,\tilde{e}_0)$ defines an inclusion from $S(\tilde{\mathbb{A}}_n)=S^{2^n-3}\hookrightarrow\mathcal{M}(\mathbb{A}_2;\mathbb{A}_n)\subset V_{2^n-1,2}$ which resembles the *bottom cell* inclusion in $V_{2^n-1,2}$.

Now we show that X_n can be identified with a subset of $\mathcal{M}(\mathbb{A}_3;\mathbb{A}_{n+1})$ for $n\geq 4$.

Lemma (4.8). *For $\alpha\in\mathbb{H}_{\varepsilon}^{\perp}\subset\mathbb{A}_{n+1}$ and $n\geq 4$,*

(1) *if $||\alpha||=1$ then $\mathbb{O}_\alpha=\mathrm{Span}\{e_0,\widetilde{\varepsilon},\varepsilon,\widetilde{e}_0,\widetilde{\alpha},\alpha\varepsilon,\widetilde{\varepsilon}\alpha,\alpha\}$ is isomorphic as an algebra to $\mathbb{A}_3$ if and only if $(\alpha,\alpha,\varepsilon)=0$;*

(2) *if $\alpha=(a,b)\in\widetilde{\mathbb{A}}_n\times\widetilde{\mathbb{A}}_n$ then $(\alpha,\alpha,\varepsilon)=(0,-(a,\widetilde{e}_0,b))\in\widetilde{\mathbb{A}}_n\times\widetilde{\mathbb{A}}_n$.*

Proof. (1) By definition $\widetilde{e}_0=e_{2^n}$ and $\varepsilon=e_{2^{n-1}}$ are elements in the canonical basis, so they are alternative elements (See [Sch]). Since $\mathbb{H}_\alpha$ is associative for all α we have $(\alpha,\alpha,\widetilde{e}_0)=0$. Clearly if $\mathbb{O}_\alpha\cong\mathbb{A}_3$ then $(\alpha,\alpha,\varepsilon)=0$ because $\mathbb{O}_\alpha$ is an alternative algebra. Conversely, assume that $(\alpha,\alpha,\varepsilon)=0$. We have the following multiplication table which under the mapping

$e_0\mapsto e_0; e_1\mapsto\widetilde{\varepsilon}; e_2\mapsto\varepsilon; e_3\mapsto\widetilde{e}_0; e_4\mapsto\widetilde{\alpha}; e_5\mapsto\alpha\varepsilon; e_0\mapsto\widetilde{\varepsilon}\alpha$ and $e_7\mapsto\alpha$

becomes an algebra monomorphism from $\mathbb{A}_3$ into $\mathbb{A}_{n+1}$:

e_0	$\widetilde{\varepsilon}$	ε	$\widetilde{e}_0$	$\widetilde{\alpha}$	$\alpha\varepsilon$	$\widetilde{\varepsilon}\alpha$	α
$\widetilde{\varepsilon}$	$-e_0$	$\widetilde{e}_0$	$-\varepsilon$	$\alpha\varepsilon$	$-\widetilde{\alpha}$	$-\alpha$	$\widetilde{\varepsilon}\alpha$
ε	$-\widetilde{e}_0$	$-e_0$	$\widetilde{\varepsilon}$	$\widetilde{\varepsilon}\alpha$	α	$-\widetilde{\alpha}$	$-\alpha\varepsilon$
$\widetilde{e}_0$	ε	$-\widetilde{\varepsilon}$	$-e_0$	α	$-\widetilde{\varepsilon}\alpha$	$+\alpha\varepsilon$	$-\widetilde{\alpha}$
$\widetilde{\alpha}$	$-\alpha\varepsilon$	$-\widetilde{\varepsilon}\alpha$	$-\alpha$	$-e_0$	$\widetilde{\varepsilon}$	ε	$-\widetilde{e}_0$
$\alpha\varepsilon$	$\widetilde{\alpha}$	$-\alpha$	$\widetilde{\varepsilon}\alpha$	$-\widetilde{\varepsilon}$	$-e_0$	$-\widetilde{e}_0$	ε
$\widetilde{\varepsilon}\alpha$	α	$\widetilde{\alpha}$	$-\alpha\varepsilon$	$-\varepsilon$	$\widetilde{e}_0$	$-e_0$	$-\widetilde{\varepsilon}$
α	$-\widetilde{\varepsilon}\alpha$	$\alpha\varepsilon$	$\widetilde{\alpha}$	$\widetilde{e}_0$	$-\varepsilon$	$\widetilde{\varepsilon}$	$-e_0$

Note that this table is skew-symmetric with $-e_0$'s along the diagonal. The nontrivial calculations are the following,

$\widetilde{\varepsilon}\widetilde{\alpha} = \alpha\varepsilon$ (by Lemma (1.3) (5)).

$\widetilde{\varepsilon}(\alpha\varepsilon) = -\widetilde{\varepsilon\alpha)} = \widetilde{\varepsilon(\varepsilon\alpha)} = \widetilde{\varepsilon^2\alpha} = -\widetilde{e_0\alpha} = -\widetilde{\alpha}$ (by Lemma (1.3) (3)).

$\varepsilon(\widetilde{\varepsilon}\alpha) = -\varepsilon\widetilde{(\varepsilon\alpha)} = -\widetilde{\varepsilon}(\varepsilon\alpha) = \widetilde{\varepsilon(\varepsilon\alpha)} = -\widetilde{\alpha}$ (by Lemma (1.3) (1) and (6).

$\widetilde{\alpha}(\alpha\varepsilon) = -\widetilde{\alpha(\alpha\varepsilon)} = -\widetilde{\alpha^2\varepsilon} = \widetilde{\varepsilon}$ because $(\alpha, \alpha, \varepsilon) = 0$ and $||\alpha|| = 1$.

$\widetilde{\alpha}(\widetilde{\varepsilon}\alpha) = -\widetilde{\alpha}\widetilde{(\varepsilon\alpha)} = +\widetilde{\widetilde{\alpha}}(\alpha\varepsilon) = -\alpha(\alpha\varepsilon) = \varepsilon$,

so we are done with (1). To prove (2) we perform similar calculation as in Theorem (3.5),

$$\begin{aligned} \alpha(\alpha\varepsilon) &= (a,b)[(a,b)(\widetilde{e}_0, 0)] = (a,b)(\widetilde{a}, -\widetilde{b}) = (a\widetilde{a} - \widetilde{b}b, -\widetilde{b}a - b\widetilde{a}) \\ &= (-||a||^2\widetilde{e}_0 - ||b||^2 e_0, -(b\widetilde{e}_0)a + b(\widetilde{e}_0 a)) \\ &= -||\alpha||^2\varepsilon - (0, (b, \widetilde{e}_0, a)). \end{aligned}$$

Therefore

$$\begin{aligned} (\alpha, \alpha, \varepsilon) &= \alpha^2\varepsilon - \alpha(\alpha\varepsilon) = -||\alpha||^2 - \alpha(\alpha\varepsilon) \\ &= (0, (b, \widetilde{e}_0, a)) = -(0, (a, \widetilde{e}_0, b)) \end{aligned}$$

by flexibility. □

Notation (4.9). For $n \geq 4$ consider the following subsets of $\mathbb{A}_{n+1}$,

$$\begin{aligned} \mathbb{E}_n &= \{\alpha \in \mathbb{H}_\varepsilon^\perp | (\alpha, \alpha, \varepsilon) = 0\}, \\ S(\mathbb{E}_n) &= \{\alpha \in \mathbb{E}_n : ||\alpha|| = 1\}, \\ P(n) &= \{(a,b) \in \widetilde{\mathbb{A}}_n \times \widetilde{\mathbb{A}}_n | a \,\text{and}\, b \,\text{are}\, \mathbb{C} - \text{collinear}\}, \\ \overline{X}_n &= \{(a,b) \in \widetilde{\mathbb{A}}_n \times \widetilde{\mathbb{A}}_n | a \neq 0, b \neq 0 \,\text{and}\, ab = 0\,\}, \end{aligned}$$

and also the following subset of monomorphisms,

$$\mathcal{M}_2(\mathbb{A}_3; \mathbb{A}_{n+1}) = \{\varphi \in \mathcal{M}(\mathbb{A}_3; \mathbb{A}_{n+1}) | \mathbb{H}_\varepsilon \subset \mathrm{Im}\varphi\}.$$

Remark (4.10). By Lemma (4.8) (1) we may identify $S(\mathbb{E}_n)$ and $\mathcal{M}_2(\mathbb{A}_3; \mathbb{A}_{n+1})$; that is, there is a one to one correspondence between these two sets.

THEOREM (4.11). *For $n \geq 4$,*
(i) *$P(n)$ and $\overline{X}_n$ are subsets of $\mathbb{E}_n$ with $P(n) \cap \overline{X}_n = \Phi$, and*
(ii) *There is a continuous retraction $R : \mathbb{E}_n \backslash P(n) \to \overline{X}_n$.*

Proof. If $(a,b) \in P(n)$, then $b \in \mathbb{H}_a$ or $a \in \mathbb{H}_b$ and $(a,\widetilde{e}_0,b)=0$ (recall that $\mathbb{H}_a$ and $\mathbb{H}_b$ are associative), so by Lemma (4.8) (2), $(\alpha,\alpha,\varepsilon)=0$ in $\mathbb{A}_{n+1}$ when $\alpha=(a,b)$ so $P(n)\subset \mathbb{E}_n$.

On the other hand, if $(a,b)\in \overline{X}_n$, then $ab=0$ and $b\in \mathbb{H}_a^{\perp}\subset \mathbb{A}_n$, so by Lemma (1.3) (1), (6) and (3),

$$(a,\widetilde{e}_0,b)=(a\widetilde{e}_0)b-a(\widetilde{e}_0 b)=\widetilde{a}b+a\widetilde{b}=\widetilde{a}b+\widetilde{a}b=2\widetilde{a}b=-2\widetilde{ab}.$$

Therefore if $ab=0$ then $(a,\widetilde{e}_0,b)=0$ and $(\alpha,\alpha,\varepsilon)=0$ in $\mathbb{A}_{n+1}$ for $\alpha=(a,b)$ by Lemma (4.8) (2), so $\overline{X}_n\subset \mathbb{E}_n$.

Now if $(a,b)\in P(n)\cap \overline{X}_n$ then $b\in \mathbb{H}_a$ and $ab=0$, but $\mathbb{H}_a$ is associative and $a=0$ or $b=0$, which is a contradiction. Therefore $P(n)\cap\overline{X}_n$ is the empty set and we are done with (i).

To prove (ii) suppose that $\alpha=(a,b)\in \widetilde{\mathbb{A}}_n\times\widetilde{\mathbb{A}}_n$ with $a\neq 0$. Since

$$\mathbb{A}_n=\mathbb{H}_a\oplus\mathbb{H}_a^{\perp},$$

there are unique elements c and d in $\mathbb{H}_a$ and $\mathbb{H}_a^{\perp}$, respectively, such that $b=c+d$. Now

$$(a,\widetilde{e}_0,b)=(a,\widetilde{e}_0,c+d)=(a,\widetilde{e}_0,c)+(a,\widetilde{e}_0,d)=0+(a,\widetilde{e}_0,d)$$

since $\mathbb{H}_a$ is associative. But by Lemma (1.3) (1), (6) and (3),

$$(a,\widetilde{e}_0,d)=(a\widetilde{e}_0)d-a(\widetilde{e}_0 d)=\widetilde{a}d+a\widetilde{d}=\widetilde{ad}+\widetilde{a}d=-2\widetilde{ad}\,.$$

Therefore $(a,\widetilde{e}_0,b)=-2\widetilde{ad}$. Suppose that $\alpha=(a,b)$ is in $\mathbb{E}_n\backslash P(n)$; then $a\neq 0$, $b\neq 0$,

$$b=c+d\in\mathbb{H}_a\oplus\mathbb{H}_a^{\perp}$$

with $d\neq 0$ and $(a,\widetilde{e}_0,b)=0$ by Lemma (1.3) (2).Thus we have that $ad=0$.

Let us define $R:\mathbb{E}_n\backslash P(n)\to\overline{X}_n$ by $R(a,b)=(a,d)$. Then $R(a,b)=(a,b)$ if $(a,b)\in\overline{X}_n$, and R is continuous because it is the restriction of the projection map

$$\begin{array}{ccccc}\widetilde{\mathbb{A}}_n\times\mathbb{A}_n & \to & \widetilde{\mathbb{A}}_n\times(\mathbb{H}_a\oplus\mathbb{H}_a^{\perp}) & \to & \widetilde{\mathbb{A}}_n\times\mathbb{H}_a^{\perp}\\ (a,b) & \to & (a,c+d) & \to & (a,d),\end{array}$$

which is obviously continuous. □

Remarks (4.12). (1) Recall that $\widetilde{\mathbb{A}}_n$ is a complex vector space by making $ia=\widetilde{a}$. By definition $(a,b)\in P(n)$ if and only if a and b are $\mathbb{C}$-collinear for $(a,b)\in\widetilde{\mathbb{A}}_n\times\widetilde{\mathbb{A}}_n$. Therefore

$$P(n)\cong(((\widetilde{\mathbb{A}}_n\backslash\{0\})\times\mathbb{C})\cup\widetilde{\mathbb{A}}_n).$$

(2) Consider the map $w_n:\widetilde{\mathbb{A}}_n\times\widetilde{\mathbb{A}}_n\to\widetilde{\mathbb{A}}_n$ given by

$$w_n(a,b)=(a,\widetilde{e}_0,b)\,.$$

Since every associator is a pure element $(a,\widetilde{e}_0,b)\perp\widetilde{e}_0$ because $(a,-,b)$ is a skew-symmetric linear transformation (see [Mo$_1$]), we have $(a,\widetilde{e}_0,b)\in\widetilde{\mathbb{A}}_n$.

Now w_n is a polynomial, in fact a quadratic map, and $\mathbb{E}_n=w_n^{-1}(0)$ is a real algebraic set in $\widetilde{\mathbb{A}}_n\times\widetilde{\mathbb{A}}_n=\mathbb{H}_{\varepsilon}^{\perp}=\mathbb{R}^{2^{n+1}-4}$ with 0 in $\widetilde{\mathbb{A}}_n$ a *singular* value, by Lemma (4.8) (2).

(3) X_n is a contraction of $\overline{X_n}$ via normalization on each coordinate.

Received August 08, 2003

Final version received April 21, 2004

DEPARTAMENTO DE MATEMÁTICAS
CINVESTAV DEL IPN
APARTADO POSTAL 14-740
07000 MÉXICO, D.F.
MÉXICO
gmoreno@math.cinvestav.mx

REFERENCES

[A] J. ADEM, *Construction of some normed maps*, Bol. Soc. Mat. Mexicana (2) **20** (1975), 59–79.

[C1] F. COHEN, *A course in some aspects of classical homotopy theory*, Lecture Notes in Math. **1139**, 1989, Springer-Verlag, Berlin, 11–120, .

[C2] F. COHEN, *On Whitehead Squares*, Cayley-Dickson algebras and rational functions. Bol. Soc. Mat. Mexicana. (2) **37** (1992), 55–62.

[D] L.E. DICKSON, *On quaternions and their generalization and the history of the 8 square theorem*, Ann. Math. **20** (1919), 155–171.

[E-K] EAKIN-SATHAYE, *On automorphisms and derivations of Cayley-Dickson algebras*, J. Pure Appl. Algebra **129** (1990), 263–280.

[L] K. LAM, *Construction of Non-singular bilinear maps*, Topology **6** (1967), 423–426.

[H-Y] S.H. KHALIL AND P. YIU, *The Cayley-Dickson algebras: A theorem of Hurwitz and quaternions*, Bol. Soc. Sci. Lett. Lódz Vol. XLVIII 117–169. 1997.

[Mo1] G. MORENO, *The zero divisors of the Cayley-Dickson algebras over the real numbers*, Bol. Soc. Mat. Mexicana (3) **4** (1998), 13–27.

[Mo2] G. MORENO, *Alternative elements in the Cayley-Dickson algebras*, hopf.math.purdue.edu/pub/Moreno.2001.

[Sch] R.D. SCHAFER, *On the algebras formed by the Cayley-Dickson process*, Amer. J. Math. **76** (1954), 435–445.

[Wh] G. WHITEHEAD, Elements of Homotopy theory, Grad. Texts Math. **61** Springer-Verlag, New York-Berlin, 1978.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

THE COMPLEX OF END REDUCTIONS OF A CONTRACTIBLE OPEN 3-MANIFOLD: CONSTRUCTING 1-DIMENSIONAL EXAMPLES

ROBERT MYERS

Dedicated to Fico González-Acuña in honor of his 60th birthday

Abstract. Given an irreducible, contractible, open 3-manifold W which is not homeomorphic to $\mathbb{R}^3$, there is an associated simplicial complex $\mathcal{S}(W)$, the complex of end reductions of W. Whenever W covers a 3-manifold M one has that $\pi_1(M)$ is isomorphic to a subgroup of the group $\mathrm{Aut}(\mathcal{S}(W))$ of simplicial automorphisms of $\mathcal{S}(W)$.

In this paper we give a new method for constructing examples W with $\mathcal{S}(W)$ isomorphic to a triangulation of $\mathbb{R}$. It follows that any 3-manifold M covered by W must have $\pi_1(M)$ infinite cyclic. We also give a complete isotopy classification of the end reductions of these W.

1. Introduction

A *Whitehead manifold* W is an irreducible, contractible, open 3-manifold which is not homeomorphic to $\mathbb{R}^3$. Given a compact, connected 3-manifold J in W which is not contained in a 3-ball in W Brin and Thickstun [1] defined a certain open submanifold V of W called an end reduction of W at J. End reductions are rather nicely behaved but badly embedded manifolds which have certain interesting engulfing and homotopy theoretic properties and are unique up to isotopy with respect to these properties.

In [12] the author showed how to associate to the set of isotopy classes of end reductions of W a certain abstract simplicial complex $\mathcal{S}(W)$ with the following properties. Every self-homeomorphism of W induces an automorphism of $\mathcal{S}(W)$. Whenever W is a non-trivial covering space of a 3-manifold M each non-trivial element of the group $\pi_1(M)$ of covering translations acts without fixed points on $\mathcal{S}(W)$. Thus information about $\mathcal{S}(W)$ gives information about what 3-manifolds W can cover.

This complex seems particularly useful when W is $\mathbb{R}^2$-irreducible, i.e. when W contains no "non-trivial" planes. In [12] the author considered an uncountable collection of $\mathbb{R}^2$-irreducible Whitehead manifolds which are modifications of an example due to Scott and Tucker [13]. He showed that each of these manifolds has $\mathcal{S}(W)$ isomorphic to a triangulation of the real line. It follows that each 3-manifold which is non-trivially covered by one of these 3-manifolds must have

2000 *Mathematics Subject Classification*: Primary: 57M10; Secondary: 57N10.
Keywords and phrases: 3-manifold, end reduction, covering space.
This research was partially supported by NSF Grant DMS-0072429.

infinite cyclic fundamental group, and in fact there are uncountably many which do cover such manifolds.

These "modified Scott-Tucker manifolds" are easy to describe, but the proof that their complexes of end reductions have the stated form is rather lengthy. In the present paper we give a different method for constructing examples of $\mathbb{R}^2$-irreducible Whitehead manifolds W which cover 3-manifolds M with $\pi_1(M) \cong \mathbf{Z}$ and have $\mathcal{S}(W)$ a triangulation of $\mathbb{R}$. This method has the advantage that the proof is much shorter. In addition we are able to classify all the end reductions of these examples. For the modified Scott-Tucker manifolds we were able to classify only those which are $\mathbb{R}^2$-irreducible (which is sufficient to determine the complex). This gives the first $\mathbb{R}^2$-irreducible Whitehead manifolds (other than those of genus one) for which the entire set of end reductions is known.

The methods of this paper can also be used to construct $\mathbb{R}^2$-irreducible Whitehead manifolds which cover 3-manifolds with non-Abelian free fundamental groups and can cover only 3-manifolds with free fundamental groups. This will be the subject of a later paper.

The paper is organized as follows. In section 2 we give general background information and terminology. In section 3 we state those properties of end reductions we will need. In section 4 we prove the existence of graphs in the 3-ball having certain properties that we will need in our construction. In section 5 we prove the main technical result needed to determine the end reductions of our examples. It is a condition on the embedding of one handlebody in the interior of another which ensures that any knot in the smaller handlebody which meets sufficiently many compressing disks for the boundary of the smaller handlebody in an essential way must meet all the compressing disks for the boundary of the larger handlebody in an essential way. This result may be of some independent interest. In section 6 we give our basic construction of the examples W. In section 7 we prove some of their important properties. In section 8 we determine $\mathcal{S}(W)$. In section 9 we show how to modify the construction to get uncountably many such W.

2. Background

In general we follow [5] or [6] for basic 3-manifold terminology. One slight difference is our use of the term ∂-incompressible. This is usually reserved for surfaces F which are properly embedded in a 3-manifold M. We extend this to the case where F is a compact surface in ∂M as follows. F is *∂-incompressible* if whenever Δ is a properly embedded disk in M with $\Delta \cap F$ an arc α and $\Delta \cap \overline{(\partial M - F)}$ an arc β, then α must be ∂-parallel in F.

When X is a submanifold of Y we denote the topological interior of X by Int X and the manifold interior of X by int X. The *exterior* of X is the closure of the complement of a regular neighborhood of X in Y. This term is also applied to the case of a graph Γ in Y. The regular neighborhood is denoted $N(\Gamma, Y)$. A *meridian* of an edge γ of Γ is the boundary of a properly embedded disk in $N(\Gamma, Y)$ which meets γ transversely in a single point.

A sequence $\{C_n\}_{n\geq 0}$ of compact, connected 3-manifolds C_n in a Whitehead manifold W such that $C_n \subseteq \text{int } C_{n+1}$ and $W - \text{int } C_n$ has no compact components is called a *quasi-exhaustion in* W. If $\cup_{n\geq 0} C_n = W$, then it is called an *exhaustion for* W.

The *genus of* $\{C_n\}_{n\geq 0}$ is the maximum of the genera of ∂C_n or ∞ if these genera are unbounded. The *genus of* W is the minimum of the genera of its exhaustions.

A plane Π in W is *proper* if for each compact $K \subseteq W$ one has that $K \cap \Pi$ is compact. A proper plane Π is *trivial* if some component of $W - \Pi$ has closure homeomorphic to $\mathbb{R}^2 \times [0, \infty)$. W is $\mathbb{R}^2$*-irreducible* if every proper plane in W is trivial. Every genus one Whitehead manifold is $\mathbb{R}^2$-irreducible [9].

A compact 3-manifold Y is *weakly anannular* if every properly embedded incompressible annulus in Y has its boundary in a single component of ∂Y.

Lemma (2.1). *Suppose that for each compact $K \subseteq W$ there is a quasi-exhaustion for W such that*

(1) *each C_n is irreducible,*

(2) *each ∂C_n is incompressible in $W - \text{int } C_n$,*

(3) *each $C_{n+1} - \text{int } C_n$ is irreducible, ∂-irreducible, and weakly anannular, and*

(4) $K \subseteq C_1$.

Then W is $\mathbb{R}^2$-irreducible.

Proof. This is Lemma 10.3 of [12], which derives from Lemma 4.2 of Scott and Tucker [13]. □

3. End reductions

In this section we collect some information about end reductions and define the complex of end reductions $\mathcal{S}(W)$ of a Whitehead manifold W.

A compact, connected 3-manifold J in W is *regular in* W if $W-J$ is irreducible and has no component with compact closure. Since W is irreducible the first condition is equivalent to the statement that J does not lie in a 3-ball in W. A quasi-exhaustion $\{C_n\}_{n\geq 0}$ in W is *regular* if each C_n is regular in W.

Let J be a regular 3-manifold in W, and let V be an open subset of W which contains J. We say that V is *end irreducible rel J in W* if there is a regular quasi-exhaustion $\{C_n\}_{n\geq 0}$ in W such that $V = \cup_{n\geq 0} C_n$, $J = C_0$, and ∂C_n is incompressible in $W - \text{int } J$ for all $n \geq 0$. We say that V has the *engulfing property rel J in W* if whenever N is regular in W, $J \subseteq \text{int } N$, and ∂N is incompressible in $W - J$, then V is ambient isotopic rel J to V' such that $N \subseteq V'$. V is an *end reduction of W at J* if V is end irreducible rel J in W, V has the engulfing property rel J in W, and no component of $W - V$ has compact closure.

Theorem (3.1) (Brin-Thickstun). *Given a regular 3-manifold J in W, an end reduction V of W at J exists and is unique up to non-ambient isotopy rel J in W.*

Proof. This follows from Theorems 2.1 and 2.3 of [1]. □

It may help the reader's intuition about V to see a brief sketch of its construction. We begin with a regular exhaustion $\{K_n\}_{n\geq 0}$ of W with $K_0 = J$.

Set $K_0^* = K_0$. If ∂K_1 is incompressible in $W - J$ set $K_1^* = K_1$. Otherwise we "completely compress" ∂K_1 in $W - K_0^*$ to obtain K_1^*. We may assume that $K_1^* \subseteq \text{int } K_2$. If ∂K_2 is incompressible in $W - J$ we set $K_2^* = K_2$. Otherwise we completely compress ∂K_2 in $W - K_1^*$ to get K_2^*. We continue in this fashion to construct a sequence $\{K_n^*\}_{n \geq 0}$. We let $V^* = \cup_{n \geq 0} K_n^*$ and then let V be the component of V^* containing J.

Proposition (3.2). *Let V be an end reduction of W at J. Then the following hold:*

(1) *(Brin-Thickstun) If J' is regular in W, $J \subseteq \text{int } J'$, $J' \subseteq V$, and $\partial J'$ is incompressible in $W - J$, then V is an end reduction of W at J'.*

(2) *There is a knot κ in* int J *such that V is an end reduction of W at (a regular neighborhood of) κ.*

(3) *V is a Whitehead manifold.*

Proof. (1) is Corollary 2.2.1 of [1]. (2) is Lemma 2.4 of [12]. (3) is Lemma 2.6 of [12]. □

An end reduction V of W at J is *minimal* if whenever U is an end reduction of W at K and $U \subseteq V$, then there is a non-ambient isotopy of U to V in W. It is easily seen that genus one end reductions are minimal; recall that they are also $\mathbb{R}^2$-irreducible.

In [14] Tucker constructed a 3-manifold W_0 whose interior and boundary are homeomorphic, respectively, to $\mathbb{R}^3$ and $\mathbb{R}^2$ but which is not homeomorphic to $\mathbb{R}^2 \times [0, \infty)$. W_0 is a monotone union of solid tori which meet ∂W_0 in a monotone union of disks. It can be shown that the double of W_0 along its boundary is a Whitehead manifold which is a minimal end reduction of itself but is not $\mathbb{R}^2$-irreducible.

In [12] and this paper examples are given of $\mathbb{R}^2$-irreducible Whitehead manifolds having $\mathbb{R}^2$-irreducible end reductions which are not minimal.

If V is an end reduction of W, then we denote the non-ambient isotopy class of V in W by $[V]$. These isotopies are not required to be rel J. From now on we will usually drop the phrase "non-ambient" from "non-ambient isotopy". The vertices of $\mathcal{S}(W)$ are those $[V]$ for which V is minimal and $\mathbb{R}^2$-irreducible.

Distinct vertices $[V_0]$ and $[V_1]$ are joined by an edge if there is an $\mathbb{R}^2$-irreducible end reduction $E_{0,1}$ of W such that (1) $E_{0,1}$ contains representatives of $[V_0]$ and $[V_1]$, (2) every $\mathbb{R}^2$-irreducible end reduction of W contained in $E_{0,1}$ is isotopic in W to V_0, V_1, or $E_{0,1}$, and (3) $[E_{0,1}]$ is unique among $\mathbb{R}^2$-irreducible end reductions of W with respect to (1) and (2).

Three distinct vertices $[V_0]$, $[V_1]$, and $[V_2]$ span a 2-simplex of $\mathcal{S}(W)$ if each pair of vertices is joined by an edge and there is an $\mathbb{R}^2$-irreducible end reduction $T_{0,1,2}$ of W such that (1) $T_{0,1,2}$ contains representatives of each $[V_i]$ and $[E_{i,j}]$, (2) every $\mathbb{R}^2$-irreducible end reduction of W contained in $T_{0,1,2}$ is isotopic in W to one of the V_i or $E_{i,j}$ or to $T_{0,1,2}$, (3) $[T_{0,1,2}]$ is unique among $\mathbb{R}^2$-irreducible end reductions of W with respect to (1) and (2).

There is an obvious generalization of these definitions which inductively defines simplices of higher dimensions.

Let $\mathrm{Hom}(W)$ denote the group of self-homeomorphisms of W. Let $\mathrm{Aut}(\mathcal{S}(W))$ denote the group of simplicial automorphisms of $\mathcal{S}(W)$. Each $g \in \mathrm{Hom}(W)$ induces a $\gamma \in \mathrm{Aut}(\mathcal{S}(W))$. Let $\Psi : \mathrm{Hom}(W) \to \mathrm{Aut}(\mathcal{S}(W))$ be the homomorphism given by $\Psi(g) = \gamma$.

Theorem (3.3). *If W is a non-trivial covering space of a 3-manifold M with group of covering translations $G \cong \pi_1(M)$, then the restriction $\Psi|G : G \to \mathrm{Aut}(\mathcal{S}(W))$ is one to one.*

Proof. This is proved in [12]. □

Corollary (3.4). *If $\mathcal{S}(W)$ is isomorphic to a triangulation of $\mathbb{R}$, then $\pi_1(M) \cong \mathbf{Z}$.*

Proof. $\pi_1(M)$ must be torsion-free. The only non-trivial torsion-free subgroups of the infinite dihedral group $\mathrm{Aut}(\mathcal{S}(W))$ are infinite cyclic. □

4. Some poly-excellent graphs in the 3-ball

A compact, connected, orientable 3-manifold is *superb* if it is irreducible, ∂-irreducible, and anannular, it contains a two-sided, properly embedded incompressible surface, and it is not a 3-ball. It is *excellent* if, in addition, it is atoroidal. In this paper superb 3-manifolds which are not excellent will occur only in the last section. A compact, properly embedded 1-manifold in a compact, connected, orientable 3-manifold is *superb* or *excellent* if its exterior is, respectively, superb or excellent. It is *poly-superb* or *poly-excellent* if for each non-empty collection of its components the union of that collection is, respectively, superb or excellent.

Define a *k-tangle* to be a disjoint union of k properly embedded arcs in a 3-ball.

Lemma (4.1). *For all $k \geq 1$ poly-excellent k-tangles exist.*

Proof. This is Theorem 6.3 of [11]. □

In this section we generalize this to certain graphs in the 3-ball. For $n \geq 2$ define an *n-frame* F to be a graph having one vertex of degree n and n vertices of degree one; thus it is the cone on a set of n points. A *subframe* of F is a subgraph of F which is an m-frame for some $m \geq 2$. Note that a single edge of F is not a subframe of F.

F is *properly embedded* in a 3-ball B if $F \cap \partial B$ is the set of vertices of F of degree one. A *system of frames in B* is a disjoint union $\mathcal{F}$ of finitely many properly embedded n_i-frames F_i in B. We say that $\mathcal{F}$ is *superb* or *excellent* if its exterior is, respectively, superb or excellent. It is *poly-superb* or *poly-excellent* if every non-empty subgraph of $\mathcal{F}$ whose components are subframes of the components of $\mathcal{F}$ is, respectively, superb or excellent. Note that the subgraph need not meet every component of $\mathcal{F}$.

Theorem (4.2). *Let $k \geq 1$. Suppose $n_1 \geq 2$. If $k \geq 2$ assume that $n_i = 2$ for $2 \leq i \leq k$. Then there exists a poly-excellent system $\mathcal{F}$ of n_i-frames F_i in the 3-ball B.*

In this paper we will need only the case $n_1 = 3$, but it is no harder to prove for $n_1 > 3$.

We will need the following lemma for gluing together superb or excellent 3-manifolds to obtain a superb or, respectively, excellent 3-manifold.

LEMMA (4.3). *Let Y be a compact, connected, orientable 3-manifold. Let S be a compact, properly embedded, two-sided surface in Y. Let Y' be the 3-manifold obtained by splitting Y along S. Let S' and S'' be the two copies of S which are identified to obtain Y. If each component of Y' is superb (respectively excellent), S', S'', and $(\partial Y') - \operatorname{int}(S' \cup S'')$ are incompressible in Y', and each component of S has negative Euler characteristic, then Y is superb (respectively excellent).*

Proof. In the excellent case this is Lemma 2.1 of [10]. The superb case follows from the proof of that lemma. □

Proof of Theorem (4.2). By Lemma (4.1) we may assume that $n_1 \geq 3$.

We first prove the case $k = 1$. Let $n = n_1$. Let (ρ, θ, ϕ), $\rho \geq 0$, $0 \leq \theta \leq 2\pi$, $0 \leq \phi \leq \pi$, be spherical coordinates in $\mathbb{R}^3$. We regard B as the set $\rho \leq 2$. Let B' be the set $\rho \leq 1$. Let Σ be the spherical shell $B - \operatorname{int} B'$. The n halfplanes $\theta = 0, 2\pi/n, \ldots, 2\pi(n-1)/n$ meet Σ in disks $D_0, D_1, \ldots, D_{n-1}$ whose union cuts Σ into 3-balls $B_0, B_1, \ldots, B_{n-1}$, where $\partial B_j = D_j \cup D_{j+1} \cup E_j \cup E'_j$ (subscripts taken mod n), where $E_j = B_j \cap \partial B$ and $E'_j = B_j \cap \partial B'$. We may think of Σ as a cantaloupe which has been cut into n wedge shaped slices and whose seeds have been removed. See Figure 1 for a schematic diagram of the following construction.

In each B_j we choose a poly-excellent $(n+1)$-tangle $\alpha_{j,0} \cup \alpha_{j,1} \cup \cdots \cup \alpha_{j,n}$. We require (taking the subscript j mod n) that $\alpha_{j,0}$ runs from int E_j to int D_{j+1}, $\alpha_{j,p}$ runs from int D_j to int D_{j+1} for $1 \leq p \leq n-1$, and $\alpha_{j,n}$ runs from int D_j to int E'_j. In addition we require that $\alpha_{j,p} \cap D_{j+1} = \alpha_{j+1,p+1} \cap D_{j+1}$. We then let $\beta_j = \alpha_{j,0} \cup \alpha_{j+1,1} \cup \cdots \cup \alpha_{j-1,n-1} \cup \alpha_{j,n}$. The β_j are disjoint arcs each of which joins ∂B to $\partial B'$ in Σ. We may think of regular neighborhoods of the β_j as tunnels eaten out of the cantaloupe by n worms who start on the outside and eat their way to the seed chamber in such a way that they each wind all the way around the cantaloupe, passing through every slice from one side to the other while coordinating their movements so that the union of the tunnels in each slice is poly-excellent.

The exterior in Σ of the union of the β_j is equal to the exterior in B of an n-frame F. We claim that F is poly-excellent. Let F' be an m-frame which is a subframe of F. Let X_j be the exterior of $F' \cap B_j$, and let $S_j = X_j \cap D_j$. Each X_j is excellent. Since $m \geq 2$ we have that F' meets each D_j at least twice. Thus $\chi(S_j) < 0$. Since no arc $\alpha_{j,p}$ joins D_j to itself or D_{j+1} to itself we have that S_j and S_{j+1} are each incompressible in X_j. Since X_j is ∂-irreducible and neither S_j nor S_{j+1} is a disk we have that $\partial X_j - \operatorname{int} S_j$ and $\partial X_j - \operatorname{int} S_{j+1}$ are incompressible in S_j. By successive applications of Lemma 4.3 we get that $X'_0 = X_1 \cup \cdots \cup X_{n-1}$ is excellent. Now X_0 and X'_0 are glued along the surface $S_0 \cup S_1$, which is a disk with $2m+1$ holes. $\partial X_0 - \operatorname{int}(S_0 \cup S_1)$ is the disjoint union of two annuli. Since X_0 is ∂-irreducible it follows that $S_0 \cup S_1$ and $\partial X_0 - \operatorname{int}(S_0 \cup S_1)$ are incompressible in X_0. Now $\partial X'_0 - \operatorname{int}(S_0 \cup S_1)$ is the disjoint union of two disks with m

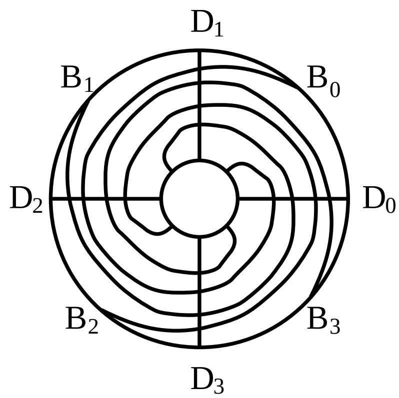


Figure 1. The cantaloupe trick

holes. Since X_0' is ∂-irreducible it follows that $S_0 \cup S_1$ and $\partial X_0' - \operatorname{int}(S_0 \cup S_1)$ are incompressible in X_0'. So by Lemma 4.3 the exterior $X_0 \cup X_0'$ of F' is excellent.

We next prove the case $k > 1$. We modify the construction of the previous case as follows. In B_0 we choose a poly-excellent $(n+k)$-tangle $\alpha_{j,0} \cup \alpha_{j,1} \cup \cdots \cup \alpha_{j,n} \cup \gamma_2 \cup \cdots \cup \gamma_k$, where each γ_q runs from int E_0 to itself. The $\alpha_{0,p}$ have the same properties as before. There is no change in the B_j for $j \neq 0$. Each γ_q is an arc and hence can be regarded as a 2-frame. The proof of poly-excellence works much as before. The only notable difference is that if the n_1-frame is deleted, then B is the union of B_0 and a 3-ball along the disk $S_0 \cup S_1 \cup E_0'$, and so $\gamma_2 \cup \cdots \cup \gamma_k$ is poly-excellent in B. $\square$

5. Disk busting knots in handlebodies

In this section we consider a knot κ in the interior of a handlebody C which is embedded in the interior of a handlebody $\widehat{C}$. We assume that C and $\widehat{C}$ each have genus at least one. Let $\mathcal{D}$ be a disjoint union of finitely many properly embedded disks in C such that $\mathcal{D}$ splits C into a collection of 3-balls and no component of $\mathcal{D}$ is ∂-parallel in C. We say that κ is *$\mathcal{D}$-busting* if no compressing disk for ∂C in $C - \kappa$ has the same boundary as a component of $\mathcal{D}$. We give conditions on the embedding of C in $\widehat{C}$ which insure that if κ is $\mathcal{D}$-busting in C, then it is *disk busting in $\widehat{C}$*, by which we mean that $\partial\widehat{C}$ is incompressible in $\widehat{C} - \kappa$.

An *n-pod* is a pair (B, P) consisting of a 3-ball B and a disjoint union P of n disks in ∂B. The components of P are called the *feet* of the n-pod. For $n = 2$ or $n = 3$ we use the term *bipod* or *tripod*, respectively.

Two compact, properly embedded surfaces S and T in a 3-manifold are in *minimal general position* if they are in general position and among all such surfaces S' isotopic to S one has that $S \cap T$ has the fewest components.

Lemma (5.1). *Let $\widehat{C}$ be a handlebody of genus at least one. Let $\widehat{\mathcal{E}}$ be a disjoint union of properly embedded disks in $\widehat{C}$ which splits $\widehat{C}$ into a union $(\widehat{\mathcal{B}}, \widehat{\mathcal{P}})$ of bipods and tripods. Let κ be a knot in* int $\widehat{C}$ *which is in general position with respect to $\widehat{\mathcal{E}}$. Let $(\kappa', \partial\kappa')$ be the 1-manifold in $(\widehat{\mathcal{B}}, \widehat{\mathcal{P}})$ obtained by splitting κ along $\kappa \cap \widehat{\mathcal{E}}$. Suppose that*

(1) *$\widehat{\mathcal{P}} - \kappa'$ is incompressible in $\widehat{\mathcal{B}} - \kappa'$,*
(2) *$\widehat{\mathcal{P}} - \kappa'$ is ∂-incompressible in $\widehat{\mathcal{B}} - \kappa'$, and*
(3) *each foot of $(\widehat{\mathcal{B}}, \widehat{\mathcal{P}})$ meets κ'.*
Then κ is disk busting in $\widehat{C}$.

Proof. Suppose D is a compressing disk for $\partial\widehat{C}$ in $\widehat{C} - \kappa$. Put D in minimal general position with respect to $\widehat{\mathcal{E}} - \kappa$.

Suppose $D \cap (\widehat{\mathcal{E}} - \kappa)$ contains a simple closed curve γ. We may assume that γ is innermost on D, so $\gamma = \partial\Delta$ for a disk Δ in D with $\Delta \cap (\widehat{\mathcal{E}} - \kappa) = \gamma$. By (1) $\gamma = \partial\Delta'$ for a disk Δ' in $\widehat{\mathcal{P}} - \kappa'$. Then $\Delta \cup \Delta'$ is a 2-sphere which bounds a 3-ball in $\mathcal{B}$ which by (3) misses κ'. Thus there is an isotopy of D in $\widehat{C} - \kappa$ which removes at least γ from the intersection, thereby contradicting minimality.

Now suppose that $D \cap (\widehat{\mathcal{E}} - \kappa)$ has a component α which is an arc. We may assume that α is outermost on D, so there is an arc β in ∂D such that $\partial\alpha = \partial\beta$ and $\alpha \cup \beta = \partial\Delta$ for a disk Δ in D with $\Delta \cap (\widehat{\mathcal{E}} - \kappa) = \alpha$. By (2) there is a disk Δ' in $\widehat{\mathcal{P}} - \kappa'$ and an arc α' in $\partial\widehat{\mathcal{P}}$ such that $\alpha \cap \alpha' = \partial\alpha = \partial\alpha'$ and $\partial\Delta' = \alpha \cup \alpha'$. Then $\Delta \cup \Delta'$ is a disk with $\partial(\Delta \cup \Delta') = \alpha' \cup \beta$. By (1) $\alpha' \cup \beta = \partial\Delta''$ for a disk Δ'' in $\partial\widehat{\mathcal{B}} - \operatorname{Int}\widehat{\mathcal{P}}$. We have that $\Delta \cup \Delta' \cup \Delta''$ is a 2-sphere bounding a 3-ball in $\widehat{\mathcal{B}}$ which by (3) misses κ'. Thus there is an isotopy of D in $\widehat{C} - \kappa$ which removes at least α from the intersection, thereby contradicting minimality.

We now have that $D \cap (\widehat{\mathcal{E}} - \kappa) = \emptyset$, so D lies in some component of $\widehat{\mathcal{B}}$. If ∂D does not bound a disk in $\partial\widehat{\mathcal{B}} - \operatorname{Int}\widehat{\mathcal{P}}$, then it is parallel in this surface to a component of $\partial\widehat{\mathcal{P}}$, thereby contradicting (1) and (3). $\square$

An n-pod (B, P) is *properly embedded* in an m-pod $(\widehat{B}, \widehat{P})$ if $B \subseteq \widehat{B}$ and $B \cap \partial\widehat{B} = B \cap \operatorname{int}\widehat{P} = P$. Note that (B, P) is a regular neighborhood of an n-frame in $\widehat{B}$.

Lemma (5.2). *Let $(\widehat{B}, \widehat{P})$ be a bipod or tripod. Let $(\mathcal{B}, \mathcal{P})$ be a disjoint union of bipods and tripods properly embedded in $(\widehat{B}, \widehat{P})$. Let λ be a disjoint union of arcs properly embedded in $\mathcal{B}$ with $\lambda \cap \partial\mathcal{B} = \partial\lambda \subseteq \mathcal{P}$. Suppose that*
(i) *$\mathcal{P} - \lambda$ is incompressible in $\mathcal{B} - \lambda$,*
(ii) *$\mathcal{P} - \lambda$ is ∂-incompressible in $\mathcal{B} - \lambda$,*
(iii) *each foot of $\mathcal{P}$ meets λ*
(iv) *each foot of $\widehat{P}$ meets $\mathcal{P}$,*
(v) *$\partial\mathcal{B} - \operatorname{int}\mathcal{P}$ and $\widehat{P} - \operatorname{int}\mathcal{P}$ are incompressible in $\widehat{B} - \operatorname{int}\mathcal{B}$, and*
(vi) *if any component of $(\mathcal{B}, \mathcal{P})$ is a tripod, then $\widehat{B} - \operatorname{Int}\mathcal{B}$ is ∂-irreducible.*

Then
(1) *$\widehat{P} - \lambda$ is incompressible in $\widehat{B} - \lambda$,*
(2) *$\widehat{P} - \lambda$ is ∂-incompressible in $\widehat{B} - \lambda$, and*
(3) *each foot of $\widehat{P}$ meets λ.*

Proof. Suppose D is a compressing disk for $\widehat{P} - \lambda$ in $\widehat{B} - \lambda$. Put D in minimal general position with respect to $\partial\mathcal{B} - \operatorname{int}\mathcal{P}$.

Suppose $D \cap (\partial\mathcal{B} - \operatorname{int}\mathcal{P})$ has a simple closed curve component γ. We may assume that γ is innermost on D, so $\gamma = \partial\Delta$ for a disk Δ in D with $\Delta \cap (\partial\mathcal{B} - \operatorname{int}\mathcal{P}) = \gamma$.

If Δ lies in $\mathcal{B} - \lambda$, then it follows from (i) and (iii) that $\gamma = \partial\Delta'$ for a disk Δ' in $\partial\mathcal{B} - \text{int } \mathcal{P}$. Then $\Delta \cup \Delta'$ is a 2-sphere which bounds a 3-ball in $\mathcal{B}$ which misses λ, so there is an isotopy of D in $\widehat{B} - \lambda$ which removes at least γ from the intersection, contradicting minimality.

If Δ lies in $\widehat{B} - \text{int } \mathcal{B}$, then by (v) there is a disk Δ' in $\partial\widehat{B} - \text{Int } \mathcal{P}$ such that $\gamma = \partial\Delta'$. Then $\Delta \cup \Delta'$ is a 2-sphere which bounds a 3-ball in $\widehat{B} - \text{Int } \mathcal{B}$ which misses λ, so there is an isotopy of D in $\widehat{B} - \lambda$ which removes at least γ from the intersection, contradicting minimality.

Thus there are no simple closed curve components. Suppose there is a component α which is an arc. We may assume that α is outermost on D, so there is an arc β in ∂D such that $\partial\alpha = \partial\beta$ and $\alpha \cup \beta = \partial\Delta$ for a disk Δ in D with $\Delta \cap (\partial\mathcal{B} - \text{int } \mathcal{P}) = \alpha$.

If Δ lies in $\mathcal{B} - \lambda$, then β lies in $\mathcal{P} - \lambda$. By (ii) there is a disk Δ' in $\mathcal{P} - \lambda$ and an arc β' in $\partial\mathcal{P}$ such that $\beta \cap \beta' = \partial\beta = \partial\beta'$ and $\partial\Delta' = \beta \cup \beta'$. Then $\Delta \cup \Delta'$ is a disk with $\partial(\Delta \cup \Delta') = \alpha \cup \beta'$. By (i) there is a disk Δ'' in $\partial\mathcal{B} - \text{int } \mathcal{P}$ with $\partial\Delta'' = \alpha \cup \beta'$. Then $\Delta \cup \Delta' \cup \Delta''$ is a 2-sphere which bounds a 3-ball in $\mathcal{B}$ that misses λ. Thus there is an isotopy of D in $\widehat{B} - \lambda$ which removes at least γ from the intersection, contradicting minimality.

If Δ lies in $\widehat{B} - \text{Int } \mathcal{B}$, then β lies in $\widehat{P} - \mathcal{P}$.

Suppose the component (B, P) of $(\mathcal{B}, \mathcal{P})$ containing α is a bipod. Then there is a disk Δ' in $\partial B - \text{int } P$ with $\partial\Delta' = \alpha \cup \alpha'$, where α' is an arc in ∂P with $\partial\alpha = \partial\alpha'$. So $\Delta \cup \Delta'$ is a disk in $\widehat{B} - \text{Int } \mathcal{B}$ with $\partial(\Delta \cup \Delta') = \alpha' \cup \beta$. By (v) there is a disk Δ'' in $\widehat{P} - \text{int } \mathcal{P}$ with $\partial\Delta'' = \alpha' \cup \beta$. Then $\Delta \cup \Delta' \cup \Delta''$ is a 2-sphere bounding a 3-ball in $\widehat{B}$ which misses λ. Thus there is an isotopy of D in $\widehat{B} - \kappa$ which removes at least α from the intersection, contradicting minimality.

Suppose the component (B, P) of $(\mathcal{B}, \mathcal{P})$ containing α is a tripod. By (vi) there is a disk Δ' in $\partial(\widehat{B} - \text{Int } \mathcal{B})$ such that $\partial\Delta' = \partial\Delta$. Since each component of $\partial\mathcal{P}$ is a non-separating curve on $\partial(\widehat{B} - \text{Int } \mathcal{B})$ we must have that $\partial\beta$ lies in a single component of $\partial\mathcal{P}$. Moreover Δ' is the union of a disk in $\partial\mathcal{B} - \text{Int } \mathcal{P}$ and a disk in $\widehat{P} - \text{int } \mathcal{P}$ which meet along an arc in $\partial\mathcal{P}$, and $\Delta \cup \Delta'$ is a 2-sphere bounding a 3-ball in $\widehat{B}$ which misses λ. Thus there is an isotopy of D in $\widehat{B} - \lambda$ which removes at least α from the intersection, contradicting minimality.

So we have that D misses $\partial\mathcal{B} - \text{int } \mathcal{P}$. If D lies in $\mathcal{B} - \lambda$, then by (i) $\partial D = \partial D'$ for a disk D' in $\mathcal{P} - \lambda$. If D lies in $\widehat{B} - \text{Int } \mathcal{B}$, then by (v) $\partial D = \partial D'$ for a disk D' in $\widehat{P} - \text{int } \mathcal{P}$. This completes the proof of (1).

Now suppose that D is a ∂-compressing disk for $\widehat{P} - \lambda$ in $\widehat{B} - \lambda$. We have that $\partial D = \gamma \cup \delta$ for arcs γ in $\widehat{P} - \lambda$ and δ in $\partial\widehat{B} - \text{int } \widehat{P}$. Put D in minimal general position with respect to $\partial\mathcal{B} - \text{int } \mathcal{P}$. As in the proof of (1) we may assume that no component of the intersection is a simple closed curve.

Suppose the intersection has a component α which is an arc. We may assume that α is outermost with respect to δ, by which we mean that there is a disk Δ in D and an arc β in γ such that $\partial\alpha = \partial\beta$, $\partial\Delta = \alpha \cup \beta$, and $\Delta \cap (\partial\mathcal{B} - \text{int } \mathcal{P}) = \alpha$. The analysis of Δ now proceeds as in the proof of (1), and we again contradict minimality.

So D misses $\partial\mathcal{B} - \text{int } P$, and D lies in $\widehat{B} - \text{Int } \mathcal{B}$. By (iv) $\partial D = \partial D'$ for a disk D' in $\partial(\widehat{B} - \text{Int } \mathcal{B})$. Since each component of $\partial\mathcal{P}$ is non-separating in $\partial(\widehat{B} - \text{Int } \mathcal{B})$ we have that $D' \cap \mathcal{P}$ is a disk. This completes the proof of (2).

(3) follows from (iii) and (iv). $\square$

A disjoint union of n_i-pods (B_i, P_i) properly embedded in an m-pod $(\widehat{B}, \widehat{P})$ is *poly-superb* or *poly-excellent* if the corresponding union of n_i-frames is, respectively, poly-superb or poly-excellent.

We suppose now that $\widehat{\mathcal{E}}$ is a disjoint union of properly embedded disks in $\widehat{C}$ which splits $\widehat{C}$ into a union $(\widehat{\mathcal{B}}, \widehat{\mathcal{P}})$ of bipods and tripods. These bipods and tripods and their feet are called *big*. We assume that $\widehat{\mathcal{E}} \cap C$ is a union $\mathcal{E}$ of properly embedded disks in C which splits C into a union $(\mathcal{B}, \mathcal{P})$ of bipods and tripods. These bipods and tripods and their feet are called *small*. We further assume that $\mathcal{D} \subseteq \mathcal{E}$. A small foot is called *hot* if it is parallel in C to a component of $\mathcal{D}$. It is *warm* if there is no compressing disk for ∂C in $C - \kappa$ which has the same boundary. It is *cold* if there is such a compressing disk. Note that every hot foot is warm.

Proposition (5.3). *Suppose that for each big bipod or tripod* $(\widehat{B}, \widehat{P})$

(1) *each big foot of* $(\widehat{B}, \widehat{P})$ *contains a small warm foot of* $(\widehat{B}, \widehat{P}) \cap (\mathcal{B}, \mathcal{P})$*, and*

(2) *either*

(a) $(\widehat{B}, \widehat{P}) \cap (\mathcal{B}, \mathcal{P})$ *is poly-superb, or*

(b) $(\widehat{B}, \widehat{P})$ *is a bipod,* $(\widehat{B}, \widehat{P}) \cap (\mathcal{B}, \mathcal{P})$ *consists of bipods, and each of these small bipods meets each of the two big feet of* $(\widehat{B}, \widehat{P})$.

Then every $\mathcal{D}$*-busting knot* κ *in* C *is disk busting in* $\widehat{C}$.

Proof. Suppose κ is $\mathcal{D}$-busting in C. Isotop κ in C so that it is in minimal general position with respect to $\mathcal{E}$. We will show that after possibly modifying $(\mathcal{B}, \mathcal{P})$ we will have that for each big bipod or tripod $(\widehat{B}, \widehat{P})$ it is the case that $(\widehat{B}, \widehat{P}) \cap (\mathcal{B}, \mathcal{P})$ satisfies the hypotheses of Lemma 5.2 with $\lambda = \kappa \cap \widehat{B}$. Note that we do not require that the components of the modified $(\mathcal{B}, \mathcal{P})$ match up along $\widehat{\mathcal{E}}$ to give a new handlebody in $\widehat{C}$.

So let $(\widehat{B}, \widehat{P})$ be a big bipod or tripod.

Suppose we are in case 2(a).

Consider a small bipod (B, P) in $(\widehat{B}, \widehat{P}) \cap (\mathcal{B}, \mathcal{P})$. If $\lambda \cap B = \emptyset$, then we discard (B, P) from $(\mathcal{B}, \mathcal{P})$ to obtain a new poly-superb system. If $\lambda \cap B \neq \emptyset$, then by minimality λ meets each small foot of (B, P) and $P - \lambda$ is incompressible in $B - \lambda$. Since (B, P) is a bipod we then have that $P - \lambda$ is ∂-incompressible in $B - \lambda$.

Consider a small tripod (B, P). If $\lambda \cap B = \emptyset$, then we discard (B, P) from $(\mathcal{B}, \mathcal{P})$ to obtain a new poly-superb system. If $\lambda \cap B \neq \emptyset$, then by minimality λ meets at least two small feet of (B, P).

Suppose λ misses the third small foot. Then we push that foot slightly into int $\widehat{B}$ to obtain a bipod. This gives a new poly-superb system. We have that λ meets each foot of the new (B, P), and $P - \lambda$ is incompressible and ∂-incompressible in $B - \lambda$.

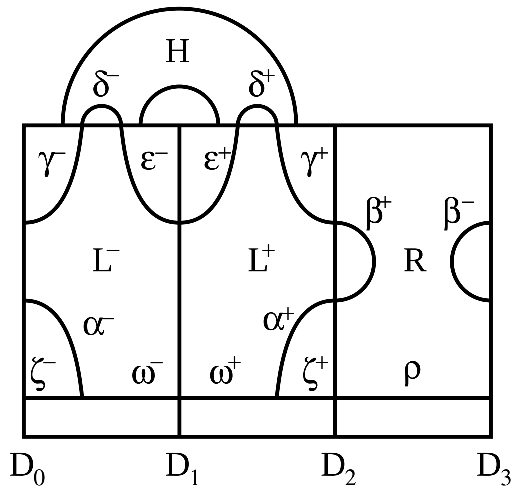


Figure 2. The systems of frames in the pieces of J

Suppose λ meets the third small foot. Then $P-\lambda$ is incompressible in $B-\lambda$. If $P-\lambda$ is ∂-compressible in $B-\lambda$, then there is a properly embedded disk Δ in $B-\lambda$ which meets a component E of P in an arc α and $\partial B - \operatorname{int} P$ in an arc β such that $\partial\alpha = \partial\beta$, $\partial\Delta = \alpha\cup\beta$, and α splits E into two disks each of which meets λ. Since $E-\lambda$ is incompressible in $B-\lambda$ we must have that the two components of $P-E$ are separated from each other by Δ. We split (B,P) along (Δ,α) to obtain two bipods (B',P') and (B'',P''). We have that $(P'\cup P'')-\lambda$ is incompressible and ∂-incompressible in $(B'\cup B'')-\lambda$. The exterior of the new system is homeomorphic to that of the old system by a homeomorphism which is the identity on the other components of $\partial\mathcal{B} - \operatorname{int}\mathcal{P}$, and so the new system is also poly-superb.

The feet discarded by our modification are precisely the cold feet of $(\widehat{B},\widehat{P})\cap(\mathcal{B},\mathcal{P})$. Some warm feet may be split into pairs of warm feet. It follows that conditions (i), (ii), and (iii) of Lemma (5.2) are satisfied. Since each component of $\widehat{P}$ contains a warm foot condition (iv) is satisfied. Since our modifications preserve poly-superbness condtions (v) and (vi) are also satisfied.

Now suppose that we are in case 2(b). As in the previous case we discard all small bipods with cold feet and get that conditions (i), (ii), (iii), and (iv) are satisfied. Since each small bipod joins the two big feet condition (v) is satisfied. Condition (vi) is vacuously satisfied.

The result now follows from Lemmas (5.1) and (5.2). $\square$

6. The construction of W

In this section we construct an $\mathbb{R}^2$-irreducible contractible open 3-manifold W which covers a 3-manifold $W^{\#}$ with $\pi_1(W^{\#})\cong\mathbf{Z}$. It will be shown that $\mathcal{S}(W)$ is a triangulation of $\mathbb{R}$ and hence every 3-manifold non-trivially covered by W must have fundamental group $\mathbf{Z}$.

Let $P = D \times [0,3]$, where D is a closed disk. Let $L^- = D \times [0,1]$, $L^+ = D \times [1,2]$, $R = D \times [2,3]$, and $D_j = D \times \{j\}$ for $j = 0,1,2,3$. Let $L = L^- \cup L^+$. Attach a 1-handle H to P so that it joins $\partial D \times (0,1)$ to $\partial D \times (1,2)$, thus giving a solid torus $J = P \cup H$. Let $J^\#$ be the genus two handlebody obtained from J by identifying D_0 and D_3. Let $P^\#$ be the solid torus in $J^\#$ which is the image of P under the identification. With the exceptions of J, $J^\#$, P, and $P^\#$ we will usually use the same symbol for subsets of J and their images in $J^\#$, relying on the context for the meaning. Thus we write, for example, $J^\# = P^\# \cup H$.

We next define a certain graph θ in $J^\#$ as follows. See Figure 2 for a schematic diagram of this construction.

Choose a poly-superb system of frames in L^- consisting of a 3-frame and two 2-frames. The 3-frame consists of arcs α^-, ζ^-, and ω^- meeting in a common endpoint in int L^-. The other endpoints of α^- and ζ^- lie in int D_0. The other endpoint of ω^- lies in int D_1. One 2-frame is an arc γ^- joining int D_0 and int $(L^- \cap H)$. The other 2-frame is an arc ε^- joining int D_1 and int $(L^- \cap H)$.

Let r be the homeomorphism $r(x,t) = (x, 2-t)$ from $D \times [0,2]$ to itself which reflects in the disk D_1. We have that $r(L^-) = L^+$. Denote $r(\alpha^-)$ by α^+, etc. This defines a poly-superb system of frames in L^+.

Next choose a poly-superb 2-tangle in H with components δ^- and δ^+ such that $\partial\delta^\pm = (\gamma^\pm \cup \varepsilon^\pm) \cap H$. Then choose a poly-superb 3-tangle in R with components β^-, β^+, and ρ, where $\partial\beta^\pm = (\alpha^\pm \cup \gamma^\pm) \cap R$ and $\partial\rho = (\zeta^- \cup \zeta^+) \cap R$.

Let η be the arc $\alpha^- \cup \beta^- \cup \gamma^- \cup \delta^- \cup \varepsilon^- \cup \varepsilon^+ \cup \delta^+ \cup \gamma^+ \cup \beta^+ \cup \alpha^+$, λ the arc $\omega^- \cup \omega^+$, and μ the arc $\zeta^+ \cup \rho \cup \zeta^-$. Set $\theta = \eta \cup \lambda \cup \mu$.

For each integer $n \geq 0$ take a copy of each of these objects. Denote the n^{th} copy of D_j by $D_{n,j}$, that of each of the other objects by a subscript n. We regard the arcs and graphs with subscripts n as embedded in the 3-manifolds with subscript $n+1$.

We embed $J_n^\#$ in int $J_{n+1}^\#$ as follows. L_n is sent to $N(\lambda_n, L_{n+1})$. R_n is sent to $N(\mu_n \cap (P_{n+1}^\# - \text{int } L_n), P_{n+1}^\# - \text{int } L_n)$. H_n is sent to $N(\eta_n \cap (J_{n+1}^\# - \text{int } P_n), J_{n+1}^\# - \text{int } P_n)$.

Now let $W^\#$ be the direct limit of the $J_n^\#$, and let $p : W \to W^\#$ be the universal covering map. Then $\pi_1(W^\#)$ is infinite cyclic. Let $h : W \to W$ be a generator of the group of covering translations. We regard $p^{-1}(P_n^\#)$ as $D_n \times \mathbb{R}$ with $P_{n,j} = D_n \times [3j, 3j+3]$, $L_{n,j}^- = D_n \times [3j, 3j+1]$, $L_{n,j}^+ = D_n \times [3j+1, 3j+2]$, and $R_{n,j} = D_n \times [3j+2, 3j+3]$. We set $L_{n,j} = L_{n,j}^- \cup L_{n,j}^+$. We have that $p^{-1}(H_n)$ is a disjoint union of 1-handles $H_{n,j}$, where $H_{n,j}$ is attached to $\partial D_n \times (3j, 3j+2)$, thereby yielding a copy $J_{n,j} = P_{n,j} \cup H_{n,j}$ of J_n. Set $D_{n,k} = D_n \times \{k\}$ for $k \in \mathbf{Z}$. For all the objects with subscript n contained in J_{n+1} denote the component of the preimage contained in $J_{n+1,j}$ by the subscripts n, j. Let η_n and μ_n be the component of the preimage of η_n and μ_n, respectively, which meets ω_n^+. We assume that h is chosen so that $h(D_{n,k}) = D_{n,k+3}$ and the image under h of any object with subscripts n, j has subscripts $n, j+1$.

We next describe certain families of quasi-exhaustions in W. Let $\mathcal{P} = \{p_1, p_2, \ldots, p_m\}$ be a finite non-empty set of distinct integers with $p_1 < p_2 < \cdots < p_m$. We say that $\mathcal{P}$ is *good* if its elements are consecutive. Otherwise $\mathcal{P}$ is *bad*. If $m = 1$, then $\mathcal{P}$ is automatically good.

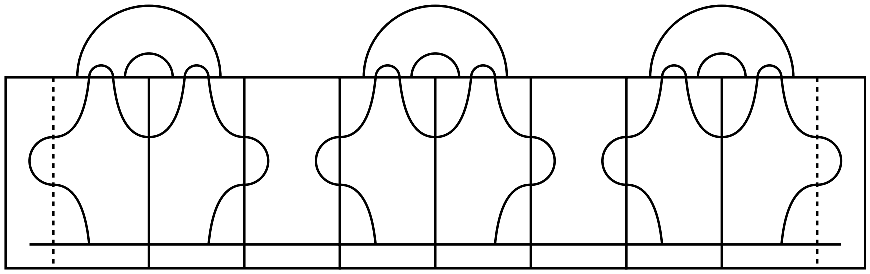


Figure 3. The embedding of $C_n^{\mathcal{P}}$ in $C_{n+1}^{\mathcal{P}}$ for a good $\mathcal{P}$

For $n \geq 0$ let $C_n^{\mathcal{P}}$ be the union of those $R_{n,j}$ with $p_1 - 1 \leq j \leq p_m$, those $L_{n,j}$ with $p_1 \leq j \leq p_m$, and those $H_{n,p}$ with $p \in \mathcal{P}$. Each $C_n^{\mathcal{P}}$ is a cube with m handles embedded in int $C_{n+1}^{\mathcal{P}}$. In Figure 3 we give a schematic diagram for the case of $\mathcal{P} = \{p, p+1, p+2\}$.

The quasi-exhaustion $\{C_n^{\mathcal{P}}\}_{n\geq 0}$ is denoted by $C^{\mathcal{P}}$; its union is denoted by $V^{\mathcal{P}}$. Whenever $\mathcal{P}$ is good and $m > 1$ we denote $V^{\mathcal{P}}$ by $V^{p,q}$, where $p = p_1$ and $q = p_m$. When $\mathcal{P} = \{p\}$ we use the notation V^p. The expressions $C_n^{p,q}$, C_n^p, $C^{p,q}$, and C^p are defined similarly.

7. Some properties of W

Given $\mathcal{P} = \{p_1, \ldots, p_m\}$ and $n > 0$, let $Y = C_{n+1}^{\mathcal{P}} - \text{int}\ C_n^{\mathcal{P}}$, $p = p_1$, and $q = p_m$. If $m > 1$ set $Z^- = Y \cap (R_{n+1,p-1} \cup L_{n+1,p}^-)$, $Z^+ = Y \cap (L_{n+1,q}^+ \cup R_{n+1,q})$, $Z = Z^- \cup Z^+$, and $X = Y - \text{Int}\ Z$.

Lemma (7.1). *Y is irreducible and ∂-irreducible.*

Proof. First consider the case $m = 1$. Then C_n^p is a solid torus in C_{n+1}^p with winding number zero. Any compressing disk for ∂C_{n+1}^p in Y would be a meridinal disk for C_{n+1}^p. Since $\delta_{n,p}^- \cup \delta_{n,p}^+$ is poly-superb in $H_{n+1,p}$ we have that $H_{n,p} \cap L_{n,p} \cap Y$ is incompressible in $H_{n+1,p} \cap Y$. It is incompressible in $(C_{n+1}^p - \text{Int}\ H_{n+1,p}) \cap Y$ for homological reasons. Thus $H_{n+1,p} \cap L_{n+1,p} \cap Y$ is incompressible in Y and thus so is ∂C_{n+1}^p. If ∂C_n^p is compressible in Y, then the union of C_n^p and a 2-handle with core the compressing disk is a 3-ball in C_{n+1}^p, and so ∂C_{n+1}^p is compressible in Y, a contradiction.

Now suppose $m > 1$. Consider the surfaces $D_{n+1,k} \cap Y$ for $3p+1 \leq k \leq 3q+1$ and $H_{n+1,p} \cap L_{n+1,p}^{\pm} \cap Y$ for $p \in \mathcal{P}$. They split Y into irreducible pieces. With the exception of $Z^{\pm}$ it follows from poly-superbness that each of these pieces is superb, and so each of those surfaces contained in its boundary is incompressible and ∂-incompressible. It follows that X is irreducible and ∂-irreducible. $Z^{\pm} \cap X$ consists of two disks with two holes, and $\partial Z^{\pm} - \text{int}\,(Z^{\pm} \cap X) = \partial(Z^{\pm} \cap X) \times [0,1]$. Thus $Z^{\pm} \cap X$ is incompressible and ∂-incompressible in $Z^{\pm}$. Thus the result follows. $\square$

Lemma (7.2). *$\partial C_n^{\mathcal{P}}$ is incompressible in $W - \text{int}\ C_n^{\mathcal{P}}$.*

Proof. A compressing disk D must lie in $C^{r,s}_m - \operatorname{int} C^{\mathcal{P}}_n$ for some $r \leq p$, $s \geq q$, and $m > n$. We can isotop D off compressing disks for $\partial C^{r,s}_m$ in $C^{r,s}_m - \operatorname{int} C^{\mathcal{P}}_n$ so that it lies in $C^{\mathcal{P}}_m - \operatorname{int} C^{\mathcal{P}}_n$. The result then follows from the previous lemma. $\square$

LEMMA (7.3). *If $\mathcal{P}$ is good and $m > 1$, then*

(1) *if A is an incompressible annulus in Y, then $\partial A = \partial A'$ for an annulus A' in ∂Y, and*

(2) *if T is an incompressible torus in Y, then T bounds a compact 3-manifold in Y.*

Proof. (1) Put A in minimal general position with respect to $X \cap Z$. Let α be a component of $A \cap X \cap Z$. Then α is not a simple closed curve bounding a disk in A.

Suppose α is an outermost arc on A, so $\partial\Delta = \alpha \cup \beta$ for an arc β in ∂A and a disk Δ in A with $\Delta \cap X \cap Z = \alpha$. If $\Delta \subseteq X$, then $\partial\Delta = \partial\Delta'$ for a disk Δ' in ∂X. Then $\Delta \cup \Delta'$ bounds a 3-ball in X, and an isotopy across it removes at least α from the intersection. If $\Delta \subseteq Z$, then β is ∂-parallel in one of the annuli comprising $\partial Z - \operatorname{int}(X \cap Z)$; it follows that one can again reduce the intersection. Thus α is not an arc.

So α is a simple closed curve. $\partial A' = \alpha \cup \beta$ for some annulus component A' of $A \cap X$ and some β in $(A \cap X \cap Z) \cup \partial A$. Then A' is parallel in X to an annulus A'' in ∂X. If A'' lies in $X \cap Z$, then we can isotop to remove at least $\alpha \cup \beta$. If A'' does not lie in $X \cap Z$, then either we can isotop to remove α or A'' contains an annulus component G of $\partial X - \operatorname{int}(X \cap Z)$. We may assume that the centerline of G is a meridian of $\beta^+_{n,q}$ and that the component of $X \cap Z$ containing $\partial A'$ is $F = H_{n+1,q} \cap L^+_{n+1,q} \cap Y$. We may further assume that all the components of $A \cap X$ are parallel to G and lie in $H_{n+1,q} \cap Y$. For homological reasons all the components of $A \cap Z$ must have their boundaries in the union of F and the two annulus components of $\partial C^{\mathcal{P}}_n \cap Z^+$. In particular, ∂A lies in the union of these two annuli and so bounds an annulus in their union with G.

Suppose $A \cap X \cap Z = \emptyset$. If $A \subseteq X$, then A is parallel in X to an annulus A' in ∂X with $\partial A'$ in $\partial X - \operatorname{int}(X \cap Z)$. It follows that A' lies in ∂Y. If $A \subseteq Z$, then for homological reasons ∂A must lie in one of the three annulus components of $\partial Z^{\pm} - \operatorname{int}(X \cap Z^{\pm})$.

(2) Suppose T is in minimal general position with respect to $X \cap Z$. T cannot lie in X since it would be ∂-parallel in X, but ∂X has no tori. If T lies in $Z^{\pm}$, then since $\partial Z^{\pm}$ is connected T must bound a compact 3-manifold in $Z^{\pm}$.

So we may assume that $T \cap Z \neq \emptyset$. Let A be a component of $T \cap X$. As in the proof of (1) we may assume that A is parallel in X to an annulus A' in ∂X which contains an annulus component G of $\partial X - \operatorname{int}(X \cap Z)$ whose centerline is a meridian of $\beta^+_{n,q}$ and that all such components are parallel to G and lie in $H_{n+1,q} \cap Y$. All the components of $T \cap Z^{\pm}$ must have their boundaries in the component F of $X \cap Z^{\pm}$ which meets G. So T lies in $(H_{n+1,q} \cup L^+_{n+1,q} \cup R_{n+1,q}) \cap Y$. Since this 3-manifold has connected boundary T must bound a compact 3-manifold in its interior. $\square$

LEMMA (7.4). *V^p does not embed in $\mathbb{R}^3$.*

Proof. Since $\beta^+ \cup \beta^- \cup \rho$ is poly-superb in R we have that $\beta^+_{n,p}$ is knotted in $R_{n,p}$. The result then follows from [4]. $\square$

PROPOSITION (7.5). *W is $\mathbb{R}^2$-irreducible. If $\mathcal{P}$ is good, then $V^{\mathcal{P}}$ is $\mathbb{R}^2$-irreducible.*

Proof. It suffices to show that for each good $\mathcal{P}$ the quasi-exhaustion $C^{\mathcal{P}}$ of W satisfies conditions (1)–(3) of Lemma (2.1). When $m = 1$ this follows from [9], so assume $m > 1$. Each $C_n^{\mathcal{P}}$ is a cube with handles, so is irreducible. We have that $\partial C_n^{\mathcal{P}}$ is incompressible in $W - \text{int } C_n^{\mathcal{P}}$ and that Y is ∂-irreducible and weakly anannular. $\square$

PROPOSITION (7.6). *If $\mathcal{P}$ is bad, then $V^{\mathcal{P}}$ is not $\mathbb{R}^2$-irreducible.*

Proof. There is an s such that $p < s < q$ and $s \notin \mathcal{P}$. We may assume that the embedding of $J_n^{\#}$ in $J_{n+1}^{\#}$ is such that $D_{n,1} \subseteq D_{n+1,1}$ for all $n \geq 0$. Then $D_{n,3s+1} \subseteq \text{int } D_{n+1,3s+1}$ for all $n \geq 0$. The union Π of these disks is a plane which is proper in $V^{\mathcal{P}}$ (but not in W!). $V^{\mathcal{P}} - \Pi$ has two components, one containing V^p and the other containing V^q. Since V^p and V^q do not embed in $\mathbb{R}^3$ we have that Π is non-trivial in $V^{\mathcal{P}}$. $\square$

A *classical knot space* is a space homeomorphic to the exterior of a non-trivial knot in S^3.

LEMMA (7.7). *If $\mathcal{P}$ is good and $m > 1$, then every incompressible torus T in $V^{\mathcal{P}} - \text{int } C_n^{\mathcal{P}}$ bounds a compact 3-manifold in $V^{\mathcal{P}} - \text{int } C_n^{\mathcal{P}}$.*

Proof. Assume that T is in minimal general position with respect to $\bigcup_{m \geq n} \partial C_m^{\mathcal{P}}$. If the intersection is empty then T lies in some Y and hence bounds a compact 3-manifold in Y. If the intersection is non-empty, then T meets a single $\partial C_m^{\mathcal{P}}$. Each annulus A into which $T \cap \partial C_m^{\mathcal{P}}$ splits T must have $\partial A = \partial A'$ for an annulus A' in $\partial C_m^{\mathcal{P}}$.

Consider an A in $S = C_m^{\mathcal{P}} - \text{int } C_n^{\mathcal{P}}$. Let $T' = A \cup A'$. Then $T' = \partial Q'$ for a compact 3-manifold Q' in $C_m^{\mathcal{P}}$. We may assume that $Q' \cap T = A$. Let $\widehat{S}$ and $\widehat{C}_m^{\mathcal{P}}$ be obtained by adding a collar C to these 3-manifolds in $V^{\mathcal{P}} - \text{int } C_m^{\mathcal{P}}$. We may assume that T meets C in a product annulus. If T' is incompressible in $\widehat{S}$, then Q' lies in S. If T' is compressible in $\widehat{S}$, then since $\widehat{S}$ is irreducible T' bounds a solid torus or a classical knot space in $\widehat{S}$. This must be Q'. So in either case Q' lies in S. Let T'' be the torus obtained from T by replacing A by A'. Then $T'' = \partial Q''$ for a compact 3-manifold Q'' in $V^{\mathcal{P}}$. If T'' is incompressible in $V^{\mathcal{P}} - \text{int } C_n^{\mathcal{P}}$, then by induction Q'' lies in $V^{\mathcal{P}} - \text{int } C_n^{\mathcal{P}}$. If T'' is compressible in $V^{\mathcal{P}} - \text{int } C_n^{\mathcal{P}}$, then by irreducibility T'' bounds a solid torus or classical knot space in $V^{\mathcal{P}} - \text{int } C_n^{\mathcal{P}}$. This must be Q''. So in either case Q'' is in $V^{\mathcal{P}} - \text{int } C_n^{\mathcal{P}}$. If $Q' \cap Q'' = A'$, then $T = \partial(Q' \cup Q'')$. If $Q' \cap Q'' \neq A'$, then $Q' \subseteq Q''$, and $T = \partial(Q'' - \text{Int } Q')$. $\square$

PROPOSITION (7.8). *$V^{\mathcal{P}}$ has finite genus. It has genus one if and only if $\mathcal{P}$ has exactly one element.*

Proof. $V^{\mathcal{P}}$ has genus at most m. Since V^p does not embed in $\mathbb{R}^3$ the genus of $V^{\mathcal{P}}$ must be at least one. So if $m = 1$, then V^p has genus one. Now suppose $m > 1$. If $V^{\mathcal{P}}$ has genus one, then it has a good exhaustion $\{K_n\}_{n\geq 0}$ by solid tori. Choose n and k such that $K_0 \subseteq \text{int } C_n^{\mathcal{P}}$ and $C_n^{\mathcal{P}} \subseteq \text{int } K_k$. Then since ∂K_k is incompressible in $V^{\mathcal{P}} - \text{int } K_0$ it is incompressible in the smaller space $V^{\mathcal{P}} - \text{int } C_n^{\mathcal{P}}$ and so bounds a compact 3-manifold in this space, which is impossible. Thus $V^{\mathcal{P}}$ has genus greater than one. □

8. The complex of end reductions of W

THEOREM (8.1). *Every $V^{\mathcal{P}}$ is an end reduction of W at each $C_n^{\mathcal{P}}$.*

Proof. We know that $V^{\mathcal{P}}$ is end irreducible rel $C_n^{\mathcal{P}}$ in W. Clearly $W - V^{\mathcal{P}}$ has no components with compact closure. Suppose N is a regular 3-manifold in W such that $C_n^{\mathcal{P}} \subseteq \text{int } N$ and ∂N is incompressible in $W - C_n^{\mathcal{P}}$. Then $N \subseteq \text{int } C_m^{r,s}$ for some $r \leq p$, $s \geq q$, and $m > n$. We isotop ∂N off a complete set of compressing disks for $\partial C_m^{r,s}$ in $C_m^{r,s} - \text{int } C_n^{\mathcal{P}}$ so that N lies in $C_m^{\mathcal{P}}$. This can be done with compact support in $W - \text{int } C_n^{\mathcal{P}}$. Running the isotopy backwards causes $V^{\mathcal{P}}$ to engulf N. □

THEOREM (8.2). *Let V be an end reduction of W at J, where $J \subseteq \text{int } C_n^{\mathcal{Q}}$. Then V is isotopic to $V^{\mathcal{P}}$ for some $\mathcal{P} \subseteq \mathcal{Q}$*

Proof. We may assume that V is an end reduction of W at a knot $\kappa \subseteq \text{int } J$. Let $\mathcal{P}$ be a minimal subset of $\mathcal{Q}$ such that, up to isotopy, $\kappa \subseteq \text{int } C_n^{\mathcal{P}}$ for some n. Let $\mathcal{D}$ be the union of the set of co-cores of the 1-handles $H_{n,p}$ with $p \in \mathcal{P}$. Then κ is $\mathcal{D}$-busting in $C_n^{\mathcal{P}}$.

If $m = 1$, then clearly κ is disk busting in C_n^p, so assume $m > 1$.

We let $\widehat{\mathcal{E}}$ be the union of the attaching disks for the $H_{n+1,p}$ with $p \in \mathcal{P}$ and the $D_{n+1,j}$ with $3p_1 + 1 \leq j \leq 3p_m + 1$. Let $\mathcal{E} = \widehat{\mathcal{E}} \cap C_n^{\mathcal{P}}$. We may assume that $\mathcal{D} \subseteq \mathcal{E}$. The conditions of Proposition (5.3) are satisfied, so κ is disk busting in $C_{n+1}^{\mathcal{P}}$. It follows that V is isotopic to $V^{\mathcal{P}}$. □

THEOREM (8.3). *$V^{\mathcal{P}}$ and $V^{\mathcal{Q}}$ are isotopic if and only if $\mathcal{P} = \mathcal{Q}$.*

Proof. We first consider the case $\mathcal{P} = \{p\}$, $\mathcal{Q} = \{q\}$, $p < q$. V^p is an end reduction of W at a knot κ in C_0^p. Let τ be the track of κ under an isotopy taking V^p to V^q and κ to κ'. Then $\tau \subseteq \text{int } C_n^{r,s}$ for some $r \leq p$, $q \leq s$, and $n \geq 0$. By the covering isotopy theorem [2, 3] there is an ambient isotopy of κ with track τ which has compact support in $C_n^{r,s}$. Let D be an attaching disk for $H_{n,p}$. Then κ is D-busting in $C_n^{r,s}$, but κ' is not. This is impossible since the isotopy is the identity on $\partial C_n^{r,s}$.

Now consider the general case. Suppoes $p \in \mathcal{P}$ and $p \notin \mathcal{Q}$. Then V^p is isotopic to $V^{\mathcal{R}}$ for some $\mathcal{R} \subseteq \mathcal{Q}$. Then we must have $\mathcal{R} = \{r\}$, where $r \neq p$, a contradiction. □

THEOREM (8.4). *$V^{\mathcal{P}}$ is minimal if and only if $\mathcal{P}$ has exactly one element.*

Proof. V^p is clearly minimal. If $m > 1$, then $V^{\mathcal{P}}$ contains V^p which is not homeomorphic to $V^{\mathcal{P}}$ since they have different genera. □

THEOREM (8.5). *$\mathcal{S}(W)$ is isomorphic to a triangulation of $\mathbb{R}$.*

Proof. The vertices of $\mathcal{S}(W)$ are the $[V^p]$, $p \in \mathbf{Z}$. We have that $[V^p]$ and $[V^{p+1}]$ are joined by the edge $[V^{p,p+1}]$. Every end reduction of W contained in $V^{p,p+1}$ is isotopic to V^p, V^{p+1}, or $V^{p,p+1}$. If V is an end reduction of W which contains representatives of V^p and V^q, where $p < q$, then V is isotopic to $V^{\mathcal{P}}$, where $p, q \in \mathcal{P}$. If $\mathcal{P} \neq \{p,q\}$, then $V^{\mathcal{P}}$ contains some V^r, $p \neq r \neq q$, so $[V]$ is not an edge joining $[V^p]$ and $[V^q]$. If $\mathcal{P} = \{p,q\}$ and $q > p+1$, then $\mathcal{P}$ is bad, so V is not $\mathbb{R}^2$-irreducible, so again $[V]$ is not an edge. The result follows. □

Corollary (8.6). *If W is a non-trivial covering space of a 3-manifold M, then $\pi_1(M) \cong \mathbf{Z}$.*

Proof. This follows immediately from Theorem (8.5) and Corollary (3.4). □

9. Uncountably many W

Theorem (9.1). *There are uncountably many pairwise non-homeomorphic W each of which has all the properties of sections 7 and 8.*

Proof. Recall that all of the genus one end reductions V^p of a fixed W resulting from our construction are homeomorphic. We will modify our construction to obtain uncountably many W such that different W have non-homeomorphic V^p.

In our construction of $W^{\#}$ we used a copy of the same 2-tangle $\delta^- \cup \delta^+$ in H for each 2-tangle $\delta_n^- \cup \delta_n^+$ in H_{n+1}. We will now change this so that the 2-tangle depends on n.

We say that a 3-manifold Q is *incompressibly embedded* in a 3-manifold X if $Q \subseteq X$ and ∂Q is incompressible in X.

Lemma (9.2). *Given an excellent classical knot space Q, there is a poly-superb 2-tangle τ in a 3-ball B with exterior X such that Q is incompressibly embedded in X and every incompressible torus in X is isotopic to ∂Q.*

Proof. Let B_0 and B_1 be 3-balls. Let D_i be a disk in ∂B_i. Let A_i be an annulus in int D_i. Let F_i be the annulus component of $D_i -$ int A_i; let E_i be the disk component. Let $\lambda_i^- \cup \lambda_i^+ \cup \mu_i^- \cup \mu_i^+$ be a poly-excellent 4-tangle in B_i. We require that $\lambda_0^{\pm}$ join $\partial B_0 -$ int D_0 to int F_0, $\mu_1^{\pm}$ join int F_1 to int E_1 , $\mu_0^{\pm}$ join E_0 to itself, and $\lambda_1^{\pm}$ join int E_1 to $\partial B_1 -$ int D_1. We now glue B_0 to B_1 by identifying $F_0 \cup E_0$ with $F_1 \cup E_1$ in such a way that $\lambda_0^{\pm} \cup \mu_1^{\pm} \cup \mu_0^{\pm} \cup \lambda_1^{\pm}$ is an arc $\delta^{\pm}$. By Lemma (2.1) $\delta^- \cup \delta^+$ is a poly-excellent system of two arcs in a 3-ball minus the interior of an unknotted solid torus with boundary $A_0 \cup A_1$. We then glue Q to this space by identifying ∂Q with $A_0 \cup A_1$ so that a meridian of Q is glued to $\partial E_0 = \partial E_1$. The result is a 3-ball B containing a 2-tangle $\tau = \delta^- \cup \delta^+$. Standard arguments then complete the proof. □

Recall that V^p is the monotone union of solid tori C_n^p, where $C_n^p = R_{n,p-1} \cup L_{n,p} \cup R_{n,p} \cup H_{n,p}$. Let $G_{n,p} = R_{n,p-1} \cup L_{n,p} \cup R_{n,p}$, $Y_{n+1,p} = C_{n+1}^p -$ int C_n^p, $X_{n+1,p} = Y_{n+1,p} \cap H_{n+1,p}$, and $Z_{n+1,p} = Y_{n+1,p} \cap G_{n+1,p}$.

Note that for all n and p the spaces $Z_{n+1,p}$ are homeomorphic. It thus follows from the Jaco-Shalen-Johannson characteristic submanifold theory [6, 7, 8] that there are, up to homeomorphism, only finitely many excellent classical knot spaces which incompressibly embed in $Z_{n+1,p}$. Denote this set by $\mathcal{N}$.

Let $\mathcal{Y}$ be the set of all homeomorphism types of excellent classical knot spaces which are not in $\mathcal{N}$. For each infinite subset $\mathcal{S}$ of $\mathcal{Y}$ we construct a W as follows. Choose a bijection of $\mathcal{S}$ with the set $\mathbf{N}$ of natural numbers. For each $n \in \mathbf{N}$ use the corresponding knot space Q_n in the construction of the 2-tangle τ_n in the previous lemma. Then use τ_n for $\delta^-_{n-1} \cup \delta^+_{n-1}$ in H_n. It follows that for each $n \geq m \geq 0$ we have that Q^p_{n+1} is incompressibly embedded in $V^p - \text{int } C^p_m$.

Lemma (9.3). *Suppose $Q \in \mathcal{Y}$ and Q is incompressibly embedded in $V^p - \text{int } C^p_m$. Then $Q \in \mathcal{S}$.*

Proof. Since Q is excellent it can be isotoped off $\cup_{n>m} \partial C^p_n$. It then lies in some $Y_{n+1,p}$. Since each $X_{n+1,p}$ is superb it can then be isotoped off $X_{n+1,p} \cap Z_{n+1,p}$. Since $Q \notin \mathcal{N}$ it must lie in $X_{n+1,p}$ and thus be isotopic to Q^p_{n+1}. $\square$

Now suppose that W' is constructed using $\mathcal{S}'$. Drop p from the notation and denote the corresponding submanifolds of W and W' by V and V', C_n and C'_n, etc.

Lemma (9.4). *If V and V' are homeomorphic, then there are finite subsets $\mathcal{S}_0$ of $\mathcal{S}$ and $\mathcal{S}'_0$ of $\mathcal{S}'$ such that $\mathcal{S} - \mathcal{S}_0 = \mathcal{S}' - \mathcal{S}'_0$.*

Proof. Suppose $h : V \to V'$ is a homeomorphism. Choose m and k such that $h(C_0) \subseteq \text{int } C'_m$ and $C'_m \subseteq \text{int } h(C_k)$. Then for all $n \geq k$ we have that $h(\partial C_n)$ is incompressible in $V' - \text{int } h(C_0)$ and hence is incompressible in the smaller space $V' - \text{int } C'_m$. It follows that $h(Q_{n+1})$ is isotopic in this space to some Q'_{j+1} with $j \geq m$. Let $\mathcal{A} = \{Q_1, \ldots, Q_k\}$. Then $\mathcal{S} - \mathcal{A} \subseteq \mathcal{S}'$. A similar argument using h^{-1} yields a finite set $\mathcal{A}' \subseteq \mathcal{S}'$ such that $\mathcal{S}' - \mathcal{A}' \subseteq \mathcal{S}$. We then let $\mathcal{S}_0 = \mathcal{A} \cup (\mathcal{S} \cap \mathcal{A}')$ and $\mathcal{S}'_0 = \mathcal{A}' \cup (\mathcal{S}' \cap \mathcal{A})$. $\square$

Define an equivalence relation on the set of infinite subsets of $\mathcal{Y}$ by setting $\mathcal{S} \sim \mathcal{S}'$ if $\mathcal{S} - \mathcal{S}_0 = \mathcal{S}' - \mathcal{S}'_0$ as in the lemma. Each equivalence class has only countably many elements, and so there are uncountably many equivalence classes. It follows that there are uncountably many non-homeomorphic V and hence uncountably many non-homeomorphic W. $\square$

Received June 30, 2003

Final version received March 23, 2004

Department of Mathematics
Oklahoma State University
Stillwater
OK 74078
USA
myersr@math.okstate.edu

References

[1] M. G. Brin and T. L. Thickstun, *Open, irreducible 3-manifolds which are end 1-movable,* Topology **26** (2), (1987), 211–233.

[2] A. V. Chernavskii, *Local contractibility of the group of homeomorphisms of a manifold,* (Russian) Mat. Sb. (N.S.) **79** (121) (1969), 307–356.

[3] R. D. Edwards and R. C. Kirby, *Deformations of spaces of imbeddings*, Ann. Math. (2) **93** (1971), 63–88.

[4] W. Haken, *Some results on surfaces in 3-manifolds*, Studies in Modern Topology, (1968), 39–98, Math. Assoc. Amer. (distributed by Prentice-Hall, Englewood Cliffs, N.J.)

[5] J. Hempel, 3-Manifolds, Ann. of Math. Stud. **86**, Princeton Univ. Press, Princeton, NJ, 1976.

[6] W. Jaco, Lectures on three-manifold topology, CBMS Regional Conference Series in Math., **43**, Amer. Math. Soc. 1980.

[7] W. Jaco and P. Shalen, *Seifert fibered spaces in 3-manifolds*, Memoirs Amer. Math. Soc. **21**, No. 220 (1979).

[8] K. Johannson, *Homotopy equivalences of 3-manifolds with boundaries*, Lecture Notes in Math. **761**, Springer-Verlag (1979).

[9] S. Kinoshita, *On infinite cyclic actions on contractible open 3-manifolds and strong irreducibility*, Proceedings of the Second Conference on Compact Transformation Groups (Univ. Mass., Amherst, Mass., 1971), Part II, 323–327. Lecture Notes in Math., **299**, Springer, Berlin 1972.

[10] R. Myers, *Excellent 1-manifolds in compact 3-manifolds*, Topology Appl. **49** (1993), 115–127.

[11] R. Myers, *Attaching boundary planes to irreducible open 3-manifolds*, Quart. J. Math. Oxford (2), **48** (1997), 363–404.

[12] R. Myers, *End reductions and covering translations of contractible open 3-manifolds*, Math ArXiv preprint math.GT/0209397 available at http://front.math.ucdavis.edu.

[13] P. Scott and T. Tucker, *Some examples of exotic noncompact 3-manifolds*, Quart. J. Math. Oxford Ser. (2) **40** (160), (1989), 481–499.

[14] T. Tucker, *On the Fox-Artin sphere and surfaces in noncompact 3-manifolds*, Quart. J. Math. Oxford Ser. (2) **28** (110), (1977), 243-253.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

A NOTE ON 2-UNIVERSAL LINKS

VÍCTOR NÚÑEZ

ABSTRACT. We show that no Montesinos knot (link) can be 2-universal.

1. Introduction

The main theorem of this note, that no Montesinos knot can be 2-universal (Corollary (3.1)), contrasts with the existence of 2-universal knots as shown in [3]. These two combined results are somewhat surprising, for most known universal knots (links) are Montesinos'.

Our main result follows easily from a result about factorization of branched coverings through cyclic coverings (Lemma (2.2)), which is interesting in its own and very useful.

Also we obtain a result on non simply connectednes of 'regular-like' branched coverings (Corollary (3.2)), as another application of Lemma (2.2).

2. Branched coverings through cyclic coverings

An *m-fold branched covering* $\varphi : M^3 \to N^3$ is a proper open map between 3-manifolds such that there is a 1-subcomplex $k \subset N$ (the *branching* of φ) with $\varphi| : M - \varphi^{-1}(k) \to N - k$ a finite m-fold covering space. For the purposes of this paper, $k \subset N$ will be a properly embedded submanifold; that is, k is a link in N. We say that 'φ is *branched along* k', and write $\varphi : M \to (N, k)$.

Given a component $\tilde{k} \subset \varphi^{-1}(k) \subset M$, the homological local degree $deg(\varphi, x)$ is the same for all $x \in \tilde{k}$; this common number is called the *ramification index* of $\tilde{k}$.

A *meridian of a component* $k_1 \subset k \subset N$ is a class $\mu \in \pi_1(N - k)$ that can be represented as $\mu = [a * m * \bar{a}]$, where m is the boundary of a disk D such that $D \cap k = \mathrm{Int}\,(D) \cap k_1 =$ one point, and a is an arc in $N - k$ connecting the base point with a point of m. Notice that meridians of the same component are conjugate. A *meridian of* k is a meridian of a component of k.

An m-fold branched covering $\varphi : M \to (N, k)$ determines (and is determined by) a representation $\omega_\varphi : \pi_1(N - k) \to S_m$ into the symmetric group on m symbols S_m. If $\omega_\varphi(\mu)$ is a product of disjoint cycles of order $c_1, c_2, \ldots$ for μ a meridian of a component k_1 of k, then the components of the preimage $\varphi^{-1}(k_1)$ have ramification indices $c_1, c_2, \ldots$. We say that φ is a *branched covering of index dividing* n, if $\omega_\varphi(\mu)^n$ is the identity permutation for all meridians μ of k.

2000 *Mathematics Subject Classification*: Primary 57M.
Keywords and phrases: Universal link, 3-manifold.
The author was supported in part by a CONACyT grant .

Let $k \subset S^3$ be a link; let us denote by $BC(n;k)$ the set of closed, connected, orientable 3-manifolds M such that there exists a branched covering $\varphi : M \to (S^3, k)$ of index dividing n. The link k is called *n-universal* if $BC(n;k)$ coincides with the set of all closed, connected, orientable 3-manifolds. It is common to call *universal link* a 0-universal link.

We let $p : B_n(k) \to (S^3, k)$ be the n-fold cyclic covering branched along all components of k; that is, the induced representation ω_p sends each meridian of k to an n-cycle in $Z_n \leq S_n$. The following lemma helps to organize the details in the proof of Lemma (2.2), and is proved for knots in [4], § 4 of Ch. 2. The proof is essentially the same for links, and we include it here for completeness.

Lemma (2.1) ([4]). *Let $k \subset S^3$ be a link, and write $\langle \mu^n \rangle_\pi$ for the normal closure of $\{\mu^n : \mu$ is a meridian of $k\}$ in $\pi_1(S^3 - k)$. Then $\pi_1(S^3 - k)/\langle \mu^n \rangle_\pi$ is a semi-direct product*

$$\frac{\pi_1(S^3 - k)}{\langle \mu^n \rangle_\pi} \cong Z_n \ltimes \pi_1(B_n(k))$$

where the generator of Z_n is the class of any meridian of k and acts on $\pi_1(B_n(k))$ as the isomorphism induced by an order n symmetry of $B_n(k)$ with quotient (S^3, k).

Proof. Let $k \subset S^3$ be a link of c components and let $H \leq \pi_1(S^3 - k)$ be the kernel of the composition $\pi_1(S^3 - k) \xrightarrow{Ab} H_1(S^3 - k) \cong Z^c \xrightarrow{\varepsilon} Z \xrightarrow{\rho} Z_n$, where Ab is the abelianization map, ε is the augmentation $\varepsilon(x_i)_i = \sum_i x_i$, and ρ is reduction (mod n). Notice that $H_1(S^3 - k) \cong Z^c$ has a basis of meridians $\mu_1, \ldots, \mu_c$, one for each component of k. We have that $H \cong p_\# \pi_1(B_n(k) - p^{-1}(k))$ where $p : B_n(k) \to (S^3, k)$ is the n-fold cyclic covering branched along all components of k. If μ is a meridian of k, then $p^{-1}(\mu)$ is a closed curve which represents, up to conjugation, the element $\mu^n \in H$, and we obtain the fundamental group of $B_n(k)$ adding the 'branching relations', $\pi_1(B_n(k)) \cong H/\langle \mu^n \rangle_H$, where $\langle \mu^n \rangle_H$ is the normal closure in H of $\{\mu^n : \mu$ is a meridian of $k\}$. Notice that $\langle \mu^n \rangle_H = \langle \mu^n \rangle_\pi$, for $\nu^{-1}\mu^n\nu = (\nu^{-1}\mu\nu)^n$ is the n-th power of a meridian, for each pair μ, ν of meridians of k. Therefore, the sequence

$$1 \to \frac{H}{\langle \mu^n \rangle_H} \to \frac{\pi_1(S^3 - k)}{\langle \mu^n \rangle_\pi} \xrightarrow{\xi} \frac{\pi_1(S^3 - k)}{H} \cong Z_n \to 1$$

is exact. The map ξ has a section $\pi_1(S^3 - k)/H \to \pi_1(S^3 - k)/\langle \mu^n \rangle_\pi$, and therefore $\pi_1(S^3 - k)/\langle \mu^n \rangle_\pi \cong Z_n \ltimes \pi_1(B_n(k))$, where the generator $\bar{\mu}$ of Z_n acts on $\pi_1(B_n(k))$ as the isomorphism induced by an order n homeomorphism of $B_n(k)$ with quotient (S^3, k). □

Lemma (2.2). *Let $k \subset S^3$ be a link, and let $\varphi : M \to (S^3, k)$ be an m-fold branched covering of index dividing n. Then there exists a commutative square of branched coverings*

$$\begin{tikzcd} & \tilde{M} \arrow[dl, "q"'] \arrow[dr, "\psi"] & \\ M \arrow[dr, "\varphi"'] & & B_n(k) \arrow[dl, "p"] \\ & (S^3, k) & \end{tikzcd}$$

where p is the n-fold cyclic covering of (S^3, k) branched along all components of k, ψ is an m-fold (unbranched) covering space, and q is an n-fold covering branched along the components of $\varphi^{-1}(k) \subset M$ with ramification index less than n.

Proof. Let $\omega : \pi_1(S^3 - k) \to S_m$ be the representation determined by the covering $\varphi : M \to (S^3, k)$. The covering subgroup of φ is $U = \omega^{-1}(St(1)) \cong \varphi_{\#}\pi_1(M - \varphi^{-1}(k))$ where $St(1) \leq S_m$ is the subgroup of permutations fixing the symbol 1. Since $\omega(\mu)^n$ is the identity permutation for each meridian μ of k, the representation ω factors:

$$\begin{tikzcd} \pi_1(S^3 - k) \arrow[rr, "\omega"] \arrow[dr] & & S_m \\ & \pi_1(S^3 - k)/\langle \mu^n \rangle_\pi \arrow[ur, "\bar{\omega}"'] & \end{tikzcd}$$

From the previous lemma we know $\pi_1(S^3 - k)/\langle \mu^n \rangle_\pi \cong Z_n \ltimes \pi_1(B_n(k))$, and, by restriction, we get $\tau = \bar{\omega}| : \pi_1(B_n(k)) \to S_m$, a representation which is perhaps not transitive. This τ induces an m-fold (unbranched) covering space $\psi : \tilde{M} \to B_n(k)$ such that $\tilde{M}$ is connected if and only if τ is transitive. The covering subgroup of ψ is $\bar{U} = \tau^{-1}(St(1)) = \pi_1(B_n(k)) \cap \bar{\omega}^{-1}(St(1)) = (H \cap U)/\langle \mu^n \rangle_H \cong \psi_{\#}\pi_1(\tilde{M})$, if $\tilde{M}$ is connected. As in the proof of the previous lemma, $H \cong \pi_1(B_n(k) - p^{-1}(k))$. We then see that $U \cap H \cong p_{\#}\psi_{\#}\pi_1(\tilde{M} - \psi^{-1}(p^{-1}(k)))$. Therefore $\tilde{M}$ is the pullback of φ and p as in [2], and the lemma follows. If $\tilde{M}$ is not connected, we perform the same analysis on subgroups for each component K of $\tilde{M}$; that is, we analyze $\psi| : K \to (S^3, k)$ for each component K and obtain that $\tilde{M}$ is again a pullback, and the lemma follows. $\square$

Remark. The previous lemma and its proof show that getting an m-fold covering $\varphi : M \to (S^3, k)$ of index dividing n is the same as finding a special representation $\pi_1(B_n(k)) \to S_m$. This point of view is exploited in [6] to construct 'dihedral-like' coverings of Montesinos knots. We thank the referee for pointing out that the construction of Lemma (2.2) is a standard pullback.

3. Branched coverings of fixed index

Let $k \subset S^3$ be a Montesinos link. Then $B_2(k)$ is an orientable Seifert manifold with orbit surface the 2-sphere, $(O, 0; \beta_1/\alpha_1, \ldots, \beta_t/\alpha_t)$, or an orientable Seifert manifold with orbit surface a non-orientable surface of (non-orientable) genus g, $(O, -g; \beta_1/\alpha_1, \ldots, \beta_t/\alpha_t)$. See [5].

COROLLARY (3.1). *If k is a Montesinos link, then k is not 2-universal.*

Proof. If $\varphi : M \to (S^3, k)$ is an m-fold branched covering of index dividing 2, then from Lemma (2.2) we obtain $\psi : \tilde{M} \to B_2(k)$ an m-fold (unbranched) covering space, and $q : \tilde{M} \to M$ a 2-fold branched covering. Since $B_2(k)$ is a Seifert manifold, we see that $\tilde{M}$ is also a Seifert manifold. Since q is 2-fold, q is a regular covering; therefore there exists an involution of $\tilde{M}$ with quotient M. We conclude that M is a Seifert orbifold ([1]), and that $BC(2; k)$ is not the set of all closed, connected, orientable 3-manifolds. Therefore k is not 2-universal. $\square$

REMARK. In particular, from the previous corollary, we see that: A hyperbolic 2-bridge knot, which is known to be 12-universal, cannot be 2-universal; the Borromean rings, known to be 4-universal, are not 2-universal.

COROLLARY (3.2). *Let $k \subset S^3$ be a link such that order of $\pi_1(B_n(k))$, possibly infinite, does not divide m. Let $\varphi : M \to (S^3, k)$ be an m-fold branched covering with induced representation $\omega : \pi_1(S^3 - k) \to S_m$ such that $\omega(\mu)$ is a product of disjoint n-cycles for each meridian μ of k. Then M is not simply connected.*

Proof. From Lemma (2.2) we obtain $\psi : \tilde{M} \to B_n(k)$ an m-fold (unbranched) covering space, and $q : \tilde{M} \to M$ an n-fold covering. By hypothesis, there are no components of $\varphi^{-1}(k) \subset M$ with ramification index less than n; therefore q is a covering (unbranched) space, and $q_\# : \pi_1(K) \to \pi_1(M)$ is an embedding for each component K of $\tilde{M}$. If $\pi_1(B_n(k))$ is infinite, each component of $\tilde{M}$ has infinite fundamental group and the corollary follows. If $\pi_1(B_n(k))$ is finite then, since its order does not divide m, at least one component of $\tilde{M}$ is not simply connected, for the index of $\pi_1(K)$ in $\pi_1(B_n(k))$ is a divisor of m for each component K of $\tilde{M}$; the corollary follows. $\square$

Received July 15, 2003

Final version received March 26, 2004

CIMAT
APDO. POSTAL 402
36000 GUANAJUATO, GTO.
MÉXICO
victor@cimat.mx

REFERENCES

[1] F. BONAHON AND L. SIEBENMANN, The classification of Seifert fibred 3-orbifolds. Low-dimensional topology (Chelwood Gate, 1982), 19–85, London Math. Soc. Lecture Notes Ser. **95**, Cambridge Univ. Press, Cambridge, 1985.

[2] J. HEMPEL, *Branched coves over strongly amphicheiral links*, Topology **29** (1990), 247–255.

[3] H. M. HILDEN, M. T. LOZANO AND J. M. MONTESINOS, *On 2-universal knots*, Bol. Soc. Mat. Mexicana (3) **10** (2004).

[4] F. GONZÁLEZ-ACUÑA, On homology spheres. Ph. D. Thesis. Princeton University, 1969.

[5] J. M. MONTESINOS, *Variedades de Seifert que son recubridores cíclicos ramificados de dos hojas*, Bol. Soc. Mat. Mexicana (2) **18** (1973), 1–32.

[6] V. NÚÑEZ AND J. RODRÍGUEZ–VIORATO, *Dihedral coverings of Montesinos knots*, Bol. Soc. Mat. Mexicana (3) **10**, (2004).

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

DIHEDRAL COVERINGS OF MONTESINOS KNOTS

VÍCTOR NÚÑEZ AND JESÚS RODRÍGUEZ–VIORATO

Abstract. We determine the family of Montesinos knots k which have the 3-sphere as a dihedral quotient $S^3 \to (S^3, k)$, and we find also 'dihedral-like' coverings of certain Montesinos knots. Through the understanding of the singular set of these coverings, we conclude the universality of many Montesinos knots.

1. Introduction

Dihedral coverings. Since the time of Reidemeister (the first part of the 20th century), it has been an interesting problem to understand the dihedral branched coverings of knots (§ 2), especially the dihedral coverings of the 3-sphere over itself. Fox's Quick Trip and [1] describe coverings of this type for 2-bridge knots; there are some examples of this kind of coverings for only a few more knots spread in other works on the subject.

We determine the family of Montesinos knots k, in terms of its invariants, which admit a dihedral quotient $S^3 \to (S^3, k)$ (Theorem (3.2)); we get this result by using the knowledge on cyclic coverings of Seifert manifolds developed in [12].

It is a very difficult problem to describe in an intelligible way what is the type of the link in the preimage of a knot k under a branched covering $S^3 \to (S^3, k)$ (see, for example, the heroic struggles in [6] and [16]). With our new understanding of dihedral quotients of Montesinos knots, we are able to describe the preimage of the knot in such a covering as a union of Montesinos knots (Montesinos knots again!), and, in some cases, we are able to compute explicitly the invariants of the knots. In general we give an algorithm for such a task (§ 5).

We extend the construction of dihedral quotient to 'dihedral-like' covering (§ 6). With these new coverings we construct, for certain Montesinos knots k, branched coverings $S^3 \to (S^3, k)$; we are able, also in these cases, to give explicit Montesinos invariants for components of the preimage of k.

Universals. In 1982 Thurston proved the striking fact that there are universal links ([14]). A link $k \subset S^3$ is called *universal* if each closed, connected, orientable 3-manifold is a branched covering over (S^3, k). A very interesting problem is to describe the family of universal links in the 3-sphere. Towards this goal the following theorems are known.

Theorem (1.1) ([6]). *All hyperbolic 2-bridge links are universal.*

2000 *Mathematics Subject Classification*: Primary 57M.

Keywords and phrases: dihedral covering, universal link, Montesinos knot, Seifert manifold.

The authors were supported in part by a CONACyT grant .

The 2-bridge link $\ell(b/a)$ is the Montesinos knot $m(b/a)$. A 2-bridge link $\ell(b/a)$ is hyperbolic if and only if $a \not\equiv \pm 1 \bmod b$.

The pretzel link $p(b; \alpha_1, \ldots, \alpha_t)$ is the Montesinos link $m(b/1, 1/\alpha_1, \ldots, 1/\alpha_t)$. An *Uchida link* is a pretzel link $p(b; \alpha_1, \ldots, \alpha_t)$ such that at least two α_i's are even.

THEOREM (1.2) ([16]). *All Uchida links are universal except for:* $p(2s, 2t)$, $p(2, -2, s)$, *with* $s, t \in Z - \{0\}$; $p(0; \pm 2, \pm 3, \mp 4)$, $p(0; \pm 3, \pm 6, \mp 2)$, $p(0; \pm 4, \pm 4, \mp 2)$, *and* $p(0; 2, 2, -2, -2)$.

THEOREM (1.3) ([13]). *If* $|n| > 1$ *is odd, then* $p(n, n, -n)$ *is universal.*
If $n \neq -2, 0$ *is even, then* $p(3, 3, n)$ *is universal.*

Once one knows that certain knots (links) are universal, to decide if a given link k is universal, one constructs branched coverings $S^3 \to (S^3, k)$ and tries to find an already known universal link in the preimage of k. With our understanding of preimages of links in dihedral and dihedral-like coverings, we give lists of Montesinos knots which are universal (§§ 6 and 7). For this end we rely heavily on the previously stated theorems.

The paper is organized as follows. In § 2 we specify precisely the notion of dihedral quotient and recall some tools to handle Seifert manifolds and Montesinos knots. In § 3 we prove the main theorem of this work (Theorem (3.2)), which is an essential ingredient to get a deeper understanding of dihedral coverings. In § 4 we give a convenient formalization of the theory of rational tangles. In § 5 we plunge into a description of branch sets of dihedral quotients. Next we extend the notion of dihedral quotient to 'dihedral-like' quotient and get some universality results in § 6. Finally in § 7 we harvest several results on universality of families of links.

2. Dihedral quotients

A *branched covering* $\varphi : M^n \to N^n$ is a proper open map between n-manifolds such that there is a codimension 2 subcomplex $K \subset N$ with $\varphi| : M - \varphi^{-1}(K) \to N - K$ a finite covering space. For the purposes of this paper, $K \subset N$ will be a properly embedded submanifold.

The map $\varphi| : M - \varphi^{-1}(K) \to N - K$ is called *the associated covering space* of φ. The submanifold K is called the *branching* of φ. We say that 'φ is *branched along* K', 'φ is a *branched covering over* (N, K)', and that 'M is a *branched covering over* (N, K)', and write $\varphi : M \to (N, K)$.

The set $\varphi^{-1}(K)$ is called the *singular set* of φ. The *pseudo-branch* of φ is the set of points $x \in \varphi^{-1}(K)$ such that φ is a homeomorphism at x. The *branch set* of φ is the complement, in $\varphi^{-1}(K)$, of the pseudo-branch of φ.

A branched covering $\varphi : M \to (N, k)$ determines (and is determined by) a representation $\omega_\varphi : \pi_1(N - k) \to S_n$ into the symmetric group S_n. It is customary to name a branched covering after the nature of the image $Im(\omega_\varphi)$. If $Im(\omega_\varphi)$ is a cyclic group, then φ is called a *cyclic* covering; if $Im(\omega_\varphi)$ is a dihedral group, then φ is called a *dihedral* covering, etc.

If $B_2(k)$ is the cyclic double branched covering over (S^3, k) branched along all components of k, and $\rho : H_1(B_2(k)) \to Z_n$ is an epimorphism, the covering

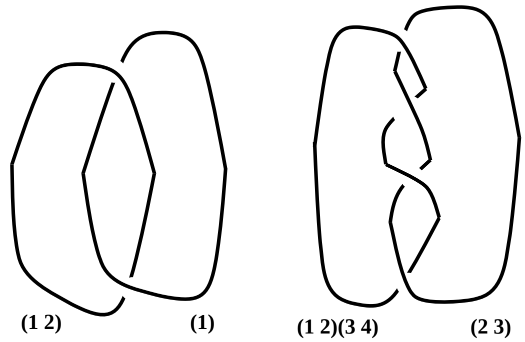


Figure 1. Dihedral quotients, but not dihedral coverings.

space $\psi : \tilde{M} \to B_2(k)$ determined by ρ has an involution $\tilde{u} : \tilde{M} \to \tilde{M}$ such that the square

$$\begin{tikzcd} \tilde{M} \arrow[r, "\tilde{u}"] \arrow[d, "\psi"'] & \tilde{M} \arrow[d, "\psi"] \\ B_2(k) \arrow[r, "u"] & B_2(k) \end{tikzcd}$$

commutes, where $u : B_2(k) \to B_2(k)$ is the involution with quotient (S^3, k). Therefore one has a commutative square

$$\begin{tikzcd} & \tilde{M} \arrow[dl, "q"'] \arrow[dr, "\psi"] & \\ M \arrow[dr, "\varphi"'] & & B_2(k) \arrow[dl, "p"] \\ & (S^3, k) & \end{tikzcd}$$

where $M = \tilde{M}/\tilde{u}$. In Fox's Quick Trip, the map φ would be called a dihedral covering, for it would correspond to a dihedral representation. Actually Fox starts a discussion, and gives a procedure to obtain metacyclic representations of knot groups in [2], pp. 160 and ff. This procedure, translated into the 'language of commutative squares', is the one presented here in the particular case of dihedral representations. But if k is not connected, or if n is even, it might be the case that $Im(\omega_\varphi) \leq S_n$ is not a dihedral group. See Figure (1).

We call $\varphi : M \to (S^3, k)$ a *dihedral quotient.* Notice that the branching of $q : \tilde{M} \to M$ is also the pseudo-branch of φ. For other works using the language of commutative squares for this kind of coverings, see [9] and [1].

(2.1) Standard double coverings of Seifert Manifolds. The *Seifert manifold* M with *Seifert symbol* $M = (O, 0; \beta_1/\alpha_1, \ldots, \beta_t/\alpha_t)$, where $(\alpha_i, \beta_i) = 1$ and $\alpha_i \neq 0$ $(i = 1, \ldots, t)$, is constructed as follows:

Let $p_1, \ldots, p_t \in S^2$ and $S^2_0 = \overline{S^2 - \cup \mathcal{N}(p_i)}$; we obtain M as the union of $S^2_0 \times S^1$ with t solid tori $V_1, \ldots, V_t$ along the boundaries. Call $q_1, \ldots, q_t$ the boundary components of S^2_0, and h the slice $\{x\} \times S^1$ for some $x \in S^2_0$, and let $g_i : \partial V_i \to q_i \times S^1$ be a homeomorphism such that $g_i(m_i) \sim q_i^{\alpha_i} h^{\beta_i}$ for m_i a meridian of V_i, $i = 1, \ldots, t$. Then $M = (S^2_0 \times S^1) \bigcup_{\cup g_i} (\cup_i V_i)$.

The core e_i of V_i is called the *fiber of M corresponding to the ratio β_i/α_i*. In case $|\alpha_i| > 1$, e_i is called *exceptional*; otherwise e_i and the curves $h = \{x\} \times S^1$ are called *ordinary fibers of M*.

Let $M = (O, 0; \beta_1/\alpha_1, \ldots, \beta_t/\alpha_t)$; there is an involution $u : M \to M$, described below, associated to the Seifert symbol $(O, 0; \beta_1/\alpha_1, \ldots, \beta_t/\alpha_t)$ (not to the manifold M), which gives rise to a 2-fold branched covering $p : M \to (S^3, k)$ branched along all components of k; the link k is called *the Montesinos knot* with symbol $k = m(\beta_1/\alpha_1, \ldots, \beta_t/\alpha_t)$ (see [8]).

Let E be a great circle in S^2 which contains the points $p_1, \ldots, p_t$ in that order with respect to a given orientation of E; then the reflection on S^2 along E induces a $180°$ rotation $u_0 : S^2_0 \times S^1 \to S^2_0 \times S^1$ along the 'axis' $E_0 = (E \cap S^2_0) \times \{1\} \cup (E \cap S^2_0) \times \{-1\}$. u_0 extends to each solid torus V_i in M with two arcs as its fixed point set in V_i, giving the involution $u : M \to M$. In the quotient $M/u \cong S^3$ we have the picture of the branching of p in t 3-balls, $B_1 = p(V_1), \ldots, B_t = p(V_t)$, joined by a 'trivial closed braid'. See Figure (2).

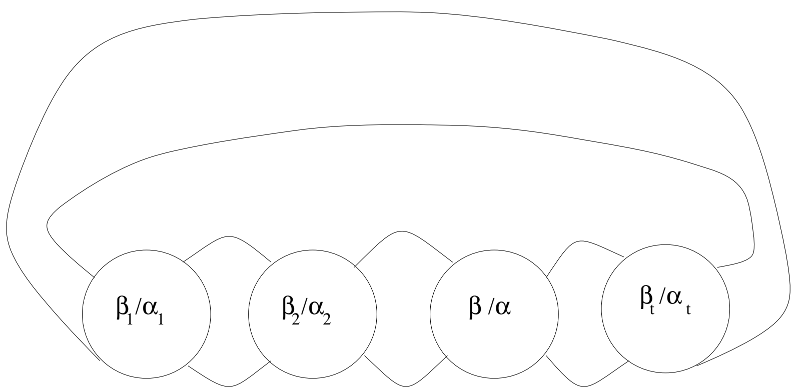


Figure 2

This branching, in each 3-ball, gives rise to a rational tangle, $\beta_i/\alpha_i = (B_i, a^i, m^i \cup \ell^i)$, with $a^i = p(fix(u) \cap V_i)$, $m^i = p(q_i \times \{1\}) \cup p(q_i \times \{-1\}) = m_0 \cup m_1$, and $\ell^i = p(\{x_0\} \times S^1) \cup p(\{x_1\} \times S^1) = \ell_0 \cup \ell_1$, where x_0, x_1 are the points in $E_0 \cap q_i$ in that order with respect to the orientation of E_0 (see § 4 and Figure (3)).

We call the covering $u : M \to M$ the *Standard involution* associated to the symbol $(O, 0; \beta_1/\alpha_1, \ldots, \beta_t/\alpha_t)$.

We will use, without warning, the next classification theorems.

Theorem (2.1.1) ([10]). *Two Seifert symbols represent homeomorphic Seifert manifolds with a fiber preserving homeomorphism if and only if one of the symbols can be transformed into the other with a finite sequence of the following moves:*

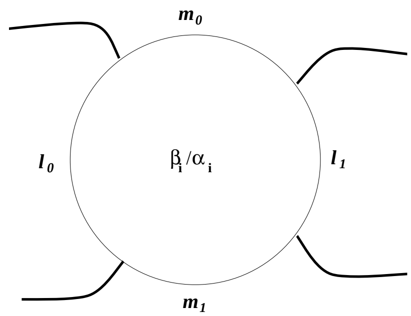


Figure 3

(0) *Permute the ratios*
(1) *Add or delete* 0/1
(2) *Replace the pair* $\beta_i/\alpha_i,\ \beta_j/\alpha_j$ *with either pair* $(\beta_i+k\alpha_i)/\alpha_i,\ (\beta_j-k\alpha_j)/\alpha_j$.

THEOREM (2.1.2) ([17]). *Two Montesinos symbols with at least three α's greater than 1 in absolute value, represent equivalent Montesinos knots if and only if one of the symbols can be transformed into the other with a finite sequence of the following moves:*
(0) *Permute the ratios cyclically*
(1) *Add or delete* 0/1
(2) *Replace the pair* $\beta_i/\alpha_i,\ \beta_j/\alpha_j$ *with either pair* $(\beta_i+k\alpha_i)/\alpha_i,\ (\beta_j-k\alpha_j)/\alpha_j$.

In a Montesinos knot $k = m(\beta_1/\alpha_1,\dots,\beta_t/\alpha_t)$ we allow that some of the $\alpha_i = 0$; but in such case k is a union of connected sums of 2-bridge links.

Notice that $m(\beta_1/\alpha_1, \beta_2/\alpha_2) \sim m(-(\alpha_1\beta_2+\alpha_2\beta_1)/(\alpha_2 r_1+\beta_2 s_1))$, the 2-bridge link $\ell(-(\alpha_1\beta_2+\alpha_2\beta_1)/(\alpha_2 r_1+\beta_2 s_1))$, where r_1 and s_1 are integers such that $\alpha_1 r_1 - \beta_1 s_1 = 1$.

3. Dihedral coverings of Montesinos knots

If $M = (O,0;\beta_1/\alpha_1,\dots,\beta_t/\alpha_t)$, then the fundamental group $\pi_1(M) = \pi_1(O,0;\beta_1/\alpha_1,\dots,\beta_t/\alpha_t) = \langle q_1,\dots,q_t,h : [q_i,h]=1, q_i^{\alpha_i}h^{\beta_i}=1, q_1\cdots q_t = 1\rangle$. We call the system $q_1,\dots,q_t,h$ a *standard system of generators* for the symbol $(O,0;\beta_1/\alpha_1,\dots,\beta_t/\alpha_t)$.

We want to understand the coverings of Seifert manifolds. The next lemma, which follows from ([12], Lemma 1 and its proof), will help. Let $\varepsilon \in S_n$ denote the standard n-cycle, $\varepsilon = (1,2,\dots,n)$.

LEMMA (3.1). *Let $\omega : \pi_1(O,0;\beta_1/\alpha_1,\dots,\beta_t/\alpha_t) \to S_n$ be the representation given by $\omega(h) = \varepsilon$, and $\omega(q_i) = \varepsilon^{r_i}$, $i = 1,\dots,t$, such that $\sum r_i = 0$. Then the covering space associated to ω is*

$$\varphi : (O,0;B_1/\alpha_1,\dots,B_t/\alpha_t) \to (O,0;\beta_1/\alpha_1,\dots,\beta_t/\alpha_t),$$

where $B_i = (\beta_i + r_i\alpha_i)/n$, for $i = 1,\dots,t$.

Also, if $\tilde e_i$ is the fiber corresponding to B_i/α_i, and e_i is the fiber corresponding to β_i/α_i, then $\varphi^{-1}(e_i) = \tilde e_i$, $i = 1,\dots,t$. At the fundamental group level, it holds that $\varphi(\tilde h) \sim h^n$, and $\varphi(\tilde q_i) \sim q_i h^{-r_i}$, for $\tilde q_1,\dots \tilde q_t,\tilde h$ a standard system *of generators of the covering. The map induced by φ on the orbit surface is a homeomorphism.*

Notice that, with the hypothesis of Lemma (3.1), $\beta_i + r_i\alpha_i \equiv 0 \bmod n$ ($i = 1,\dots,t$), for ω is a group homomorphism.

If k is the Montesinos knot $m(\beta_1/\alpha_1,\dots,\beta_t/\alpha_t)$, the number $\Delta(k) = \beta_1\alpha_2\cdots\alpha_t + \alpha_1\beta_2\cdots\alpha_t + \cdots + \alpha_1\alpha_2\cdots\beta_t$ is called the *determinant* of k. Let $B_2(k)$ be the double branched cover of (S^3,k); if $H_1(B_2(k))$ is infinite, then $\Delta(k) = 0$, and if $H_1(B_2(k))$ is finite, then $|H_1(B_2(k))| = |\Delta(k)|$.

Theorem (3.2). *Let k be the Montesinos knot $m(\beta_1/\alpha_1,\dots,\beta_t/\alpha_t)$, and let n be a positive divisor of $\Delta(k)$. If $(n,\alpha_i) = 1$ for $i = 1,\dots,t$, then there exists a dihedral quotient $\varphi : S^3 \to (S^3,k)$.*

Any n-fold dihedral quotient $\varphi : S^3 \to (S^3,k)$ has the Montesinos knot

$$m(b_1/\alpha_1,\dots,b_t/\alpha_t)$$

as pseudo-branch, where $b_i = (\beta_i + r_i\alpha_i)/n$, $i = 1,\dots,t$, and the integers $r_1,\dots,r_t$ satisfy the conditions:

$$\begin{cases} \beta_1 + r_1\alpha_1 \equiv 0 \quad \textit{mod } n \\ \qquad \vdots \\ \beta_t + r_t\alpha_t \equiv 0 \quad \textit{mod } n \\ \sum r_i = 0. \end{cases}$$

Proof. Assume $n|\Delta(k)$; then there is an epimorphism $\omega : H_1(B_2(k)) \to Z_n$; we regard $Z_n = \langle\varepsilon\rangle \le S_n$, where $\varepsilon = (1,2,\dots,n)$. Write $\omega(q_i) = \varepsilon^{r_i}$ and $\omega(h) = \varepsilon^s$.

We have $H_1(B_2(k)) = \langle q_1,\dots,q_t,h : q_i^{\alpha_i}h^{\beta_i} = 1, q_1\cdots q_t = 1,$ everything commutes $\rangle$. We can write $H_1(B_2(k)) = \langle x_2,\dots,x_t,v : x_i^{\delta_i} = 1, v^\lambda x_2^{\lambda_2}\cdots x_t^{\lambda_t} = 1,$ everything commutes $\rangle$, with $h = v^\ell$, $\ell = lcm\{\alpha_1,\dots,\alpha_t\}$, $\alpha'_1 = \alpha_1$, $\delta_1 = 1$, and $\delta_{i+1} = gcd\{\alpha_{i+1},\alpha'_1\cdots\alpha'_i\}$, $\alpha'_{i+1} = \alpha_{i+1}/\delta_{i+1}$ for $i \ge 1$ ([11], Lemma 2.2).

Assume now that $(\alpha_i,n) = 1$ for $i = 1,\dots,t$. If $\omega(x_i) = \varepsilon^{s_i}$ for some integer s_i, then $\omega(x_i^{\delta_i}) = \varepsilon^{\delta_i s_i} = (1)$; therefore $n|\delta_i s_i$. Since $\delta_i|\alpha_i$, by definition, we have that $(n,\delta_i) = 1$; then $n|s_i$ and, therefore, $\omega(x_i) = (1)$. Since ω is epimorphism, $\omega(v)$ is an n-cycle; then $\omega(h) = \omega(v)^\ell$ is an n-cycle, because $(n,\ell) = (n, lcm\{\alpha_1\cdots\alpha_t\}) = 1$. Since $\omega(h) = \varepsilon^s$, we conclude that $(n,s) = 1$. By conjugating in $Z_n \le S_n$, if necessary, we may assume that $\omega(h) = \varepsilon$. Since $\omega(q_1\cdots q_t) = \varepsilon^{\Sigma r_i} = (1)$, therefore $n|\sum r_i$, or $\sum r_i = m\cdot n$ for some m; replacing r_t by $r_t - m\cdot n$, we get $\sum r_i = 0$.

By Lemma (3.1), $\psi : \tilde M = (O,0;b_1/\alpha_1,\dots,b_t/\alpha_t) \to B_2(k)$ is the n-fold cyclic covering space associated to ω, with $b_i = (\beta_i + r_i\alpha_i)/n$; also $\psi(\tilde h) \sim h^n$ and $\psi(\tilde q_i) \sim q_i h^{-r_i}$. If $\tilde u : \tilde M \to \tilde M$ is the standard involution for the symbol $(O,0;b_1/\alpha_1,\dots,b_t/\alpha_t)$, and $u : B_2(k) \to B_2(k)$ is the standard involution with $B_2(k)/u = (S^3,k)$, then the square

$$\begin{tikzcd} \tilde M \arrow[r, "\tilde u"] \arrow[d, "\psi"'] & \tilde M \arrow[d, "\psi"] \\ B_2(k) \arrow[r, "u"'] & B_2(k) \end{tikzcd}$$

commutes, for $\psi \circ \tilde{u}(\tilde{h}) \sim \psi(\tilde{h}^{-1}) \sim h^{-n} \sim u(h^n) \sim u \circ \psi(\tilde{h})$, and $\psi \circ \tilde{u}(\tilde{q}_i) \sim \psi(\tilde{q}_i^{-1}) \sim q_i^{-1}h^{r_i} \sim u(q_i h^{-r_i}) \sim u \circ \psi(\tilde{q}_i)$. If $q : \tilde{M} \to \tilde{M}/\tilde{u}$ is the canonical projection, then q is a double branched covering over $(S^3, m(b_1/\alpha_1, \dots, b_t/\alpha_t)) = \tilde{M}/\tilde{u}$; we have the commutative diagram

$$\begin{array}{ccc} \tilde{M} & \xrightarrow{\psi} & B_2(k) \\ {\scriptstyle q}\downarrow & & \downarrow{\scriptstyle p} \\ M & \xrightarrow[\varphi]{} & (S^3, k) \end{array}$$

with $\varphi : S^3 \to (S^3, k)$ a dihedral quotient, and $m(b_1/\alpha_1, \dots, b_t/\alpha_t) \subset \varphi^{-1}(k)$ the pseudo-branch of φ. $\square$

(3.3) Remarks. Let $k = m(\beta_1/\alpha_1, \dots, \beta_t/\alpha_t)$, and let n be a divisor of $\Delta(k)$ such that $(n, \alpha_i) = 1$ for $i = 1, \dots, t$. Then $k = m(nb_1/\alpha_1, \dots, nb_t/\alpha_t)$ with $b_i = (\beta_i + r_i\alpha_i)/n$, and $\sum r_i = 0$.

3.3.1. The conclusion of Theorem (3.2) can be rephrased as: *If $n|\Delta(k)$, and $(n, \alpha_i) = 1$, then $m(b_1/\alpha_1, \dots, b_t/\alpha_t)$ is the pseudo-branch of an n-fold dihedral quotient $S^3 \to (S^3, m(nb_1/\alpha_1, \dots, nb_t/\alpha_t))$.*

3.3.2. Redrawing $k = m(nb_1/\alpha_1, \dots, nb_t/\alpha_t) = m(\beta_1/\alpha_1 + r_1, \dots, \beta_t/\alpha_t + r_t)$ one can easily visualize the representation ω in a projection of k: the meridians of each piece of one strand of the trivial 2-braid connecting the rational tangles of k (see Figure (2)) go to the same permutation under ω. Then using techniques similar to those in [6] one can draw at once the preimage of φ of Theorem (3.2) (see Figure (4) for the 5-fold dihedral covering of $m(1/3, 2/3, 2/3)$).

3.3.3. A possible partial converse of Theorem (3.2) is:

Let k be the Montesinos knot $m(\beta_1/\alpha_1, \dots, \beta_t/\alpha_t)$. Assume n is odd and $t \geq 3$. If there exists a dihedral quotient $S^3 \to (S^3, k)$, then $(n, \alpha_i) = 1$ for $i = 1, \dots, t$.

Notice that one cannot skip the hypothesis '$t \geq 3$ and n odd'.

For example, when $t \leq 2$, the 2-bridge knot

$$\ell = m(\beta_1/n\alpha_1, \beta_2/n\alpha_2) = m(-n(\alpha_1\beta_2 + \alpha_2\beta_1)/(n\alpha_2 r_1 + \beta_2 s_1)),$$

where $n\alpha_1 r_1 - \beta_1 s_1 = 1$, has an n-fold dihedral quotient $S^3 \to (S^3, \ell)$ with pseudo-branch $m(\beta_1/\alpha_1, \beta_2/\alpha_2)$.

When n is even, if $\ell = m(\beta_1/2\alpha_1, \beta_2/2\alpha_2, \beta_3/\alpha_3, \dots, \beta_t/\alpha_t)$, the representation $\omega : \pi_1(B_2(\ell)) \to S_2$ such that $\omega(q_1) = (1,2)$, $\omega(q_2) = (1,2)$, $\omega(q_i) = (1)$, for $i = 3, \dots, t$, and $\omega(h) = (1)$, gives rise to a 2-fold dihedral quotient $S^3 \to (S^3, \ell)$ with pseudo-branch $m(\beta_1/\alpha_1, \beta_2/\alpha_2, \beta_3/\alpha_3, \dots, \beta_t/\alpha_t)$ ([12], Lemma 2).

COROLLARY (3.3.4). *Let $k = m(\beta_1/\alpha_1, \dots, \beta_t/\alpha_t)$ be a universal link, and let n be a positive integer such that $(n, \alpha_i) = 1$ $(i = 1, \dots, n)$. Then*

$$m(n\beta_1/\alpha_1, \dots, n\beta_t/\alpha_t)$$

is a universal link.

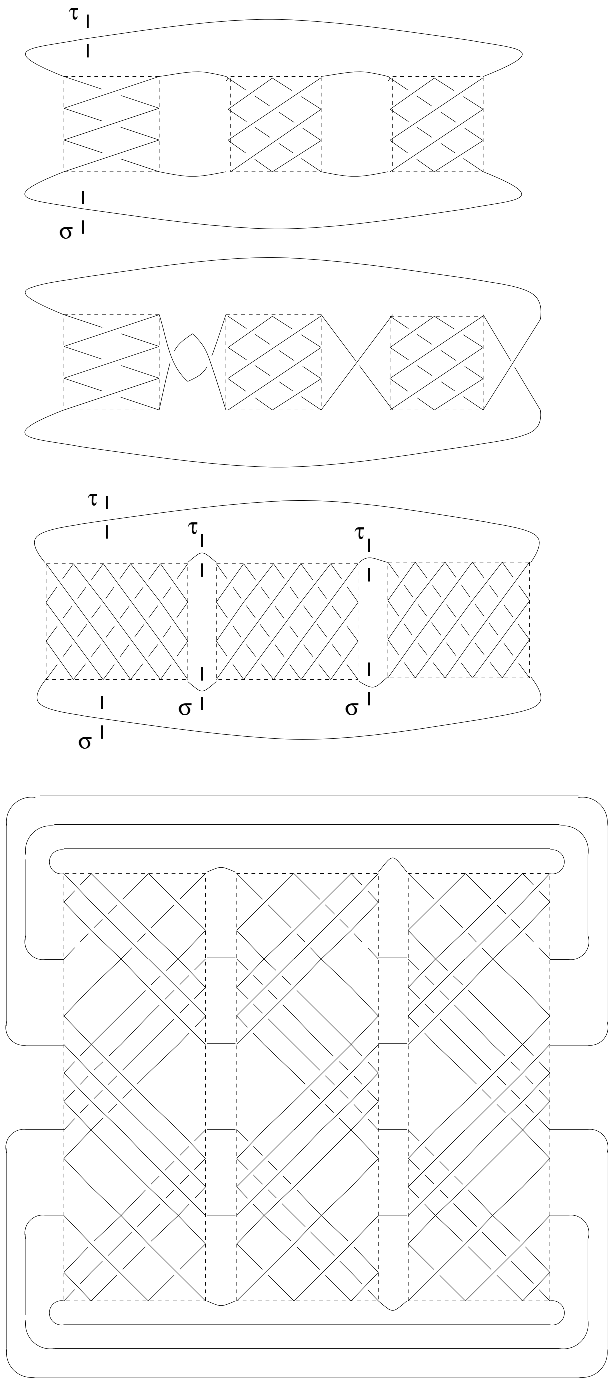


Figure 4

4. Rational tangles

A *rational tangle* is a triple $(B, a, m \cup \ell)$, where B is a 3-ball, a is a disjoint union of two properly embedded arcs in B, $a = a_0 \sqcup a_1$, and each of m and ℓ are a disjoint union of two arcs in ∂B, $m = m_0 \sqcup m_1$, $\ell = \ell_0 \sqcup \ell_1$, such that $m \cap \ell = \partial m = \partial \ell = \partial a$ and $m \cup \ell \cong S^1$, and such that the pair (B, a) is homeomorphic to $(D^2 \times I, \{x_1, x_2\} \times I)$ with $x_1, x_2 \in int(D^2)$. See Figure (5).

Two rational tangles are *equivalent* if they are homeomorphic as triples. A rational tangle $(B, a, m \cup \ell)$ contains a 'meridional disk' d_B properly embedded in B, which is the image of $\delta \times I \subset D^2 \times I$ under the homeomorphism $(B, a) \cong (D^2 \times I, \{x_1, x_2\} \times I)$ for $\delta \subset D^2$ a diameter separating x_1 and x_2. Notice that $\overline{B - \mathcal{N}(d_B)}$ is a regular neighborhood of a, and, therefore, a_i is the arc, unique up to isotopy fixing ∂B, which connects the ends of a_i in $B - d_B$ $(i = 0, 1)$; the isotopy class of ∂d_B in ∂B, with isotopies being the identity in ∂a, determines the isotopy class of d_B in $B - a$, and, therefore, *The curve ∂d_B determines the equivalence class of* $(B, a, m \cup \ell)$.

If $p : \tilde{B} \to B$ is the double branched cover of B branched along both a_0 and a_1, then $\tilde{m} = p^{-1}(m_0)$ and $\tilde{\ell} = p^{-1}(\ell_0)$ are a meridian-longitude pair in $\partial \tilde{B}$; if we fix an orientation on ∂B, an orientation of $\partial \tilde{B}$ is fixed; we orient $\tilde{m}$ and $\tilde{\ell}$ such that $\tilde{m} \cdot \tilde{\ell} = +1$. Now if $\tilde{d}$ is a lifting of ∂d_B in $\tilde{B}$, then $\tilde{d} \sim \tilde{\ell}^{\beta} \tilde{m}^{\alpha}$ in ∂B for some orientation of $\tilde{d}$; we associate the rational number β/α (or ∞, if $\alpha = 0$) to the tangle $(B, a, m \cup \ell)$. It is well known that the number β/α determines the equivalence class of $(B, a, m \cup \ell)$.

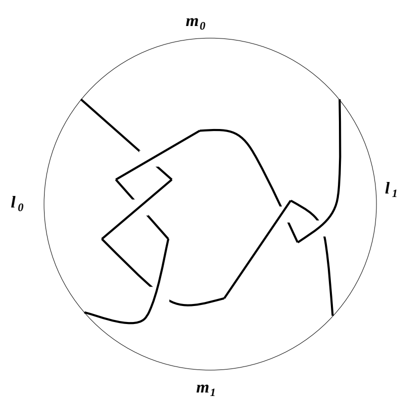


Figure 5

By pushing the arcs a into ∂B, it is possible to draw the rational tangle β/α on a square 'pillowcase' with lines of slope $\pm\beta/\alpha$ starting at the points $\{0, 1\} \times \{i/\alpha\}$ for $i = 0, 1, \ldots, \alpha$, and $\{j/\beta\} \times \{0, 1\}$ for $j = 0, 1, \ldots, \beta$ (Figure (6)).

In this square pillowcase the boundary of the meridional disk $d_B = d_{\beta/\alpha}$ is drawn "in between" the arcs a_0, a_1, that is, with lines of slope $\pm\beta/\alpha$ starting in the points $\{0, 1\} \times \{(2i - 1)/2\alpha\}$ for $i = 1, 2, \ldots, \alpha$, and $\{(2j - 1)/2\beta\} \times \{0, 1\}$ for $j = 1, 2, \ldots, \beta$ (Figure (7)).

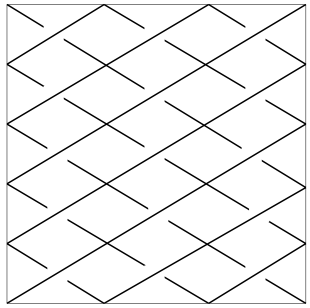


Figure 6. The rational tangle 3/5

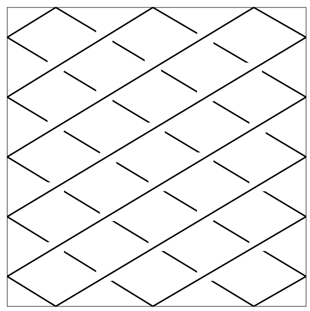


Figure 7. The disk 3/5

5. Branch sets in dihedral coverings

Let k be the Montesinos knot $k = m(\beta_1/\alpha_1, \dots, \beta_t/\alpha_t)$. Let n be a positive integer such that n divides the determinant of k, $n|\Delta(k)$, and $(n, \alpha_i) = 1$ for $i = 1, \dots, t$. Let $\varphi : S^3 \to (S^3, k)$ be the dihedral quotient guaranteed by Theorem (3.2). We will show that:

The preimage $\varphi^{-1}(k)$ is a disjoint union of Montesinos knots.

We write $B_2(k) = (O, 0; \beta_1/\alpha_1, \dots, \beta_t/\alpha_t)$, and $p : B_2(k) \to (S^3, k)$ the double covering branched along all components of k. Let $\psi : \tilde{M} \to B_2(k)$ be an n-fold cyclic covering space (unbranched), $\rho : \tilde{M} \to \tilde{M}$ the homeomorphism of order n such that $\tilde{M}/\rho = B_2(k)$, and $\psi : \tilde{M} \to \tilde{M}/\rho = B_2(k)$ the canonical projection.

Let $u : B_2(k) \to B_2(k)$ be the standard involution for $B_2(k)$ with respect to the symbol $(O, 0; \beta_1/\alpha_1, \dots, \beta_t/\alpha_t)$, and let $\tilde{u} : \tilde{M} \to \tilde{M}$ be the standard involution for $\tilde{M}$ with respect to the symbol $(O, 0; b_1/\alpha_1, \dots, b_t/\alpha_t)$, where, if $\tilde{e}_i$ is the exceptional fiber of $\tilde{M}$ corresponding to the ratio b_i/α_i, and e_i is the exceptional fiber of $B_2(k)$ corresponding to the ratio β_i/α_i, then $\psi^{-1}(e_i) = \tilde{e}_i$; we see that $k^2 = p^{-1}(k)$ is the axis of u, that is, $k^2 = Fix(u)$; let $\tilde{k}_0$ be the axis of $\tilde{u}$, $\tilde{k}_0 = Fix(\tilde{u})$; the preimage of k^2 is $\psi^{-1}(k^2) = \tilde{k}_0 \sqcup \tilde{k}_1 \sqcup \cdots \sqcup \tilde{k}_{n-1} = \tilde{k}_0 \sqcup \rho(\tilde{k}_0) \sqcup \cdots \sqcup \rho^{n-1}(\tilde{k}_0)$. Replacing ρ by ρ^m for some m coprime with n, if necessary, we may assume that, starting with $\tilde{k}_0$, we find successively $\tilde{k}_1, \tilde{k}_2, \dots, \tilde{k}_{n-1}, \tilde{k}_0, \tilde{k}_1, \tilde{k}_2, \dots, \tilde{k}_{n-1}$ as we traverse along $\tilde{h} \subset \tilde{M}$, for $\tilde{h}$ a preimage of an ordinary fiber h of $B_2(k)$ with $h \cap k^2 =$ two points. Notice that, then, $\tilde{k}_\ell = \tilde{u}(\tilde{k}_{n-\ell})$ for each ℓ.

We have $q : \tilde{M} \to \tilde{M}/\tilde{u} = (S^3, k_0)$ a double branched covering, and $\varphi : S^3 \to (S^3, k)$ a dihedral quotient. Then $k_0 = q(\tilde{k}_0)$ is the branching of q, and also the pseudo-branch of φ, and the diagram of coverings

$$\begin{tikzcd} & \tilde{M} \arrow[dl, "q"'] \arrow[dr, "\psi"] & \\ M \arrow[dr, "\varphi"'] & & B_2(k) \arrow[dl, "p"] \\ & (S^3, k) & \end{tikzcd}$$

commutes.

Now we define

$$\begin{aligned} k_1 &= q(\tilde{k}_1) = q(\tilde{k}_{n-1}) \\ k_2 &= q(\tilde{k}_2) = q(\tilde{k}_{n-2}) \\ &\vdots \\ k_{(n-1)/2} &= q(\tilde{k}_{(n-1)/2}) = q(\tilde{k}_{(n+1)/2}) \end{aligned}$$

if n is odd, and

$$\begin{aligned} k_1 &= q(\tilde{k}_1) = q(\tilde{k}_{n-1}) \\ k_2 &= q(\tilde{k}_2) = q(\tilde{k}_{n-2}) \\ &\vdots \\ k_{(n-2)/2} &= q(\tilde{k}_{(n-2)/2}) = q(\tilde{k}_{(n+2)/2}) \\ k_{n/2} &= q(\tilde{k}_{n/2}) \end{aligned}$$

if n is even.

Then $\varphi^{-1}(k) = k_0 \sqcup k_1 \sqcup k_2 \sqcup \cdots$. Notice that, if n is even, $k_{n/2}$ is the trivial knot (always).

We show that each k_ℓ is a Montesinos knot (see Figure (8)).

Let $i \in \{1, \ldots, t\}$; and let $j \in \{1, 2, \ldots, (n-1)/2\}$ if n is odd, or $j \in \{1, 2, \ldots, (n-2)/2\}$ if n is even.

In the solid torus $\tilde{V}_i \subset \tilde{M}$ there are two meridional disks $\tilde{D}_i, u(\tilde{D}_i)$, which separate the two arcs $\tilde{k}_j \cap \tilde{V}_i$; the image $D_i = q(\tilde{D}_i) = q(u(\tilde{D}_i))$ is a 2-disk properly embedded in the 3-ball $B_i = q(\tilde{V}_i)$ which separates the two arcs $q(\tilde{k}_j \cap \tilde{V}_i) = q(\tilde{k}_{n-j} \cap \tilde{V}_i)$ (see Figure (9)). From these arcs we get a rational tangle as follows:

Recall that, to construct $\tilde{M}$ (see § (2)), the torus $\partial\tilde{V}_i$ is glued to $\tilde{q}_i \times S^1$ in a punctured $\tilde{M}$ ($\cong S^2_0 \times S^1$). Let $\tilde{h}^0 = \{x_0\} \times S^1$ and $\tilde{h}^1 = \{x_1\} \times S^1$, where x_0 and x_1 are the points in $\tilde{k}_0 \cap (\tilde{q}_i \times \{1\})$. Starting at x_a and following the orientation of $\tilde{h}^a$ we get three subarcs: one subarc $\tilde{m}^a_1 \subset \tilde{h}^a$ connecting x_a with $\tilde{h}^a \cap \tilde{k}_j$, one subarc $\tilde{\ell}_a \subset \tilde{h}^a$ connecting $\tilde{h}^a \cap \tilde{k}_j$ with $\tilde{h}^a \cap \tilde{k}_{n-j}$, and one

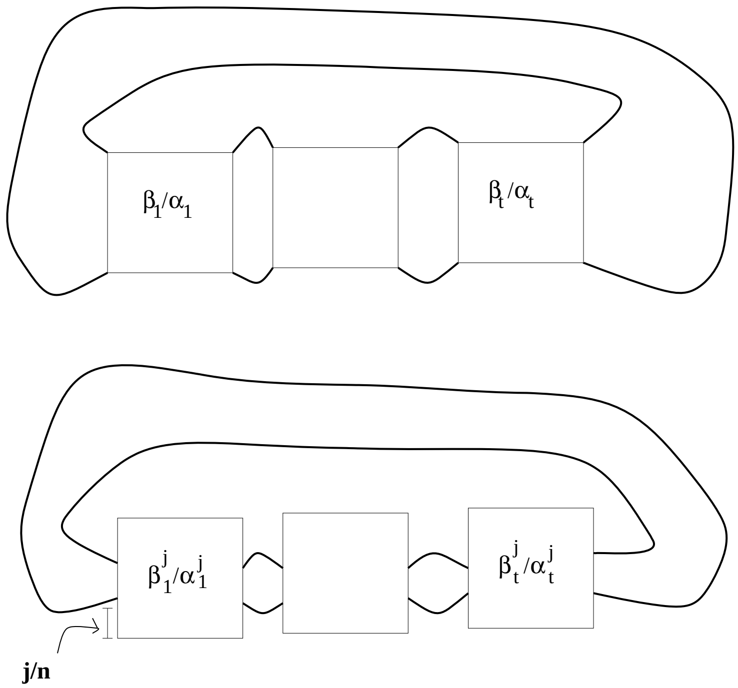


Figure 8

subarc $\tilde{m}_0^a \subset \tilde{h}^a$ connecting $\tilde{h}^a \cap \tilde{k}_j$ with $y_a \in \tilde{h}^a \cap \tilde{k}_0 - \{x_a\}$ $(a = 0, 1)$. Write $m_1 = q(\tilde{m}_1^0 \cup \tilde{q}_i \times \{1\} \cup \tilde{m}_1^1)$, $m_0 = q(\tilde{m}_0^0 \cup \tilde{q}_i \times \{y_0\} \cup \tilde{m}_0^1)$, $\ell_0 = q(\tilde{\ell}_0)$, and $\ell_1 = q(\tilde{\ell}_1)$; set $m^j = m_0 \sqcup m_1$ and $\ell^j = \ell_0 \sqcup \ell_1$; then

$$b_i^j/\alpha_i^j = (B_i, q(\tilde{k}_j \cap \tilde{V}_i), m^j \cup \ell^j)$$

is a rational tangle for some number $b_i^j/\alpha_i^j \in Q \cup \{\infty\}$, and it is defined by ∂D_i. Therefore

1. *For $j \in \{1, 2, \ldots, (n-1)/2\}$ if n is odd, or $j \in \{1, 2, \ldots, (n-2)/2\}$ if n is even, k_j is the Montesinos knot $m(b_1^j/\alpha_1^j, \ldots, b_t^j/\alpha_t^j)$.*

Now we will compute the ratios b_i^j/α_i^j.

Draw the ball B_i as a square pillowcase; then the knot k_j intersects ∂B_i in the points $\{0, 1\} \times \{j/n, 1 - j/n\}$.

We will compute the defining numbers $b_i^j/\alpha_i^j \in Q \cup \{\infty\}$ of a lifting d of ∂D_i in the double cover T of ∂B_i branched along $\partial B_i \cap k_j$. The contour of the square pillowcase, $m^j \cup \ell^j$, is composed of the arcs (see Figure (10)):

$$m_0 = (\{0\} \times [1 - j/n, 1]) \cup (I \times \{1\}) \cup (\{1\} \times [1 - j/n, 1]) = v_0 \cup \hat{m}_0 \cup w_0,$$

$$m_1 = (\{0\} \times [0, j/n]) \cup (I \times \{0\}) \cup (\{1\} \times [0, j/n]) = v_1 \cup \hat{m}_1 \cup w_1,$$

$$\ell_0 = \{0\} \times [j/n, 1 - j/n],$$

$$\ell_1 = \{1\} \times [j/n, 1 - j/n].$$

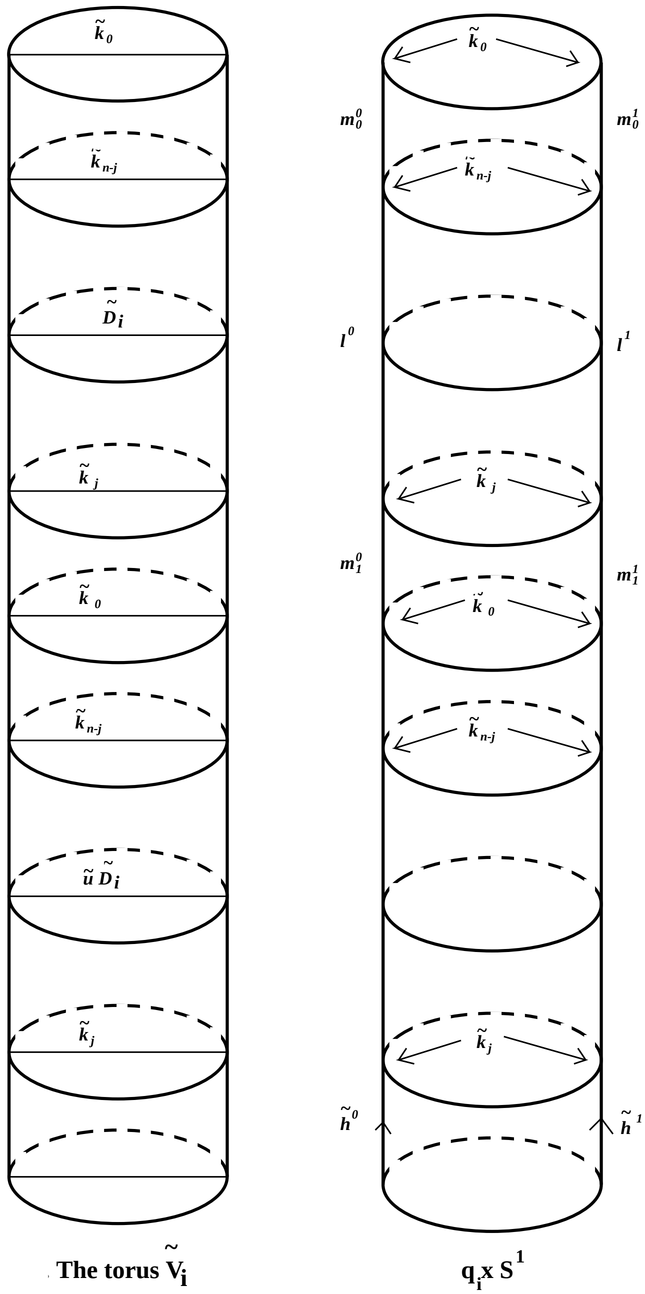


Figure 9. The solid torus $\tilde{V}_i$ and $q_i \times S^1$

Let V be the cylinder obtained by cutting ∂B_i along m^j; the boundary of V is the disjoint union of two circles $m_0^+ \cup m_0^-$ and $m_1^+ \cup m_1^-$ where $m_a^+$and m_a^- are copies of m_a in V $(a = 0, 1)$.

To construct T glue two copies, V_1, V_2, of V identifying $m^+_{0,1} \sim m^-_{0,2}$, $m^-_{0,1} \sim m^+_{0,2}$, $m^+_{1,1} \sim m^-_{1,2}$, and $m^-_{1,1} \sim m^+_{1,2}$, where $\partial V_a = (m^+_{0,a} \cup m^-_{0,a}) \cup (m^+_{1,a} \cup m^-_{1,a})$ $(a = 1, 2)$.

To compute the defining numbers b^j_i/α^j_i of a lifting $d \subset T$ of ∂D_i with respect to m_0 and ℓ_0, we will construct a new torus $\bar{T}$ and a curve $\bar{d}$ in $\bar{T}$ which will have the defining numbers b^j_i/α^j_i with respect to some basis of $\bar{T}$.

Glue V_1 and V_2 but only along the arcs $\hat{m}_0$ and $\hat{m}_1$, that is, identify $\hat{m}^+_{0,1} \sim \hat{m}^-_{0,2}$, etc., and obtain a four-punctured torus $\hat{T}$; if in $\hat{T}$ we identify the arcs v_0, w_0, v_1, w_1 as if they were not originally cut, that is, if we identify $v^+_{0,1} \sim v^-_{0,1}$, $w^+_{0,1} \sim w^-_{0,1}$, $v^+_{0,2} \sim v^-_{0,2}$, $w^+_{0,2} \sim w^-_{0,2}$, etc., we will obtain T^0, the double covering of ∂B_i branched along the corners of the pillowcase $\partial(\hat{m}_0 \cup \hat{m}_1)$. In T^0 a lifting of ∂D_i is the curve b_i/α_i, $\widetilde{\partial D_i} \sim \hat{m}_0^{\alpha_i} \hat{\ell}_0^{b_i}$, where $\hat{\ell}_0$ is the preimage in T^0 of $v_1 \cup \ell_0 \cup v_0$. We regard $\hat{T}$ as subset of T^0, and call $\hat{d} = \hat{T} \cap (\text{preimage of } \partial D_i \text{ in } T^0)$.

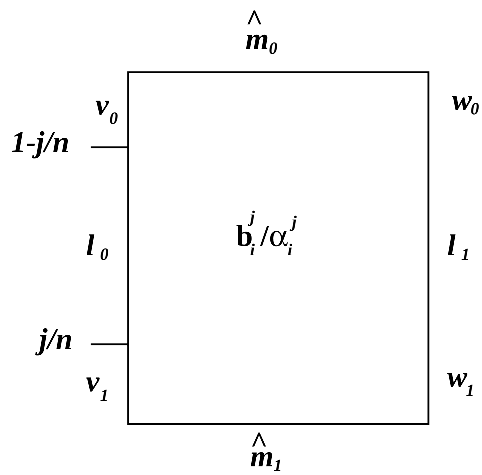


Figure 10. The ball B_i

To obtain T from $\hat{T}$ we still have to identify, as before, $v^+_{0,1} \sim v^-_{0,2}$, $w^+_{0,1} \sim w^-_{0,2}$, etc. Therefore we can construct the preimage of ∂D_i in T from $\hat{d}$ as follows: each point $p^+ \in v^+_{0,1} \cap \hat{d}$ is identified with a point $p^- \in v^-_{0,2} \cap \hat{d}$, where $p^+ \sim p^-$ under the identification $v^+_{0,1} \sim v^-_{0,2}$; each point $p^+ \in w^+_{0,1} \cap \hat{d}$ is identified with a point $p^- \in w^-_{0,2} \cap \hat{d}$, etc.

If we close $\hat{T}$ with four rhombi $R_1 = v^+_{0,1}v^-_{0,2}v^+_{0,2}v^-_{0,1}$, $R_2 = w^-_{0,1}w^+_{0,2}w^-_{0,2}w^+_{0,1}$, $R_3 = v^+_{1,2}v^-_{1,1}v^+_{1,1}v^-_{1,2}$, $R_4 = w^+_{1,2}w^-_{1,1}w^+_{1,1}w^-_{1,2}$, in the resulting torus $\bar{T}$, we close the arcs of $\hat{d}$ with vertical lines in the rhombi connecting equivalent points of the v's and w's; one of the two components, $\bar{d}_0$, of the resulting $\bar{d}$ has defining numbers b^j_i/α^j_i, $\bar{d}_0 \sim \bar{m}_0^{\pm\alpha^j_i} \bar{\ell}_0^{\pm b^j_i}$, where $\bar{m}_0$ is the union of the preimage of $\hat{m}_0$ in $\hat{T}$ with the horizontal diagonals of the rhombi R_1 and R_2 and $\bar{\ell}_0$ is the union of the preimage of ℓ_0 in $\hat{T}$ with the vertical diagonals of the rhombi R_1 and R_3. See Figure (11).

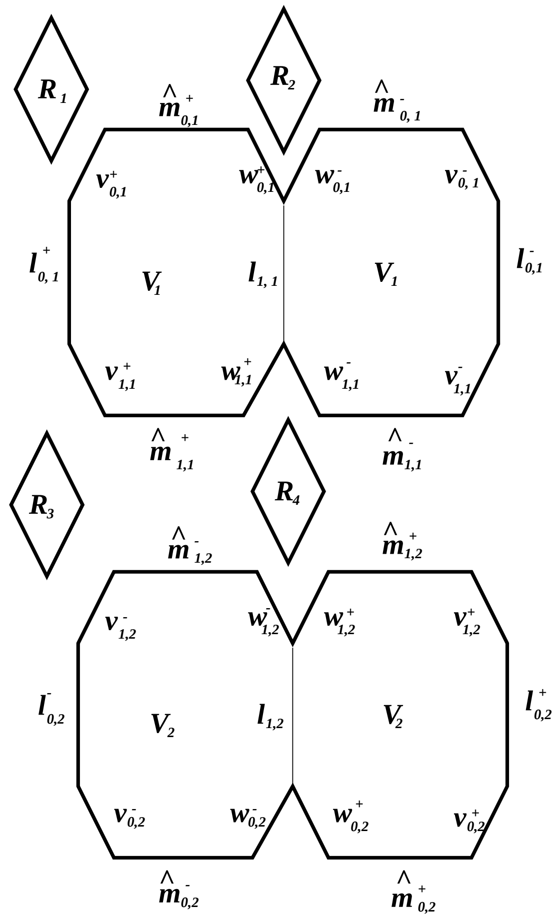


Figure 11

Notice that if either $b_i = 0$ or $\alpha_i = 0$, then $b_i^j/\alpha_i^j = 0$ or ∞, resp.

For practical computation of the number b_i^j/α_i^j, we may assume that $b_i \neq 0$ and $\alpha_i \neq 0$, and visualize the universal cover of $\bar{T}$ as follows:

A fundamental region for $\bar{T}$ is a square $[2u, 2u+2] \times [2v, 2v+2]$ in the plane, for u, v integers; at each vertex $(u, v) \in Z^2$, we cut the plane along the interval $J_{(u,v)} = \{u\} \times [v - j/n, v + j/n]$, and glue a thin rhombus along the resulting boundaries $(\{u\} \times [v - j/n, v])^+$, $(\{u\} \times [v, v + j/n])^-$, $(\{u\} \times [v, v + j/n])^+$, $(\{u\} \times [v - j/n, v])^-$.

A lifting of $\bar{d}_0$ is made out of segments of straight lines of slope b_i/α_i and vertical segments in the glued rhombi:

Start in the point, say, $p_0 = (0, 1/2\alpha_i)$; draw the segment L_1 of the line of slope b_i/α_i starting in p_0 to the right of the interval $J_{(0,0)}$, until it hits for the first time an interval $J_{(u_1,v_1)}$ in a point $p_1 = (u_1, v_1 \pm (2j_1 - 1)/2\alpha_i)$; jump to $\bar{p}_1 = (u_1, v_1 \mp (2j_1 - 1)/2\alpha_i)$ with a vertical segment in the corresponding rhombus; now draw the segment L_2 of the line of slope b_i/α_i starting in $\bar{p}_1$ to the *left* of the interval $J_{(u_1,v_1)}$, until it hits for the first time an interval $J_{(u_2,v_2)}$ in a point $p_2 = (u_2, v_2 \pm (2j_2 - 1)/2\alpha_i)$; jump to $\bar{p}_2 = (u_2, v_2 \mp (2j_2 - 1)/2\alpha_i)$ with a vertical segment in the corresponding rhombus; now draw the segment L_3 of the line of slope b_i/α_i starting in $\bar{p}_2$ to the *right* of the interval $J_{(u_2,v_2)}$. Continuing in this fashion, eventually we draw a segment L_r which ends in an interval $J_{(u_r,v_r)}$ in a point $p_r = (u_r, v_r + 1/2\alpha_i)$ with u_r and v_r even integers, and u_q and v_q are not both even for $q < r$ if $p_q = (u_q, v_q + 1/2\alpha_i)$. We see that $b_i^j/\alpha_i^j = v_r/u_r$.

Remark. With this representation in the plane it is easy to see that: *If* $b_i\ell = -1+\alpha_i k$, $b_i r = 1+\alpha_i s$ *with* ℓ *and* r *minimal positive, and* $j/n \in (1/2\alpha_i, 3/2\alpha_i)$, *then* $b_i^j/\alpha_i^j = (k-s)/(\ell - r)$.

Similar expressions for b_i^j/α_i^j when $j/n > 3/2\alpha_i$ seem to be extremely complicated.

(5.1) An algorithm. We obtain the following algorithm in the square pillowcase B_i to compute b_i^j/α_i^j.

Orient the curve $d = \partial D_i$, and let ε be the sign of b_i. Assume $j/n \in [0, 1/2)$, and write $m_0 = \{0,1\}\times[1-j/n, 1]\cup[0,1]\times\{1\}$, $m_1 = \{0,1\}\times[0, j/n]\cup[0,1]\times\{0\}$, $\ell_0 = \{0\} \times [j/n, 1 - j/n]$

1. Mark the point $p_0 = (0, 1/2\alpha_i)$ with $+1$.

2. If the point $p_u \in d \cap \partial I^2$ is marked with $\varepsilon_u \in \{-1, +1\}$, then the straight line segment of d starting in p_u and following the given orientation of d, hits ∂I^2 in a point p_{u+1}.

If p_{u+1} has no mark, then

If ($p_u \in m_0$ and $p_{u+1} \in m_0$) or ($p_u \in m_1$ and $p_{u+1} \in m_1$),

then mark p_{u+1} with $\varepsilon_{u+1} = -\varepsilon_u$;

else mark p_{u+1} with $\varepsilon_{u+1} = \varepsilon_u$.

Go to **2.** with '$u := u + 1$'.

If p_{u+1} is already marked, then call $b =$ the sum of marks of points in m_0, and $\alpha =$ the sum of marks of points in ℓ_0. Then $b_i^j/\alpha_i^j = \varepsilon b/\alpha$.

(5.2) Rules. Let $\beta/\alpha \in Q^* = Q \cup \{\infty\}$ with $\alpha > 0$. Let $J^0 = [0, 1/2\alpha)$, and $J^i = ((2i-1)/2\alpha, (2(i+1)-1)/2\alpha)$ for $i = 1, \dots, [\alpha/2] - 1$, and $J^{[\alpha/2]} = ((2[\alpha/2] - 1)/2\alpha, 1/2)$. Let $J = J_{\beta/\alpha} = \bigcup_{i=0}^{[\alpha/2]} J^i \subset [0, 1/2)$.

Let $\lambda \in J$. In the square pillowcase for β/α, $(B, a, m \cup \ell)$, with meridional disk $d_{\beta/\alpha}$, define

$$\begin{aligned} m_0^\lambda &= m_0 \cup \{0,1\} \times [1-\lambda, 1], \\ m_1^\lambda &= m_1 \cup \{0,1\} \times [0, \lambda], \\ \ell_0^\lambda &= \{0\} \times [\lambda, 1-\lambda], \\ \ell_1^\lambda &= \{1\} \times [\lambda, 1-\lambda], \end{aligned}$$

and $m^\lambda = m_0^\lambda \cup m_1^\lambda$, $\ell^\lambda = \ell_0^\lambda \cup \ell_1^\lambda$.

We have a function $J \to Q^*$ such that $\lambda \mapsto (\beta/\alpha)_\lambda =$ the defining rational number for $(B, a^\lambda, m^\lambda \sqcup \ell^\lambda)$ with meridional disk $d_{\beta/\alpha}$.

With this notation we get the rules:

0) If $\beta < 0$, *then* $(\beta/\alpha)_\lambda = -(|\beta|/\alpha)_\lambda$.

a) If $\beta\ell = -1 + \alpha k$, *and* $\beta r = 1 + \alpha s$ *with* ℓ *and* r *minimal positive, and* $\lambda \in J^1$, *then* $(\beta/\alpha)_\lambda = (k-s)/(\ell-r)$.

b) If $\lambda, \mu \in J^i$, *then* $(\beta/\alpha)_\lambda = (\beta/\lambda)_\mu$.

c) Let λ *be in the last interval* $J^{[\alpha/2]}$. *Then* $(\beta/\alpha)_\lambda = b/1$ *for some integer* b *if* α *is odd, and* $(\beta/\alpha)_\lambda = \infty$ *if* α *is even.*

In view of rule (c), we get: *If* $n > 2\alpha_i$, *and* $(n, \alpha_i) = 1$ $(i = 1, \dots, t)$, *then* $\varphi^{-1}(k)$ *contains a rational link, where* $k = m(n\beta_1/\alpha_1, \dots, n\beta_t/\alpha_t)$, *and* $\varphi : S^3 \to (S^3, k)$ *is an* n*-fold dihedral quotient,* for, by the assumption $n > 2\alpha_i$, there is an integer $j \in \{1, \dots, (n-1)/2\}$ such that $j/n \in J^{[\alpha_i/2]}$, the last interval for the tangle β_i/α_i, $(i = 1, \dots, t)$.

d) If $\beta/\alpha = \pm 1/\alpha$ *and* $\lambda \in J^u$, *then* $(\beta/\alpha)_\lambda = \pm 1/(\alpha - 2u)$.

More generally, recalling the construction of the tangle with number β/α in [8], we get

e) Assume $\alpha > 2k|\beta|$ *for some* k *positive. If* $\lambda \in J^{k|\beta|}$, *then* $(\beta/\alpha)_\lambda = \beta/(\alpha - 2k|\beta|)$.

In the special case $\alpha \equiv \pm 1 \bmod \beta$, examining carefully the signs in the square pillowcase given by Algorithm 5.1, one can see that:

f) Assume $\alpha = \beta\ell \pm 1$, *and* $\beta > 0$*; let* r *be such that* $0 \le r \le [\ell/2]$.

If $t \in \{0, 1, \dots, [\beta/2]\}$ *and* $\lambda \in J^{r\beta+t}$, *then*

$$\left(\frac{\beta}{\alpha}\right)_\lambda = \frac{\beta - 2t}{(\beta - 2t)(\ell - 2r) \pm 1}.$$

If $t \in \{[\beta/2]+1, [\beta/2]+2, \dots, \beta - 1\}$ *and* $\lambda \in J^{r\beta+t}$, *then*

$$\left(\frac{\beta}{\alpha}\right)_\lambda = \frac{2t - \beta}{(2t - \beta)(\ell - 2r - 2) \pm 1}.$$

In particular we compute that the first $\beta + 1$ terms (when $r = 0$) of the sequence $\left\{\left(\dfrac{\beta}{\alpha}\right)_\lambda\right\}_{t=1}^{\beta}$ are:

$$\frac{\beta}{\alpha} = \frac{\beta}{\beta\ell \pm 1}, \frac{\beta - 2}{(\beta-2)\ell \pm 1}, \frac{\beta - 4}{(\beta - 4)\ell \pm 1}, \dots$$

$$\dots, \begin{cases} \dfrac{4}{4\ell \pm 1}, \dfrac{2}{2\ell \pm 1}, 0, \dfrac{2}{2(\ell-2) \pm 1}, \dfrac{4}{4(\ell - 2) \pm 1}, \dots, \dfrac{\beta}{\alpha - 2\beta} & \text{if } \beta \text{ even} \\ \dfrac{3}{3\ell \pm 1}, \dfrac{1}{\ell \pm 1}, \dfrac{1}{(\ell - 2) \pm 1}, \dfrac{3}{3(\ell-2) \pm 1}, \dots, \dfrac{\beta}{\alpha - 2\beta} & \text{if } \beta \text{ odd} \end{cases}$$

Proposition (5.2.1). *Let* q *be an odd integer,* $q \notin \{-11, -7, -5, -3, -1, 1, 3, 5\}$, *and let* k *be the pretzel knot* $k = p(2, q, q) = m(1/2, \pm 1/|q|, \pm 1/|q|)$. *Then there exists a* $|q+4|$*-fold dihedral covering* $\varphi : S^3 \to (S^3, k)$ *such that*

1) If $|q| \equiv 1 \bmod 4$, *then one of the Montesinos knots* $m(1/2, -1/5, -1/5)$ *or* $m(1/2, -2/9, -2/9)$ *lies in* $\varphi^{-1}(k)$.

2) If $|q| \equiv -1 \bmod 4$, *then one of the Montesinos knots* $m(-1/2, 3/5, 3/5)$ *or* $m(-1/2, 2/3, 2/3)$ *lies in* $\varphi^{-1}(k)$.

Proof. Since q is odd, $(q+4, q) = (q, 2) = 1$; since $\Delta(k) = q^2 + 2q + 2q = q(q+4)$, by Theorem (3.2), there exists a $|q+4|$-dihedral covering $\varphi : S^3 \to (S^3, k)$.

"(1)" Assume $|q| = 4\beta+1$; then $k = m(\frac{|q+4|}{2}, \frac{-|q+4|\beta}{|q|}, \frac{-|q+4|\beta}{|q|})$, and, therefore, the pseudo-branch of φ is $k_0 = m(\frac{1}{2}, \frac{-\beta}{|q|}, \frac{-\beta}{|q|})$.

If β is odd, write $i = (\beta-1)/2$; then $i/|q+4| \in J^i$ for the tangle $-\beta/|q|$, that is, $(2i-1)/2|q| < i/|q+4| < (2i+1)/2|q|$; and $i/|q+4| \in J^0$ for the tangle $1/2$, that is, $i/|q+4| < 1/4$.

By Rule (f), $\left(\frac{-\beta}{|q|}\right)_{i/|q+4|} = -1/5$, and $\left(\frac{1}{2}\right)_{i/|q+4|} = 1/2$. We conclude that $k_i = m(1/2, -1/5, -1/5) \subset \varphi^{-1}(k)$.

If β is even, write $i = \beta/2 - 1$; then $i/|q+4| \in J^i$ for the tangle $-\beta/|q|$; and $i/|q+4| \in J^0$ for the tangle $1/2$. By Rule (f), $\left(\frac{-\beta}{|q|}\right)_{i/|q+4|} = -2/9$, and $\left(\frac{1}{2}\right)_{i/|q+4|} = 1/2$. We conclude that $k_i = m(1/2, -2/9, -2/9) \subset \varphi^{-1}(k)$.

"(2)" Assume $|q| = 4\beta - 1$; then $k = m(\frac{-|q+4|}{2}, \frac{|q+4|\beta}{|q|}, \frac{|q+4|\beta}{|q|})$, and, therefore, the pseudo-branch of φ is $k_0 = m(-\frac{1}{2}, \frac{\beta}{|q|}, \frac{\beta}{|q|})$.

If β is odd and $q \geq 7$, write $i = (\beta+5)/2$. Then $i/(q+4) \in J^{(\beta+3)/2}$ for the tangle $\beta/|q|$; and $i/(q+4) \in J^0$ for the tangle $-1/2$. By Rule (f), $\left(\frac{\beta}{|q|}\right)_{i/(q+4)} = 3/5$, and $\left(\frac{-1}{2}\right)_{i/(q+4)} = -1/2$. We conclude that $k_i = m(-1/2, 3/5, 3/5) \subset \varphi^{-1}(k)$.

If β is odd and $q < -11$, then $|q+4| = -(q+4) > 7$; write $i = (\beta+1)/2$. Then $-i/(q+4) \in J^{(\beta+3)/2}$ for the tangle $\beta/|q|$; and $-i/(q+4) \in J^0$ for the tangle $-1/2$. Again $k_i = m(-1/2, 3/5, 3/5) \subset \varphi^{-1}(k)$.

If β is even and $q \geq 7$, write $i = \beta/2+2$. Then $i/(q+4) \in J^{\beta/2+1}$ for the tangle $\beta/|q|$; and $i/(q+4) \in J^0$ for the tangle $-1/2$. By Rule (f), $\left(\frac{\beta}{|q|}\right)_{i/(q+4)} = 2/3$, and $\left(\frac{-1}{2}\right)_{i/(q+4)} = -1/2$. We conclude that $k_i = m(-1/2, 2/3, 2/3) \subset \varphi^{-1}(k)$.

If β is even and $q < -11$, then $|q+4| = -(q+4) > 7$; write $i = \beta/2$. Then $\left(\frac{\beta}{|q|}\right)_{-i/(q+4)} = 2/3$; and $\left(\frac{-1}{2}\right)_{-i/(q+4)} = -1/2$. Again $k_i = m(-1/2, 2/3, 2/3) \subset \varphi^{-1}(k)$. □

We finish this section rewriting a version of Theorem (3.2).

Corollary (5.2.2). *Let $(n, \alpha_i) = 1$, and $k = m(n\beta_1/\alpha_1, \dots, n\beta_t/\alpha_t)$. If $\varphi : S^3 \to (S^3, k)$ is an n-fold dihedral quotient, and $k_j \subset \varphi^{-1}(k)$ is as before with $j < n/2$, then*

$$k_j = m((\beta_1/\alpha_1)_{j/n}, \dots, (\beta_t/\alpha_t)_{j/n}).$$

6. Dihedral-like coverings of Montesinos knots

Let $k \subset S^3$ be a link. Let $\varphi : M \to (S^3, k)$ be an m-fold branched covering, and let $\omega = \omega_\varphi : \pi_1(S^3 - k) \to S_m$ be the representation determined by φ.

Assume that $order(\omega(\mu)) = p$ for each meridian μ of k. Write $\langle \mu^p \rangle_\pi$ for the normal closure of $\langle \mu^p; \mu \text{ a meridian of } k \rangle$ in $\pi_1(S^3 - k)$; then ω factors through the quotient $\pi_1(S^3-k)/\langle\mu^p\rangle_\pi$

$$\begin{tikzcd}
\pi_1(S^3-k) \arrow[rr, "\omega"] \arrow[dr] & & S_m \\
& \pi_1(S^3-k)/\langle\mu^n\rangle_\pi \arrow[ur, "\bar\omega"'] &
\end{tikzcd}$$

It is known (see [4], Ch. 2, § 4, with the slight necessary modifications to the argument) that this quotient is a semi-direct product

$$\pi_1(S^3-k)/\langle\mu^p\rangle_\pi \cong Z_p \ltimes \pi_1(B_p(k)),$$

where $B_p(k)$ denotes the p-fold cyclic covering of S^3 branched along all components of k. The generator of Z_p is the class of a meridian, $\bar\mu$, and acts on $\pi_1(B_p(k))$ as the isomorphism induced by the homeomorphism of order p, $\rho : B_p(k) \to B_p(k)$, such that $B_p(k)/\rho = (S^3, k)$.

Let $\psi : \tilde M \to B_p(k)$ be the m-fold covering space determined by $\bar\omega| : \pi_1(B_p(k)) \to S_m$, which might be not transitive; therefore, $\tilde M$ might be not connected.

Now assume $p = 2$. Then $\bar\mu$ (which is the isomorphism induced by the involution ρ) leaves invariant $\psi_\#\pi_1(\tilde M) \le \pi_1(B_2(k))$, and, therefore, ρ lifts to an involution $\tilde\rho : \tilde M \to \tilde M$, with quotient a 2-fold cyclic branched covering $q : \tilde M \to M = \tilde M/\tilde\rho$. The quotient-induced m-fold branched covering is the original $\varphi : M \to (S^3, k)$, and we have a commutative diagram

$$\begin{tikzcd}
& \tilde M \arrow[dl, "q"'] \arrow[dr, "\psi"] & \\
M \arrow[dr, "\varphi"'] & & B_n(k) \arrow[dl, "p"] \\
& (S^3, k) &
\end{tikzcd}$$

The branching of q is the pseudo-branch of φ (if any).

Conversely, if we start with a representation

$$\omega : \pi_1(S^3-k)/\langle\mu^2\rangle_\pi = Z_2 \ltimes \pi_1(B_2(k)) \to S_m,$$

we get a commutative diagram of coverings

$$\begin{tikzcd}
& \tilde M \arrow[dl, "q"'] \arrow[dr, "\psi"] & \\
M \arrow[dr, "\varphi"'] & & B_2(k) \arrow[dl, "p"] \\
& (S^3, k) &
\end{tikzcd}$$

with q a 2-fold covering branched along the pseudo-branch of φ.

1. Let $k = m(\beta_1/\alpha_1, \dots, \beta_t/\alpha_t)$. Then $\pi_1(S^3 - k)/\langle \mu^2 \rangle_\pi$ equals the semi-direct product $Z_2 \ltimes \pi_1(O, 0; \beta_1/\alpha_1, \dots, \beta_t/\alpha_t)$; the quotient, obtained by 'killing' an ordinary fiber of $B_2(k)$, gives the epimorphism

$$\pi_1(S^3 - k)/\langle \mu^2 \rangle_\pi \to Z_2 \ltimes \Delta(\alpha_1, \dots, \alpha_t).$$

In terms of generators and relations, $Z_2 \ltimes \Delta(\alpha_1, \dots, \alpha_t) = \langle \mu, q_1, \dots, q_t : \mu^2 = 1, q_1^{\alpha_1} = 1, \dots, q_t^{\alpha_t} = 1, q_1 \cdots q_t = 1, q_i^\mu = q_i^{-1} \rangle = \langle \mu, q_1, \dots, q_{t-1} : \mu^2 = 1, q_1^{\alpha_1} = 1, \dots, q_{t-1}^{\alpha_{t-1}} = 1, (q_1 \cdots q_{t-1})^{\alpha_t} = 1, q_i^\mu = q_i^{-1} \rangle$. To find a representation of $Z_2 \ltimes \Delta(\alpha_1, \dots, \alpha_t)$ into S_n, it is equivalent to find permutations $\tau, \sigma_1, \dots, \sigma_{t-1} \in S_n$ such that $\tau^2 = (1)$, $\sigma_i^{\alpha_i} = (1)$, $\sigma_i^\tau = \sigma_i^{-1}$ $(i = 1, \dots, t-1)$, and $(\sigma_1 \cdots \sigma_{t-1})^{\alpha_t} = (1)$. This task simplifies when $t = 3$, for it is possible to easily draw the product of two permutations ([7]), which amounts to drawing the induced branched covering on the orbit surfaces of $\psi : \tilde{M} \to B_2(k)$.

In the computations of the following representations, the GAP programming language ([3]) was very useful. We will use, for computations of Seifert symbols, Lemma 2 of [12].

2. Consider $k = m(\beta_1/2, \beta_2/3, \beta_3/3)$. We find $\sigma_1 = (1)$, $\sigma_2 = (1, 2, 3)$, $\tau = (2, 3)$ (the other two meridians of k go to $(2, 3)$ and $(1, 2)$); we compute $\tilde{M} = (O, 0; \beta_1/2, \beta_1/2, \beta_1/2, \beta_2/1, \beta_3/1)$ ([12], Lemma 2), and the involution of $B_2(k)$ lifts to the involution $\tilde{u}$ in $\tilde{M}$ with axis crossing the fibers corresponding to one of the $\beta_1/2$ ratios, and $\beta_2/1$, $\beta_3/1$, and interchanging the other two fibers corresponding to $\beta_1/2$, as seen in Figure (12).

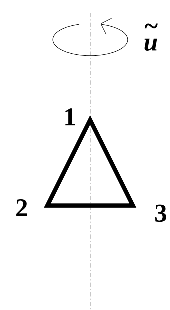


Figure 12. $\sigma_1 = (1)$, and $\sigma_2 = (1, 2, 3)$

We have $\varphi_1 : M = \tilde{M}/\tilde{u} \to (S^3, k)$ a 3-fold (simple) branched covering with pseudo-branch the branching of the quotient $\tilde{M} \to M$.

Let $\tilde{M}_0$ be the result of drilling out the fibers $\beta_1/2$ in $\tilde{M}$ interchanged by $\tilde{u}$; then $\tilde{M}_0/\tilde{u}$ is the complement of a trivial knot T in S^3 and the branching looks as in Figure (13): the link $m(\beta_1/2, \beta_2/1, \beta_3/1) \subset \tilde{M}_0/\tilde{u}$. To construct $\tilde{M}/\tilde{u}$, we have to fill $\tilde{M}_0/\tilde{u}$ with a solid torus whose meridian goes around $m^2\ell^{\beta_1}$ for m, ℓ a meridian-longitude pair of T.

Therefore $M = \tilde{M}/\tilde{u}$ is the lens space $L(2, \beta_1)$. If $\varphi_2 : S^3 \to M$ is the universal covering space, then $\varphi_2^{-1}(\tilde{M}_0/\tilde{u})$ is the complement of a trivial knot where we still have to perform surgery $1/\beta_1$; then the preimage in S^3 of the pseudo-branch

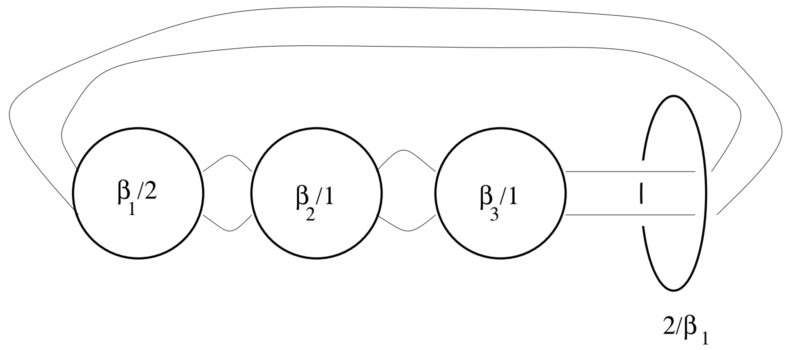


Figure 13

of φ_1 is the Montesinos knot

$$\tilde{k} = m\left(\frac{\beta_1}{2}, \frac{\beta_2}{1}, \frac{\beta_3}{1}, \frac{\beta_1}{2}, \frac{\beta_2}{1}, \frac{\beta_3}{1}, \frac{2\beta_1}{1}\right) = m\left(-\frac{12\beta_1 + 8\beta_2 + 8\beta_3}{6\beta_1 + 4\beta_2 + 4\beta_3 - 1}\right).$$

This link is a torus link if and only if $6\beta_1 + 4\beta_2 + 4\beta_3 - 1 \equiv \pm 1 \bmod (12\beta_1 + 8\beta_2 + 8\beta_3)$, that is, if and only if $\Delta(k)/3 = 3\beta_1 + 2\beta_2 + 2\beta_3 = \pm 1$. Then $\tilde{k}$ is a universal link if and only if $\Delta(k) \neq \pm 3$ (Theorem (1.1)). Notice that the equality $\Delta(k) = \pm 3$ implies that k is the pretzel knot $p(\pm 2, \mp 3, \mp 3)$, which is a torus knot. We have $\varphi = \varphi_1 \circ \varphi_2 : S^3 \to (S^3, k)$ a branched covering with $\tilde{k} \subset \varphi(k)$; we conclude that:

The Montesinos knot $k = m(\beta_1/2, \beta_2/3, \beta_3/3)$ *is a universal link if and only if* $\Delta(k) \neq \pm 3$.

3. Consider $k = m(\beta_1/2, \beta_2/3, \beta_3/5)$. We find $\sigma_1 = (2,3)(4,5)$, $\sigma_2 = (1,2,4)$, $\tau = (1,2)(3,5)$ (the other two meridians of k go to $(2,5)(3,4)$ and $(2,4)(3,5)$); we compute $\tilde{M} = (O, 0; \beta_1/2, \beta_1/1, \beta_1/1, \beta_2/3, \beta_2/3, \beta_2/1, \beta_3/1)$ ([12], Lemma 2), and the involution of $B_2(k)$ lifts to the involution $\tilde{u}$ in $\tilde{M}$ with axis crossing the fibers corresponding to $\beta_1/2$, $\beta_2/1$, $\beta_3/1$, and pairwise interchanging the fibers corresponding to $\beta_1/1$ and $\beta_2/3$, as seen in Figure (14).

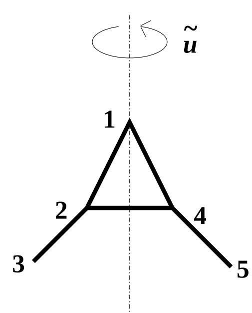


Figure 14. $\sigma_1 = (2,3)(4,5)$, and $\sigma_2 = (1,2,4)$

We have $\varphi : M = \tilde{M}/\tilde{u} \to (S^3, k)$ a 5-fold branched covering with pseudo-branch the branching of the quotient $\tilde{M} \to M$. To draw this pseudo branch we still have to perform surgeries $1/\beta_1$, and $3/\beta_2$ as before (Figure (15)); passing to

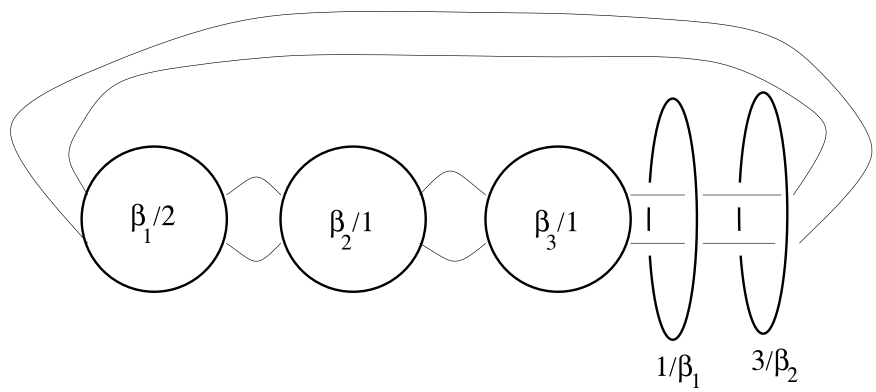


Figure 15

the universal cover of $M = L(3, 3\beta_1 + \beta_2)$ we obtain

$$\begin{aligned}\tilde{k} &= m\left(\frac{\beta_1}{2}, \frac{\beta_2}{1}, \frac{\beta_3}{1}, \frac{\beta_1}{2}, \frac{\beta_2}{1}, \frac{\beta_3}{1}, \frac{\beta_1}{2}, \frac{\beta_2}{1}, \frac{\beta_3}{1}, \frac{2(3\beta_1+\beta_2)}{1}\right)\\ &= m\left(\frac{\beta_1}{2}, \frac{\beta_1}{2}, \frac{\beta_1}{2}, \frac{6\beta_1+5\beta_2+3\beta_3}{1}\right).\end{aligned}$$

If $\beta_1 = \pm 1$, then $\tilde{k}$ is an Uchida pretzel chain,

$$\tilde{k} = p(\pm 6 + 5\beta_2 + 3\beta_3; \pm 2, \pm 2, \pm 2) = \begin{cases} (3, 6+5\beta_2+3\beta_3) & \text{if } \beta_1 = 1 \\ (3, -9+5\beta_2+3\beta_3) & \text{if } \beta_1 = -1 \end{cases}$$

which is universal if $6 + 5\beta_2 + 3\beta_3 \neq -2, -1$ for $\beta_1 = 1$, and is universal if $-9 + 5\beta_2 + 3\beta_3 \neq -2, -1$ for $\beta_1 = -1$ ([15]). If $\beta_1 = 1$, then the equality $6 + 5\beta_2 + 3\beta_3 = -2, -1$ is equivalent to $\Delta(k) = \pm 1$; also if $\beta_1 = -1$, the equality $-9 + 5\beta_2 + 3\beta_3 = -2, -1$ is equivalent to $\Delta(k) = \pm 1$. Note that $\Delta(k) = \pm 1$ implies that k is the pretzel knot $= p(\pm 2, \mp 3, \mp 5)$, which is a torus knot. We conclude that:

The Montesinos knot $k = m(\pm 1/2, \beta_2/3, \beta_3/5)$ *is a universal link if and only if* $\Delta(k) \neq \pm 1$.

Note that any knot $m(\gamma_1/2, \gamma_2/3, \gamma_3/5)$ is also of the form $m(\pm 1/2, \beta_2/3, \beta_3/5)$.

4. Consider $k = m(\beta_1/2, \beta_2/3, \beta_3/7)$. We find $\sigma_1 = (2,3)(4,6)(5,7)(8,9)$, $\sigma_2 = (1,2,4)(3,5,7)(6,9,8)$, $\tau = (2,3)(4,6)(5,7)(8,9)$ (the other two meridians of k go to $(2,6)(3,4)(5,9)(7,8)$ and $(2,4)(3,6)(5,8)(7,9)$); we compute

$$\tilde{M} = (O, 0; \beta_1/2, \beta_1/1, \beta_1/1, \beta_1/1, \beta_1/1, \beta_2/1, \beta_2/1, \beta_2/1, \beta_3/1, \beta_3/7, \beta_3/7)$$

([12], Lemma 2), and the involution of $B_2(k)$ lifts to the involution $\tilde{u}$ in $\tilde{M}$ with axis crossing the fibers corresponding to $\beta_1/2$, $\beta_2/1$, $\beta_3/1$, and interchanging by pairs the fibers corresponding to $\beta_1/1$, $\beta_2/1$, and $\beta_3/7$, as shown in Figure (16).

We have $\varphi : M = \tilde{M}/\tilde{u} \to (S^3, k)$, a 9-fold branched covering with pseudo-branch the branching of the quotient $\tilde{M} \to M$. To draw this pseudo-branch we still have to perform surgeries $1/\beta_1$, $1/\beta_1$, $1/\beta_2$, and $7/\beta_3$ as before (Figure (17)); passing to the universal cover of $M = L(7, 14\beta_1 + 7\beta_2 + \beta_3)$ we obtain

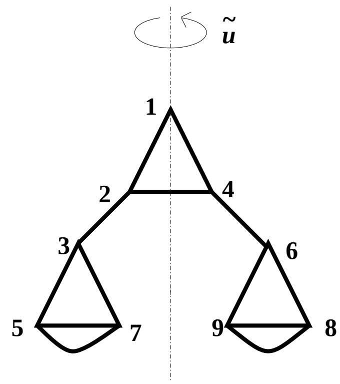


Figure 16. $\sigma_1 = (2,3)(4,6)(5,7)(8,9)$, and $\sigma_2 = (1,2,4)(3,5,7)(6,9,8)$

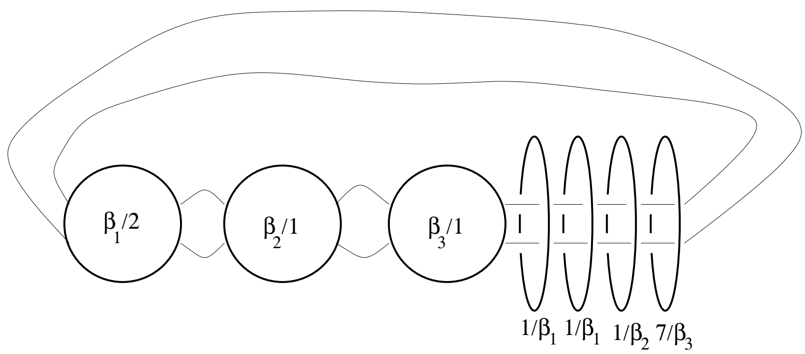


Figure 17

$$\begin{aligned}\tilde{k} &= m\Bigg(\overbrace{\frac{\beta_1}{2},\dots,\frac{\beta_1}{2}}^{7\text{ times}},\overbrace{\frac{\beta_2}{1},\dots,\frac{\beta_2}{1}}^{7\text{ times}},\overbrace{\frac{\beta_3}{1},\dots,\frac{\beta_3}{1}}^{7\text{ times}},\frac{2(14\beta_1+7\beta_2+\beta_3)}{1}\Bigg)\\ &= m\Bigg(\overbrace{\frac{\beta_1}{2},\dots,\frac{\beta_1}{2}}^{7\text{ times}},\frac{28\beta_1+21\beta_2+9\beta_3}{1}\Bigg).\end{aligned}$$

If $\beta_1 = \pm 1$, then $\tilde{k}$ is a universal Uchida pretzel link,

$$\tilde{k} = p(\pm 28 + 21\beta_2 + 9\beta_3;\overbrace{\pm 2,\dots,\pm 2}^{7\text{ times}}).$$

We conclude that:

The Montesinos knot $k = m(\pm 1/2, \beta_2/3, \beta_3/7)$ is a universal link.

Note that any $m(\gamma_1/2, \gamma_2/3, \gamma_3/7)$ is of the form $m(\pm 1/2, \beta_2/3, \beta_3/7)$, and that, in particular, the Fintushel-Stern knot, $k = p(-2, 3, 7)$, is universal; this famous knot has $\Delta(k) = 1$ and, therefore, has no dihedral quotients.

5. Using the same representations, and following the ideas below in paragraphs **2.**, **3.**, and **4.**, we obtain also:

a) If $|x| > 1$, then the pretzel link $k = p(e; 2x, 3y, 3z)$ is universal, for, in this case, following the constructions as in **2.**, $\tilde{k}$ is the universal Uchida pretzel link

$$\tilde{k} = p(\pm(2 + (4x+1)e); \overbrace{2x, \ldots, 2x}^{2|x| \text{ times}}, \overbrace{y, \ldots, y}^{2|x| \text{ times}}, \overbrace{z, \ldots, z}^{2|x| \text{ times}}).$$

a.1) Similarly, if $|y| > 1$ or $|z| > 1$, then the pretzel link $k = p(2, 3y, 3z)$ is universal.

a.2) If $|y| > 1$ or $|z| > 1$, and $\beta_2 \equiv \pm 1$ mod y and $\beta_3 \equiv \pm 1$ mod z, then the Montesinos knot $k = m(1/2, \beta_2/3y, \beta_3/3z)$ is universal, for, as before, it holds that $p(-2(-1+k_2+k_3); 2, 2, \pm y, \pm y, \pm z, \pm z)$, which is a universal Uchida link, is in the preimage of k in a branched covering, where $\beta_2 = k_2 y \pm 1$ and $\beta_3 = k_3 z \pm 1$.

b) If $|y| > 1$, or $|z| > 1$, then the pretzel link $k = p(\pm 2, \pm 3y, \pm 5z)$ is universal, for, in this case, following the constructions as in **3.**, $\tilde{k}$ is the universal Uchida pretzel link

$$\tilde{k} = p(2(\pm 3y \pm 1);\ \overbrace{\pm 2, \ldots, \pm 2}^{3|y| \text{ times}}, \overbrace{\pm y, \ldots, \pm y}^{3|y| \text{ times}}, \overbrace{\pm z, \ldots, \pm z}^{3|y| \text{ times}}).$$

c) If $z > 0$, then the pretzel link $k = p(\pm 2, \pm 3, \pm 7z)$ is universal, for, in this case, following the constructions as in **4.**, $\tilde{k}$ is the universal Uchida pretzel link

$$\tilde{k} = p(2(\pm 14z \pm 7z \pm 1);\ \overbrace{\pm 2, \ldots, \pm 2}^{7z \text{ times}}, \overbrace{\pm 3, \ldots, \pm 3}^{7z \text{ times}}, \overbrace{\pm z, \ldots, \pm z}^{7z \text{ times}}).$$

6. Consider $k = p(\pm 2, 5y, 5z)$ $(y, z \neq 0)$. We find permutations $\sigma_1 = (2,3)(5,7)(6,9)(10,11)$, $\sigma_2 = (1,2,4,8,5)(3,6,7,10,11)$, and $\tau = (2,5)(3,7)(4,8)(10,11)$ (the other two meridians of k go to $(1,2)(3,10)(4,5)(6,7)$ and $(2,7)(3,5)(4,8)(6,9)$); we compute

$$\tilde{M} = (O, 0; \pm 1/2, \pm 1, \pm 1, \pm 1, \pm 1, \pm 1/2, \pm 1/2, 1/5y, 1/y, 1/y, 1/5z, 1/z, 1/z),$$

([12], Lemma 2), and the involution $\tilde{u}$ in $\tilde{M}$ (see Figure (18)) gives $M = L(2,1)$ as quotient. In the universal cover of M we find, as preimage of k,

$$\tilde{k} = p(\pm 10; \pm 2, \pm 2, 5y, 5y, y, y, y, y, 5z, 5z, z, z, z, z)$$

which is a universal Uchida link. We conclude

The pretzel knot $p(\pm 2, 5y, 5z)$ is universal $(y, z \neq 0)$.

7. We think that it is not possible to obtain the 3-sphere as a branched covering over $(S^3, m(\beta_1/2, \beta_2/p, \beta_3/q))$ with this kind of dihedral-like coverings, except for the pairs (p, q) as in paragraphs **2.**–**6.**.

7. Universal Montesinos knots

We collect the following results.

THEOREM (7.1). *Let q be an odd integer, $q \neq -1, -3, -7, -11$. Then the pretzel knot $k = p(2, q, q)$ is universal.*

Proof. If $q = 1$, then $k = m(5/2)$ is universal (Theorem (1.1)). If $q = 3$, k is universal by § 6.6. If $q = \pm 5$, k is universal by § 6.6. Assume $|q| > 5$, $q \neq -7, -11$. By Proposition 1 in § 5.1, there is a dihedral quotient $\varphi : S^3 \to$

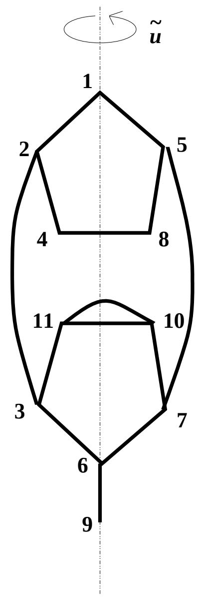


Figure 18. $\sigma_1 = (2,3)(5,7)(6,9)(10,11)$, and $\sigma_2 = (1,2,4,8,5)(3,6,7,10,11)$

(S^3, k) such that $\varphi^{-1}(k)$ contains one of the knots $p_1 = m(1/2, -1/5, -1/5)$, $p_2 = m(1/2, -2/9, -2/9)$, $p_3 = m(-1/2, 3/5, 3/5)$, or $p_4 = m(-1/2, 2/3, 2/3)$.

Now p_1 is universal by § 6.6; p_2 is universal by § 6.5 (a.2); p_4 is universal by § 6.2; $p_3 = m(7/2, -7/5, -7/5)$ has a 7-fold dihedral quotient with pseudo-branch p_1 which is universal, so, p_3 is universal. We conclude that k is universal. □

We remark that $p(2, -1, -1) =$ trefoil knot, and $p(2, -3, -3) = \tau_{3,4}$ are not universal.

Question. Are the knots $p(2, -7, -7)$ and $p(2, -11, -11)$ universal?

THEOREM (7.2). *Let n be a positive integer such that $n > \alpha(\alpha - 2)/2$ and $(n, \alpha) = (n, \alpha_i) = 1$, where $\alpha_1, \ldots, \alpha_t$ are odd positive integers with $\alpha \geq \alpha_i + 2$ for each i. Let $\varepsilon_i \in \{-1, +1\}$ $(i = 1, \ldots, t)$ such that $\sum \varepsilon_i \neq -2, -1, 0, 1$. Then*

$$k = m(n\varepsilon_1/\alpha_1, \ldots, n\varepsilon_t/\alpha_t, n/\alpha)$$

is a universal link.

Proof. The hypothesis $n > \alpha(\alpha-2)/2$ implies that $(\alpha-2)/2\alpha - (\alpha_i - 2)/2\alpha_i \geq (\alpha-2)/2\alpha - (\alpha-4)/2(\alpha-2) > 1/n$; therefore there is some $j \in \{1, \ldots, (n-1)/2\}$ such that j/n is in the last interval $J^{[\alpha_i/2]}$ for the tangle ε_i/α_i, and in the next-to-last interval $J^{[\alpha_i/2]-1}$ for the tangle $1/\alpha$ (see § 5.2.). Then, in an n-fold dihedral quotient $\varphi : S^3 \to (S^3, k)$, we have $k_j = m(\varepsilon_1/1, \ldots, \varepsilon_t/1, 1/3) = m(\sum \varepsilon_i + 1/3) \subset \varphi^{-1}(k)$ (§ 5.2.(c)) if α odd, and $k_j = m(\varepsilon_1/1, \ldots, \varepsilon_t/1, 1/2) = m(\sum \varepsilon_i + 1/2)$ if α even. Since $m(\sum \varepsilon_i + 1/3)$ and $m(\sum \varepsilon_i + 1/2)$ are non-torus 2-bridge links, by the choice of the ε_i's, we see that $m(\sum \varepsilon_i + 1/3)$ and $m(\sum \varepsilon_i + 1/2)$ are universal (Theorem (1.1)); therefore k is universal. □

From Theorem (1.2) in the Introduction and Corollary (3.3.4), it follows that:

Theorem (7.3). *Let $b, \alpha_1, \ldots, \alpha_t$ be integers such that $p(b; \alpha_1, \ldots, \alpha_t)$ is a universal Uchida link. Let n be an integer such that $(n, \alpha_i) = 1$ for $i = 1, \ldots, t$. Then*

$$m(nb/1, n/\alpha_1, \ldots, n/\alpha_t)$$

is a universal link.

From Theorem (1.3) in the Introduction and Corollary (3.3.4) it follows that:

Theorem (7.4). *If $|p| > 1$ is odd and $(n, p) = 1$, then $m(n/p, n/p, -n/p)$ is a universal link.*

If $p \neq -2, 0$ is even and $(n, p) = 1$, then $m(n/3, n/3, n/p)$ is a universal link.

(7.5) Montesinos knots up to ten crossings and Hatcher-Oertel knots.

Montesinos knots. We analyze the Montesinos knots up to 10 crossings. We do not consider here 2-bridge knots, but only 'real' Montesinos knots.

1. *Non-universal.* The knots $8_{19} = p(-2, 3, 3) = \tau_{3,4}$, and $10_{124} = p(-2, 3, 5) = \tau_{3,5}$ are torus knots and not universal.

2. *Universal.* By § 6.2, the knots 8_5, 8_{10}, 8_{15}, 8_{20}, 8_{21}, 9_{16}, 9_{24}, 9_{28}, 10_{76}, 10_{77}, and 10_{78} are universal.

By § 6.3, the knots 9_{22}, 9_{25}, 9_{30}, 9_{36}, 9_{42}, 9_{43}, 9_{44}, 9_{45}, 10_{46}, 10_{47}, 10_{48}, 10_{49}, 10_{70}, 10_{71}, 10_{72}, 10_{73}, 10_{125}, 10_{126}, and 10_{127} are universal.

By § 6.4, the knots 10_{50}, 10_{51}, 10_{52}, 10_{53}, 10_{54}, 10_{55}, 10_{56}, 10_{57}, 10_{128}, 10_{129}, 10_{130}, 10_{131}, 10_{132}, 10_{133}, 10_{134}, and 10_{135} are universal (also all of them have a dihedral quotient with the Fintushel-Stern knot as pseudo-branch).

By § 6.5 (a), or Theorem (7.3), the knots 10_{61}, 10_{62}, 10_{63}, 10_{64}, 10_{65}, 10_{66}, 10_{139}, 10_{140}, 10_{141}, 10_{142}, 10_{143}, and 10_{144} are universal.

The knots 10_{58}, 10_{60}, 10_{136}, and 10_{138} have the pretzel knot $p(-2, 5, 5)$ as pseudo-branch in a dihedral covering of 13, 17, 3 and 7 sheets, resp. The knot 10_{59} has the pretzel knot $p(-2, -5, 5)$ as pseudo-branch in a 3-fold dihedral covering. Then, using § 6.6, we see that 10_{58}, 10_{59}, 10_{60}, 10_{136}, and 10_{138} are universal knots.

The knots $9_{37} = m(1/3, 2/3, 2/3) = m(-5/3, 5/3, 5/3)$, $9_{46} = m(-1/3, 1/3, 1/3)$, and $10_{74} = m(1/3, 1/3, 5/3) = m(-7/3, 7/3, 7/3)$ are universal by Theorem (7.4).

We found sixty six universal knots.

3. *Undecided.* At this point, we cannot decide about the universality of ten knots: 9_{35}, 9_{48}, 10_{67}, 10_{68}, 10_{69}, 10_{75}, 10_{137}, 10_{145}, 10_{146}, and 10_{147}.

4. *No dihedral quotients.* The following fourteen knots do not have the 3-sphere as a dihedral quotient (yet some of them have S^3 as a dihedral-like covering, as noted before): 8_{10}, 8_{19}, 8_{20}, 9_{35}, 9_{46}, 9_{48}, 10_{75}, 10_{124}, 10_{137}, 10_{139}, 10_{140}, 10_{143}, 10_{145}, and 10_{147}.

Hatcher-Oertel knots. The following knots are mentioned in [5].

1. $k = m(2/5, 3/7, -1/3, -5/8)$, $\Delta(k) = 109$. In the 109-fold dihedral quotient, $k_{33} = m(2/1, 1/1, 1/3, 1/4, -1/1) = m(-31/22)$ which is universal and, therefore, k is universal.

2. $k = m(-15/32, 3/11, 7/41)$, $\Delta(k) = 365$. In the 365-fold dihedral quotient, $k_{33} = m(5/14, 2/5, -1/1) = m(-17/4)$ which is universal and, therefore, k is universal.

4. $k = m(11/53, 17/43, -13/21)$, $\Delta(k) = 773$. In the 773-fold dihedral quotient, $k_{19} = m(-24/49, 13/33) = m(-155/3)$ which is universal and, therefore, k is universal.

5. $k = m(1/3, 3/5, -3/4, -2/7, 3/11, -5/13)$, $\Delta(k) = 12869$. In the 12869-fold dihedral quotient, $k_{2758} = m(1/1, -1/2, 1/1, 1/1) = m(7/2)$ which is universal and, therefore, k is universal.

6. $k = m(1/3, 1/3, -1/3, -2/5, 1/5, -3/4, 2/3)$, $\Delta(k) = 405$. Therefore k has not the 3-sphere as a dihedral quotient.

Received June 22, 2003

Final version received May 11, 2004

CIMAT
Apdo. Postal 402
36000 Guanajuato, Gto.
México
victor@cimat.mx
jesusr@cimat.mx

References

[1] G. Burde, *On branched coverings of S^3*, Can. J. Math. XXIII (1971), 84–89.

[2] R. Fox, *A quick trip through knot theory. Topology of 3-manifolds and related topics*, Proc. Univ. Georgia institute (1961), 120–167.

[3] The GAP Group, GAP — Groups, Algorithms, and Programming, Version 4.2; 2002. (http://www.gap-system.org)

[4] F. González-Acuña, *On homology spheres*, Ph. D. Thesis, Princeton University, 1969.

[5] A. Hatcher and U. Oertel, *Boundary slopes for Montesinos knots*, Topology **28** (1989), 453–480.

[6] H. Hilden, M. Lozano and J. Montesinos, *On knots that are universal*, Topology **24** (1985), 499–504.

[7] R. Lyndon and J. Brenner, *The orbits of the product of two permutations*, European J. Combin. **4** (1983), 279–293.

[8] J. Montesinos, *Variedades de Seifert que son recubridores cíclicos ramificados de dos hojas*, Bol. Soc. Mat. Mexicana **18** (1973), 1–32.

[9] J. Montesinos, *Representación de enlaces en relación con recubridores dobles ramificados*, Collectanea Math. **25** (1974), 145–157.

[10] W. Neumann and F. Raymond, *Seifert manifolds, plumbing, μ-invariant and orientation reversing maps*, Springer, Lecture Notes in Math. **664** (1978), 163–196.

[11] V. Núñez, *On the Heegaard genus and tri-genus of non-orientable Seifert 3-manifolds*, Top. Appl. **98** (1999), 241–267.

[12] V. Núñez and E. Ramírez, *The trefoil knot is as universal as it can be*, Top. Appl. **130** (2003), 1–17.

[13] J. Rodríguez, *Some universal pretzel knots*, Preprint.

[14] W. Thurston, *Universal links*, 1982, Preprint.

[15] Y. Uchida, *Universal chains*, Kobe J. Math. **8** (1991), 55–65.

[16] Y. Uchida, *Universal pretzel links*, Knots '90, 241–270. Walter de Gruyter, Berlin-New York, 1992.

[17] H. Zieschang, *Classification of Montesinos knots*, Springer, Lecture Notes in Math. **1060** (1984), 378–389.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

CROSSCAP NUMBER TWO KNOTS IN S^3 WITH (1,1) DECOMPOSITIONS

ENRIQUE RAMÍREZ-LOSADA AND LUIS G. VALDEZ-SÁNCHEZ

Dedicated to Fico on the occasion of his 60th birthday.

Abstract. M. Scharlemann has recently proved that any genus one tunnel number one knot is either a satellite or 2-bridge knot, as conjectured by H. Goda and M. Teragaito; all such knots admit a (1,1) decomposition. In this paper we give a classification of the family of (1,1) knots in S^3 with crosscap number two (i.e., bounding an essential once-punctured Klein bottle).

1. Introduction

H. Goda and M. Teragaito classified in [6] the family of non-simple genus one tunnel number one knots, and conjectured that any genus one tunnel number one simple knot is a 2-bridge knot. This conjecture was shown by H. Matsuda [9] to be equivalent to the statement that any genus one tunnel number one knot in S^3 admits a $(1,1)$ decomposition; it is in this form that M. Scharlemann has recently settled it in [13].

In this paper we explore the family of crosscap number two tunnel number one knots in S^3. Recall (cf. [1]) that a knot in S^3 has crosscap number two if it bounds a once-punctured Klein bottle but not a Moebius band; it was shown in [12] that a knot K has crosscap number two iff its exterior contains a properly embedded *essential* (incompressible and boundary incompressible, in the geometric sense) once-punctured Klein bottle F, in which case K is not a 2-torus knot, and F has integral boundary slope by [8].

In contrast with genus one knots, a crosscap number two knot can bound once-punctured Klein bottles with distinct boundary slopes; however, as shown in [8, 12], such knots are all satellite knots, with the exception of the figure-8 knot and the Fintushel-Stern $(-2,3,7)$ pretzel knot. Here we restrict our attention to the family of crosscap number two knots in S^3 which admit a $(1,1)$ decomposition; the special cases of tunnel number one satellite knots, 2-bridge knots, and torus knots, are also discussed.

In order to state our main result we need to define a particular family of $(1,1)$ knots in S^3. Let S be a Heegaard torus of S^3, and let $S\times I$ be a product regular neighborhood of S, with S corresponding to $S\times\{1/2\}$. An arc β embedded in $S\times I$ is called *monotone* if the natural projection map $S\times I\to I$ is monotone on β. For $i=0,1$, let t_i be an embedded nontrivial circle in $S\times\{i\}$; t_i^* will denote

2000 *Mathematics Subject Classification*: Primary 57M25; Secondary 57N10.

Keywords and phrases: Crosscap number two knot, tunnel number one knot, $(1,1)$ decomposition.

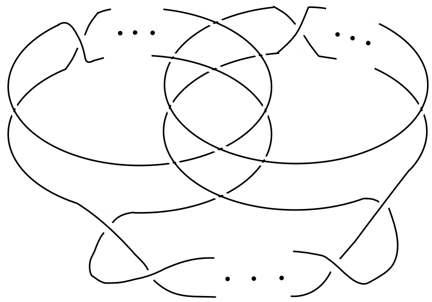


Figure 1. A knot of the form $K(t_0^*, t_1^*, R)$.

a $(\pm 1, 2)$ cable of t_i relative to $S \times \{i\}$; that is, t_i^* is the boundary of a Moebius band B_i obtained by giving a half-twist to a thin annulus intersecting $S \times \{i\}$ transversely in a core circle isotopic to t_i. Let $R = \beta \times I$ be a rectangle in $S \times I$ such that $(B_0 \cup B_1) \cap R = (\partial B_0 \cup \partial B_1) \cap R = \partial\beta \times I$, and such that β is a monotone arc in $S \times I$. Now let $K(t_0^*, t_1^*, R)$ be the boundary of $B_0 \cup R \cup B_1$ (see Fig. 1). With this notation, the following theorem summarizes our main result.

Theorem (1.1). *Let K be a crosscap number two knot in S^3. If K admits a $(1,1)$ decomposition, then K is either a torus knot, a 2-bridge knot, a satellite knot, or a knot of the form $K(t_0^*, t_1^*, R)$.*

The families of 2-bridge knots and tunnel number one satellite knots, both of which admit $(1,1)$ decompositions, are of independent interest, and we classify those having crosscap number two explicitly; we note here (see Section 3) that the exterior X_K of a tunnel number one satellite knot $K \subset S^3$ can be decomposed as the union $X_L \cup_T X_{K_0}$ for some 2-bridge link L and torus knot K_0. We call any (p, q) torus knot with $|p| = 2$ or $|q| = 2$ a 2-*torus knot.*

Theorem (1.2). *Let K be a crosscap number two knot in S^3; then,*

(a) *K is a 2-bridge knot iff K is a plumbing of an annulus and a Moebius band, i.e., iff K is of the form $(2m(2n+1)-1)/(2n+1)$ for $m \neq 0$ (see Fig. 2(a));*

(b) *K is a tunnel number one satellite knot, with $X_K = X_L \cup_T X_{K_0}$, iff, for some integer m, either*

  (i) *K_0 is any nontrivial torus knot and L is the $4(4m+2)/(4m+1)$ or $8(m+1)/(4m+3)$ 2-bridge link (see Fig. 2(b),(c)),*

  (ii) *K_0 is any nontrivial 2-torus knot and L is the $(8m+6)/(2m+1)$ 2-bridge link (see Fig. 2(d)).*

We remark that in Theorem (1.2) (b) the knot K is an iterated torus knot iff $m = 0$; in such case, combining the classifications of crosscap number two cable knots in [15] and of tunnel number one cable knots in [2], it follows that K must be an iterated torus knot of the form $[(4pq \pm 1, 4), (p, q)]$ or $[(6p \pm 1, 3), (p, 2)]$ for some integers p, q.

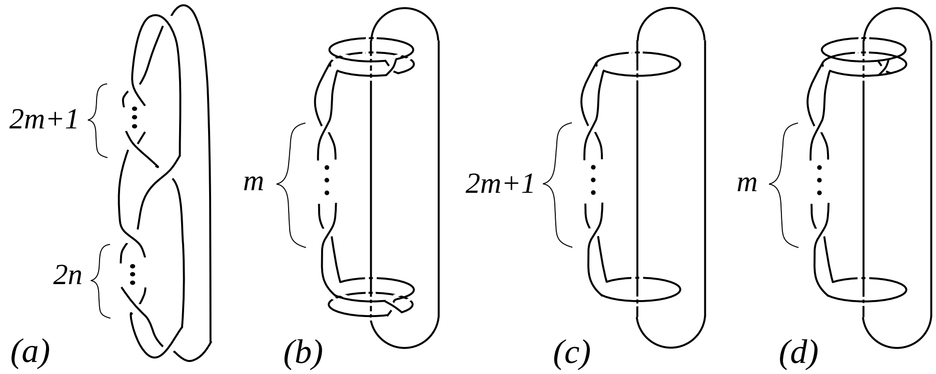


Figure 2. The knots and links of Theorem (1.2); the integers denote half-twists.

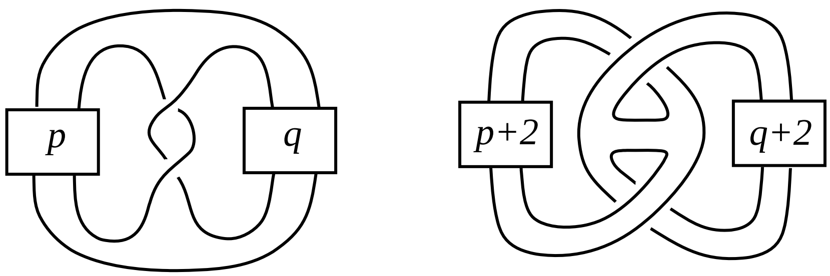


Figure 3. The $(p, q, 2)$ pretzel knot.

Examples of $(1, 1)$ knots of the form $K(t_0^*, t_1^*, R)$ which are neither torus, 2-bridge, nor satellites are provided by the $(p, q, \pm 2)$ pretzel knots with p, q odd integers distinct from ± 1, as shown in Fig. 3; in fact, by [11], these are the only tunnel number one pretzel knots which are not 2-bridge.

Finally, the crosscap number two torus knots are also classified in [15]; the crosscap number of torus knots in general are determined in [14].

Theorem (1.3) ([15]). *A (p, q) torus knot has crosscap number two iff (p, q) or (q, p) is of the form $(3, 5)$, $(3, 7)$, or $(2(2m + 1)n \pm 1, 4n)$ for some integers $m, n,\ n \neq 0$.* □

We will work in the smooth category. In Section 2 we discuss (g, n) decompositions for knots in S^3 and prove Theorem (1.2) (a). This first case, involving 2-bridge knots, has a pleasant solution arising directly from the classification of π_1-injective surfaces in 2-bridge knot exteriors by Hatcher and Thurston [7]; we will follow and extend the basic ideas of [4, 5, 7] to handle the remaining cases along similar lines, via Morse position of essential surfaces relative to a Heegaard surface product structure. In the process it becomes necessary to deal with essential surfaces Σ in knot or link exteriors, all satisfying $\chi(\Sigma) = -1$. Section 3 improves slightly on the theme of [5] to allow for nonorientable essential surfaces in a 2-bridge link exterior; this is the content of Lemma (3.4), which

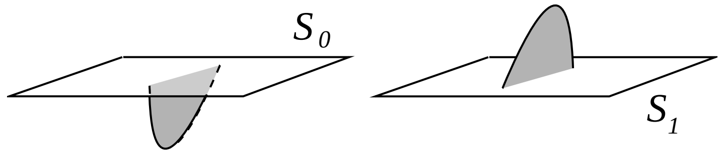


Figure 4.

leads to the proof of Theorem (1.2) (b). To handle the case of knots with a $(1,1)$ decomposition along the same lines it is necessary to prove a statement similar to Lemma (3.4); this is done in Section 4, where Lemma (4.1) is established and which, along with some results from [9], leads to a proof of Theorem (1.1). Since a once-punctured torus Σ also satisfies $\chi(\Sigma) = -1$, the results of this paper can be modified to obtain the classification of genus one knots in S^3 with a $(1,1)$ decomposition as well.

We want to thank Mario Eudave-Muñoz for making his preprint [3] accessible to us, which motivated the line of argument used in Lemma (4.1).

2. (g, n) decompositions and 2-bridge knots

A knot or link L in S^3 is said to be *of type* (g, n) if there is a genus g Heegaard splitting surface S in S^3 bounding handlebodies H_0, H_1 such that, for $i = 0, 1$, L intersects H_i transversely in a trivial n-string arc system. Let $S \times I$ be a product regular neighborhood of S in S^3 and let $h : S \times I \to I$ be the natural projection map. We denote the level surfaces $h^{-1}(r) = S \times \{r\}$ by S_r for each $0 \leq r \leq 1$, and assume that $S_0 \subset H_0, S_1 \subset H_1$, and that $h|S \times I \cap L$ has no critical points (so $S \times I \cap L$ consists of monotone arcs).

Let F be an essential surface properly embedded in the exterior $X_L = S^3 \setminus \operatorname{int} N(L)$ of L; such a surface can always be isotoped in X_L so that:

(M1) F intersects $S_0 \cup S_1$ transversely; we denote the surfaces $F \cap H_0, F \cap H_1, F \cap S \times I$ by $F_0, F_1, \widetilde{F}$, respectively;

(M2) each component of ∂F is either a level meridian circle of ∂X_L lying in some level set S_r or it is transverse to all the level meridians circles of ∂X_L in $S \times I$;

(M3) for $i = 0, 1$, any component of F_i containing parts of L is a *cancelling disk* for some arc in $L \cap H_i$ (see Fig. 4); in particular, such cancelling disks are disjoint from any arc of $L \cap H_i$ other than the one they cancel;

(M4) $h|\widetilde{F}$ is a Morse function with a finite set $Y(F)$ of critical points in the interior of $\widetilde{F}$, located at different levels; in particular, $\widetilde{F}$ intersects each noncritical level surface transversely.

We define the complexity of any surface F satisfying (M1)–(M4) as the number

$$c(F) = |\partial F_0| + |\partial F_1| + |Y(F)|,$$

where $|Z|$ stands for the number of elements in the finite set Z, or the number of components of the topological space Z.

We say that F is *meridionally incompressible* if whenever F compresses in S^3 via a disk D with $\partial D = D \cap F$ such that D intersects L transversely in one point interior to D, then ∂D is parallel in F to some boundary component of F which is a meridian circle in ∂X_L; otherwise, F is *meridionally compressible*. Observe that if F is essential and meridionally compressible then a 'meridional surgery' on F produces a new essential surface in X_L.

In the sequel we will concentrate in the case of knots and 2-component links L of types $(0,2)$ or $(1,1)$ and certain essential surfaces F in X_L with $\chi(F) = -1$. We close this section with a proof of the first part of Theorem (1.2).

Proof of Theorem (1.2) (a). Suppose K is a 2-bridge knot with a $(0,2)$ decomposition relative to some 2-sphere S in S^3. In this context, it is proved in [7, Lemma 2] that once F has been isotoped so as to satisfy (M1)–(M4) with minimal complexity, then F lies in $S \times I$ except for the cancelling disk components of $F_0 \cup F_1$, $h|\widetilde{F}$ has only saddle critical points, $F \cap S_r$ has no circle components for any r, and each saddle joins distinct level arc components. As $\chi(F) = -1$ and F_0 consists of two cancelling disks only, $h|\widetilde{F}$ has exactly three critical points, so F is a plumbing of an annulus and a Moebius band by [7]; thus that K must be a 2-bridge knot of the form $(2m(2n+1)-1)/(2n+1)$ for $m \neq 0$ follows from Fig. 2(a), and the claim follows. $\square$

3. Satellite knots

In this section we assume that K is a tunnel number one satellite knot in S^3 of crosscap number two. By [2, 10], the exterior $X_K = S^3 \setminus \operatorname{int} N(K)$ of K can be decomposed as $X_L \cup_T X_{K_0}$, where $X_L = S^3 \setminus \operatorname{int} N(L)$ is the exterior of some 2-bridge link $L \subset S^3$ other than the unlink or the Hopf link and $X_{K_0} = S^3 \setminus \operatorname{int} N(K_0)$ is the exterior of some nontrivial torus knot $K_0 \subset S^3$, glued along a common torus boundary component T in such a way that a meridian circle of L in T becomes a regular fiber of the Seifert fibration of X_{K_0}.

If F is any once-punctured Klein bottle, then any orientation preserving nontrivial circle embedded in F either cuts F into a pair of pants, splits off a Moebius band from F, or is parallel to ∂F; in the first case we call such circle a *meridian* of F, while in the second case we call it a *longitude* (cf. [12, §2]). Notice any meridian and longitude circles of F intersect nontrivially.

As mentioned in the Introduction, K has crosscap number two iff its exterior X_K contains a properly embedded essential once-punctured Klein bottle F, in which case K is not a 2-torus knot and F has integral boundary slope. We first show the existence of some once-punctured Klein bottle in X_K which intersects the torus T transversely in a simple way.

LEMMA (3.1). *Let K be a tunnel number one satellite knot in S^3 of crosscap number two. Then there is an essential once-punctured Klein bottle $F \subset X_K = X_L \cup_T X_{K_0}$ which intersects T transversely and such that either:*

(i) *F lies in X_L, or*

(ii) *$F \cap X_L$ is a once-punctured Moebius band F_L and $F \cap X_{K_0}$ is a Moebius band; in particular, K_0 is a 2-torus knot.*

Proof. Let F be an essential once-punctured Klein bottle in X_K, necessarily having integral boundary slope; we may assume F has been isotoped so as to intersect T transversely and minimally. Hence $T \cap F$ is a disjoint collection of circles which are nontrivial and orientation preserving in both T and F, so each such circle is either a meridian or longitude of F, or parallel to ∂F in F. Thus, the closure of any component of $F \setminus T$ is either an annulus, a Moebius band, a once-punctured Moebius band, a pair of pants, or a once punctured Klein bottle.

Suppose $\gamma \subset T \cap F$ is a component parallel to ∂F in F; let ρ denote the slope of a fiber of X_{K_0} in T. Then the component of $F \cap X_L$ containing ∂F is an annulus with the same boundary slope as γ on T. If the slope of γ on T is integral then K is isotopic to K_0, which is not the case; thus γ has nonintegral slope on T and so K is a tunnel number one iterated torus knot with $\Delta(\gamma, \rho) = 1$ by [2, Lemma 4.6].

In particular, as any component of $F \cap X_{K_0}$ must be incompressible and not boundary parallel in X_{K_0} by minimality of $T \cap F$, no such component can be an annulus, a Moebius band, or a pair of pants. Therefore, $F' = F \cap X_{K_0}$ is a once-punctured Klein bottle in X_{K_0} with nonintegral boundary slope γ on T, and so, by [12, Lemma 4.5], F' must boundary compress in X_{K_0} into a Moebius band B such that $\Delta(\partial F', \partial B) = 2$. But then K_0 is a 2-torus knot and ∂B is a fiber of X_{K_0}, so $\Delta(\gamma, \rho) = 2$, which is not the case. Therefore, no component of $F \cap T$ is parallel to ∂F in F, so either $T \cap F$ is empty and (i) holds or its components are either all meridians or all longitudes of F. We now deal with the last two options.

Case (3.2). *The circles $T \cap F$ are all meridians of F.*

Then the component P of $F \cap X_L$ containing ∂F is a pair of pants with two boundary components c_1, c_2 on T. If A is an annulus in T cobounded by c_1, c_2, then $P \cup A$ is necessarily a once-punctured Klein bottle for K which, after pushing slightly into X_L, satisfies (i).

Case (3.3). *The circles $T \cap F$ are all longitudes of F.*

If the circles $T \cap F$ are not all parallel in F then there are two components of $F \setminus T$ whose closures are disjoint Moebius bands B_1, B_2 with boundaries on T. But then, if A is an annulus in T cobounded by $\partial B_1, \partial B_2$, the surface $B_1 \cup B_2 \cup A$ is a closed Klein bottle in $X_K \subset S^3$, which is not possible. Hence the circles $T \cap F \subset F$ are mutually parallel in F, and so the component of $F \cap X_L$ which contains ∂F is a once-punctured Moebius band F_L. Moreover, there is a component of $F \setminus T$ whose closure is a Moebius band B, properly embedded in X_L or X_{K_0}. If B lies in X_L then $F \cap X_{K_0}$ is a nonempty collection of disjoint essential annuli in X_{K_0}, hence ∂B is the meridian circle of a component of L, which implies that B closes into a projective plane in S^3, an impossibility. Therefore B lies in X_{K_0}, and if A is an annulus in T cobounded by ∂F_L and ∂B then $F_L \cup A \cup B$ can be isotoped into a once punctured Klein bottle for K satisfying (ii). □

Denote the components of L by K_1, K_2, with ∂F isotopic to K_1. We assume that a fixed 2-bridge presentation L is given relative to some 2-sphere S in S^3, and that F has been isotoped so as to satisfy (M1)–(M4) and have minimal complexity. Notice that H_0, H_1 are 3-balls in this case. The next result will be useful in the sequel.

LEMMA (3.4). *Let Σ' be a surface in S^3 spanned by K_1 (orientable or not) and transverse to K_2, such that $\Sigma = \Sigma' \cap X_L$ is essential and meridionally*

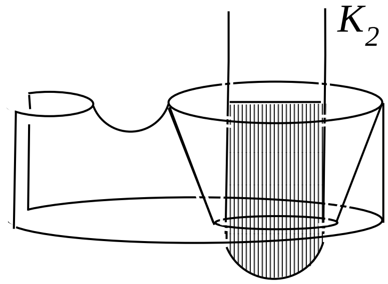


Figure 5.

incompressible in X_L. If Σ is isotoped so as to satisfy (M1)–(M4) with minimal complexity, then $|Y(\Sigma)| = 2 - (\chi(\Sigma) + |\partial\Sigma|)$, and

(i) *each critical point of $h|\widetilde{\Sigma}$ is a saddle,*
(ii) *for $0 \leq r \leq 1$ any circle component of $S_r \cap \Sigma$ is nontrivial in $S_r \setminus L$ and Σ, and*
(iii) *Σ_0 and Σ_1 each consists of one cancelling disk.*

Proof. If Σ is orientable the statement follows from the proof of [5, Theorem 3.1] without any constraints on the boundary of Σ. If Σ is nonorientable, the given hypothesis on Σ are sufficient for the arguments of [4, Proposition 2.1] and [7, Lemma 2] to go through and establish (i)–(iii); the meridional incompressibility condition is needed only for (iii), as in [5, Theorem 3.1], while the fact that any circle component of $S_r \cap \Sigma$ is nontrivial in Σ follows by the argument of Lemma (4.1) (ii). That $|Y(F)| = 2 - (\chi(\Sigma) + |\partial\Sigma|)$ follows now from (i) and (iii). □

Proof of Theorem (1.2) (b). We will split the argument into several parts, according to Lemma (3.1).

Case (A): *$F \subset X_L$ and F is meridionally incompressible.*

In this case Lemma (3.4) applies with $\Sigma = F$, so $|Y(F)| = 2$ and $F \cap S_0, F \cap S_1$ have no circle components. Let $0 < r_1 < r_2 < 1$ be the levels at which the two saddles of $h|\widetilde{F}$ are located, and let α_0, α_1 denote the arcs $F \cap S_0, F \cap S_1$, respectively. For any level $0 < r < 1$, any circle component of $F \cap S_r$ either separates or does not separate the points $S_r \cap K_2$; the first option is not possible by Lemma (3.4) (ii) since F is meridionally incompressible, while in the second option it is not hard to see that, with the aid of the cancelling disk F_0, F compresses in X_L along one such level circle (see Fig. 5).

Hence $S_r \cap F$ has no circle components for $0 \leq r \leq 1$, so the saddles, when seen from bottom to top and top to bottom, join the arcs α_0, α_1, respectively, in a nonorientable fashion (see Fig. 6(a)) and so, for a sufficiently small $\varepsilon > 0$, $B_1 = F \cap S \times [r_1 - \varepsilon, r_1 + \varepsilon]$ and $B_2 = F \cap S \times [r_2 - \varepsilon, r_2 + \varepsilon]$ are Moebius bands in F. For $i = 1, 2$, the core circle C_i of B_i in S_{r_i} necessarily separates the points $K_2 \cap S_{r_i}$, else C_i bounds a disk D_i in S_{r_i} disjoint from K_2 as in Fig. 6(b), and a boundary compression disk for F can be constructed from the subdisk D_i' of D_i as in Fig. 6(c); also, ∂B_i is a $(\pm 1, 2)$ cable of C_i. Let R be the rectangle $F \cap S \times [r_1 + \varepsilon, r_2 - \varepsilon] \subset F$. As $h|R$ has no critical points, there exists an embedded

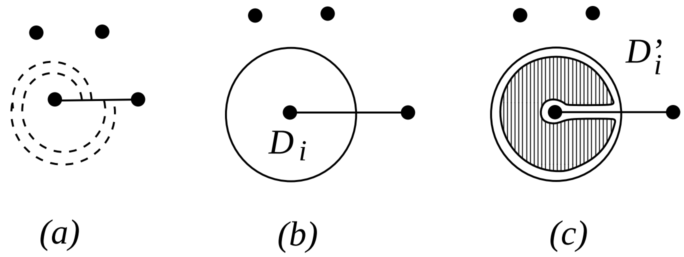


Figure 6. Boundary compression of F.

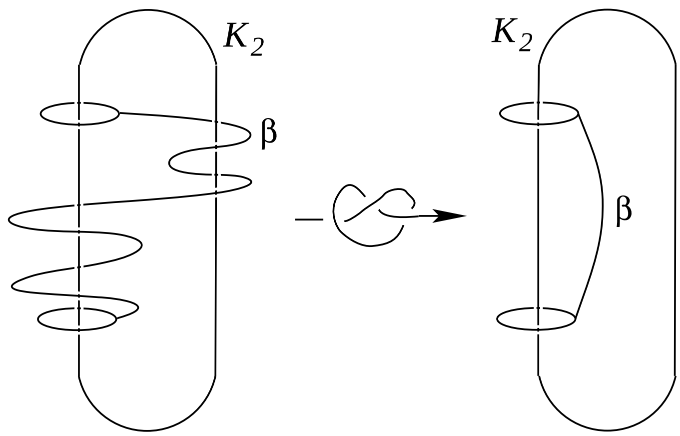


Figure 7. Isotoping the arc β.

arc β in R with one endpoint in ∂B_1 and the other in ∂B_2, and such that $h|N(\beta)$ has no critical points for some small regular neighborhood $N(\beta)$ of β in R; thus β is monotone. As the once-punctured Klein bottle $F' = B_1 \cup N(\beta) \cup B_2$ is isotopic in X_{K_2} to F, it follows that the link L has the form of Fig. 2(b) up to isotopy (see Fig. 7), and hence that L is a $4(4m+2)/(4m+1)$ 2-bridge link. $\square$(Case (A))

Case (B): *$F \subset X_L$ and F is meridionally compressible.*

Observe that if F meridionally compresses along a circle $\gamma \subset F$ then γ must be a meridian circle of F: for if γ is trivial in F then a 2-sphere in S^3 can be constructed which intersects K_2 in one point, if γ is a longitude in F then S^3 contains RP^2, and if γ is parallel to ∂F then L is the Hopf link. Thus, F meridionally compresses into an essential pair of pants Δ in X_L, which is necessarily meridionally incompressible. By Lemma (3.4), we may therefore assume that Δ satisfies (M1)–(M4) and lies within the region $S \times I$ except for the cancelling disks Δ_0, Δ_1, and $|Y(\Delta)| = 0$.

Since Δ is orientable, the saddles must join the corresponding arcs $\alpha_0 = \Delta \cap S_0, \alpha_1 = \Delta \cap S_1$ to themselves in an orientable fashion or to a level circle component, when seen from bottom to top and top to bottom, respectively. Let C_1, C_2 be the two level boundary circles of Δ, and let C_3, C_4 be the *limiting* circles in the saddle levels (see Fig. 8); we assume that, for $1 \leq i \leq 4$, the C_i's

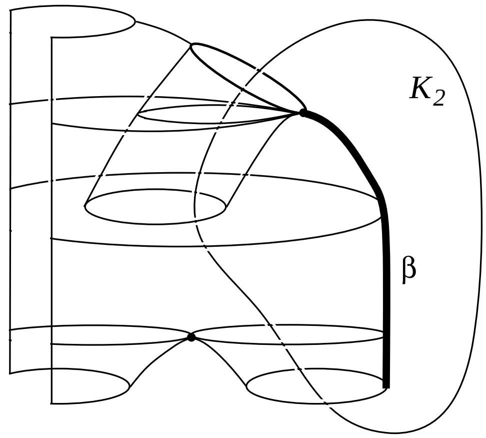


Figure 8.

are located at distinct levels r_i, respectively. If r_j and r_k are the lowest and highest levels in this list, respectively, then there exists an embedded arc β in $\widetilde{\Delta}$ with one endpoint in C_j and the other in C_k, such that $h|N(\beta)$ has no critical points for some small regular neighborhood $N(\beta)$ of β in $\widetilde{\Delta}$ (see Fig. 8). Then a small regular neighborhood $N(C_j \cup \beta \cup C_k)$ in Δ yields a 2-punctured disk with boundary isotopic to K_1 in X_{K_2}. As in Case (A), it follows that L can be isotoped into the form of Fig. 2(c), so L is a $8(m+1)/(4m+3)$ 2-bridge link. □(Case (B))

Therefore part (i) holds when $F \subset X_L$. We now handle the last possible case.

Case (C): $F \cap X_L = F_L$.

As for any level $0 \leq r \leq 1$ each circle component of $F_L \cap S_r$ is either parallel to the boundary circle of F_L isotopic to K_1, or parallel to the boundary circle of F_L which is a level meridian of K_2, and L is neither the unlink nor the Hopf link, it follows that F_L is incompressible and meridionally incompressible, hence Lemma (3.4) applies. Therefore, the method of proof used in Case (B) above immediately implies that L is isotopic to a link of the form of Fig. 2(d), hence (ii) holds in this case.

Since clearly any knot constructed as above has crosscap number two, the theorem follows. □

4. Knots with $(1,1)$ decompositions

In this section we assume that K is a crosscap number two knot in S^3 admitting a $(1,1)$ decomposition relative to some Heegaard torus S of S^3. In this case the handlebodies H_0, H_1 are solid tori with meridian disks of slope μ_0, μ_1 in S_0, S_1, respectively. For $\{i,j\} = \{0,1\}$, we project μ_j onto S_i, continue to denote such projection by μ_j, and frame S_i via the circles μ_i, μ_j, so that a (p,q)-*circle* in S_i means a circle embedded in S_i isotopic to $p\mu_i + q\mu_j$; thus S_i gets the standard framing as the boundary of the exterior of the core of H_i, and a (p,q)-circle in S_0 is isotopic in $S \times I$ to a (q,p)-circle in S_1.

Before studying the associated essential once-punctured Klein bottle for K, we prove a statement similar to Lemma (3.4) in the present context.

LEMMA (4.1). *Suppose K is not a torus knot. Let Σ' be a spanning surface for K in S^3 (orientable or not) such that $\Sigma = \Sigma' \cap X_K$ is essential in X_K. If Σ is isotoped so as to satisfy (M1)–(M4) with minimal complexity, then $|Y(\Sigma)| = 1 - \chi(\Sigma)$, and*

- (i) *each critical point of $h|\widetilde{\Sigma}$ is a saddle,*
- (ii) *for $0 \leq r \leq 1$ any circle component of $S_r \cap \Sigma$ is nontrivial in $S_r \setminus K$ and Σ, and not parallel in Σ to $\partial\Sigma$,*
- (iii) *for $i = 0, 1$ Σ_i consists of one cancelling disk and either one Moebius band and some annuli components, or a collection of disjoint annuli each having boundary slope (p_i, q_i) in S_i with $|q_i| \geq 2$, and*
- (iv) *the saddle closest to either the 0-level or 1-level does not join circle components.*

Proof. Part (i) follows from the argument of [4, Proposition 2.1].

Suppose now that γ is a circle component of $S_r \cap \Sigma$ for some level $0 \leq r \leq 1$. If γ bounds a disk D in $S_r \setminus K$ then γ bounds a disk D' in Σ, since Σ is incompressible in X_K. Construct a surface Σ'' isotopic to Σ from $(\Sigma \setminus D') \cup D$ by pushing D slightly above or below S_r so that Σ'' satisfies (M1)–(M4) and the singularities of $h|\widetilde{\Sigma}''$ are exactly those of $h|\widetilde{\Sigma} \setminus D'$ with an additional local extremum in the interior of D; thus, $h|\widetilde{\Sigma}''$ has at most $|Y(\Sigma)| + 1$ critical points.

If D' is disjoint from $S_0 \cup S_1$ then D' lies in $S \times I$ and, since $\partial D'$ is level, $h|D'$ has a local extremum in $\operatorname{int} D'$, contradicting (i). If D' intersects $S_0 \cup S_1$ then $|\partial\Sigma''_0| + |\partial\Sigma''_1| < |\partial\Sigma_0| + |\partial\Sigma_1|$ while $|Y(\Sigma'')| \leq |Y(\Sigma)| + 1$, hence $c(\Sigma'') \leq c(\Sigma)$ and so $c(\Sigma'') = c(\Sigma)$ by minimality of $c(\Sigma)$, again contradicting (i). Therefore, γ is nontrivial in $S_r \setminus K$ and, since K is not a torus knot, γ is not parallel in Σ to $\partial\Sigma$. Thus it only remains to verify that γ is nontrivial in Σ for (ii) to hold, which we will do by the end of the proof.

If some component of Σ_0, other than the cancelling disk, compresses in H_0, then there is one such component σ which compresses in H_0 via a disk D disjoint from all other components of Σ_0. Since Σ is essential in X_K, ∂D bounds a disk D' in Σ. Let $\Sigma'' = (\Sigma \setminus D') \cup D$. Then $h|\widetilde{\Sigma}''$ has at most $|Y(\Sigma)|$ singular points and, since $\operatorname{int} D'$ necessarily intersects $S_0 \cup S_1$, $|\partial\Sigma''_0| + |\partial\Sigma''_1| < |\partial\Sigma_0| + |\partial\Sigma_1|$ and so $c(\Sigma'') < c(\Sigma)$, an impossibility. Therefore, any component of Σ_0 is incompressible in H_0, hence it must be either an annulus, a Moebius band, or a disk; since H_0 is a solid torus, Σ_0 may have at most one Moebius band component.

Suppose Σ_0 has an annulus component σ; then σ separates H_0 into two pieces, one of which contains the cancelling disk component of Σ_0. If σ is parallel in H_0 into S_0 away from all other components of Σ_0 then σ can be pushed into the region $S \times I$; notice this is the case if the slope of σ in S_0 is of the form (p_0, q_0) with $|q_0| = 1$. It is then possible to isotope σ and Σ appropriately, so that $h|\sigma$ has one saddle and one local minimum and Σ continues to satisfy (M1)–(M4); hence $|\partial\Sigma_0|$ will decrease by two while $|Y(\Sigma)|$ will increase by two, and so $c(\Sigma)$ will remain minimal. However, this time $h|\widetilde{\Sigma}$ has a local minimum critical point in σ, contradicting (i). Therefore, since σ is incompressible in H_0, any boundary

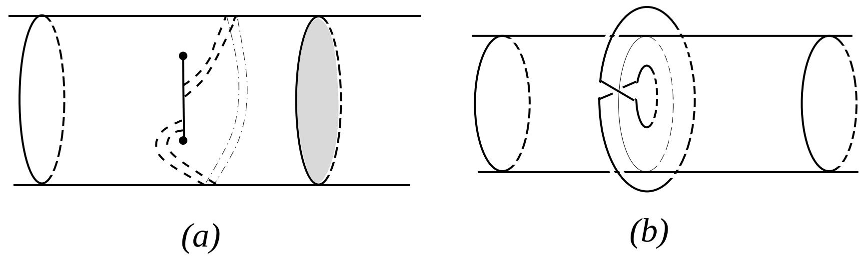


Figure 9.

component of σ must be nontrivial in $S_0 \setminus K$ and distinct from μ_0, so it follows that the boundary slope of σ in S_0 is of the form (p_0, q_0) with $|q_0| \geq 2$.

Consider the first saddle above level 0; if it joins a circle component γ of $\Sigma \cap S_0$ to itself or to another such circle component then it is possible to lower the saddle below level S_0 while satisfying (M1)–(M4), thus reducing the value of $c(\Sigma)$, which is not possible. Hence (iv) holds, and the first saddle above level 0 joins the arc component α_0 of $S_0 \cap \Sigma$ to itself or to a circle component.

Suppose now that σ is a disk component of Σ_0 other than the cancelling disk; then σ is either a trivial disk or a meridian disk of H_0. In the first case, σ separates H_0 into a 3-ball B^3 and a solid torus, with the cancelling disk of Σ_0 contained in B^3 by the first part of (ii); we may further assume that $\partial\sigma$ and α_0 are adjacent in S_0. Consider the first saddle above level 0. If it joins the arc component α_0 of $\Sigma \cap S_r$ to itself then either a Moebius band is created by the saddle with core a circle bounding a disk in the saddle level, so Σ is boundary compressible (see Fig. 6), or a trivial circle component is created in a level slightly above the saddle level, contradicting the first part of (ii). If the saddle joins $\partial\sigma$ to α_0 then pushing down the saddle slightly below level 0 isotopes Σ so as to still satisfy (M1)–(M4) but lowers its complexity. Since by (iv) these are the only possibilities for the first saddle, if Σ_0 contains any disk components other than the cancelling disk then all such components are meridian disks of H_0. The analysis of the possible scenarios for the first saddle above level 0 is similar to that of the previous cases, except for when the saddle joins α_0 to itself as in Fig. 9(a). In such case, if r is the level of the first saddle above level 0, the Moebius band created by the saddle has as core a circle in S_r which bounds a meridian disk of the solid torus bounded by S_r below the level S_r (see Fig 9(b)). The situation is similar to that of Fig. 6, so Σ is boundary compressible, which is not the case. Hence Σ_0, and similarly Σ_1, has no such disk components and (iii) holds.

Now let $0 \leq r \leq 1$ and γ be any circle component of $(S_0 \cup S_r \cup S_1) \cap \Sigma$. If γ is trivial and innermost in Σ then it bounds a subdisk D in Σ with interior disjoint from $S_0 \cup S_r \cup S_1$, hence D lies either in Σ_0, Σ_1, or $S \times I$. But, as shown above, neither Σ_0 nor Σ_1 have disk components other than the cancelling disks, and if D lies in $S \times I$ then, as $\partial D = \gamma$ is level, $h|D$ must have a local extremum in $\operatorname{int} D$, contradicting (i). Hence γ is nontrivial in Σ and so the proof of (ii) is complete. That $|Y(\Sigma)| = 1 - \chi(\Sigma)$ now follows from (i) and (iii). □

In preparation for the proof of Theorem (1.1), the following result specializes Lemma (4.1) to the case when Σ is a once punctured Klein bottle F; its first part is a slight generalization of a construction by Matsuda in [9, pp. 2161–2162]. We will say that an essential annulus A properly embedded in $S \times I$ is an F-*spanning annulus* if A can be isotoped so as to be disjoint from the component of $\widetilde{F} = F \cap S \times I$ containing parts of K, and its boundary slope in S_0 is of the form (p,q) for some $|p|, |q| \geq 2$. Notice that an F-spanning annulus A is isotopic in $S \times I$ to the annulus $(\partial A \cap S_0) \times I$, and its boundary component in S_1 has slope (q,p).

Lemma (4.2). *Let F be an essential once-punctured Klein bottle spanned by K which has been isotoped so as to satisfy (M1)–(M4) with minimal complexity. If there is an F-spanning annulus in $S \times I$ having boundary slope (p,q) in S_0 then K is either a (p,q) torus knot or a satellite of a (p,q)-torus knot; otherwise, $F \cap (S_0 \cup S_1)$ has at most two circle components.*

Proof. Let F' denote the component of $\widetilde{F} = F \cap S \times I$ containing parts of K. Let A be an F-spanning annulus with boundary slope (p,q) in S_0, and suppose K is not a (p,q) torus knot. By Lemma (4.1) (ii),(iii), F' is either a once-punctured Moebius band or a pair of pants embedded in the solid torus $V = S \times I \setminus \operatorname{int} N(A)$, where $N(A)$ is a small regular neighborhood of A in $S \times I$. In either case, $\partial F'$ has one component $K' \subset \operatorname{int} V$ which is isotopic to K in S^3, and one or two more components embedded in ∂V, each running once around V. Notice that V is a regular neighborhood of a (p,q) torus knot, so K' is not a core of V.

If F' is a once-punctured Moebius band then K' is a nontrivial knot in V with odd winding number. If F' is a pair of pants then, by Lemma (4.1) (ii),(iii), the closure of $F \setminus F'$ consists either of two Moebius bands or an annulus with core a meridian circle of F. In the first case F_0 and F_1 each have a Moebius band component which, due to the presence of the spanning annulus A, have boundary slopes (p,q) and (q,p), respectively, an impossibility since then $|p| = |q| = 2$; in the latter case, the closure of the annulus $F \setminus F'$ intersects V in annuli running once around V, thus it can be isotoped in S^3, away from F', into $S^3 \setminus \operatorname{int} V$, and so the components of $\partial F'$ other than K' must be coherently oriented in ∂V; therefore K' has winding number two in V and hence it is a nontrivial satellite of the core of V. The first part of the lemma follows.

Suppose now that $F \cap (S_0 \cup S_1)$ has at least three circle components; if, say, three such components lie in S_0, or at least two lie in S_0 and at least one in S_1, then, since $|Y(F)| = 2$ by Lemma (4.1) and the saddles do not join circle components, at least one of the circle components of $F \cap S_0$ must flow along an annulus component of $\widetilde{F}$ from S_0 to S_1 without interacting with the saddles. Thus $\widetilde{F}$ has at least one annulus component which, by Lemma (4.1) (iii), has boundary slope of the form (p,q) in S_0 for some $|p|, |q| \geq 2$, and so must be an F-spanning annulus. Thus the second part of the lemma follows. □

Proof of Theorem (1.1). Let K be a crosscap number two knot in S^3, and let F be an essential once-punctured Klein bottle spanned by K; we assume F has been isotoped so as to satisfy (M1)–(M4) with minimal complexity. To simplify notation, let F_0', F_1' denote the components of F_0, F_1, respectively, other than

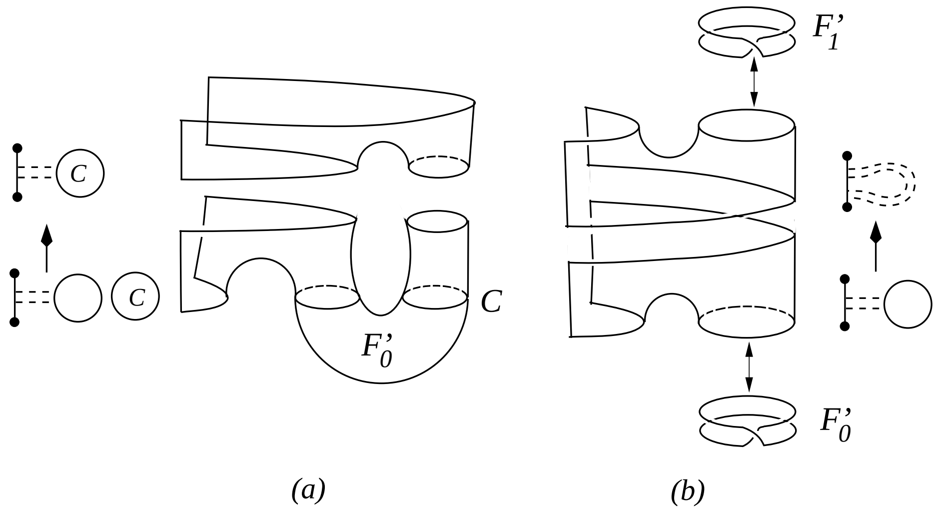


Figure 10. Construction of F via saddles in Cases (A) and (B).

the cancelling disks. By Lemma (4.2), if K is neither a torus nor a satellite knot then $S \times I$ contains no F-spanning annuli and $F \cap (S_0 \cup S_1)$ has at most two circle components; thus, without loss of generality, F_0' and F_1' fit in one of the following cases.

Case (A): *F_0' is an annulus and F_1' is empty.*

Fig. 10(a) shows the only possible construction (abstractly) of the surface F, starting from F_0, via the two saddles of $h|\widetilde{F}$. By Lemma (4.1) (iii), the boundary slope of the annulus F_0' in S_0 is of the form (p, q) with $|q| \geq 2$. It is not hard to see that the boundary circle C of the annulus $\partial F_0'$ in Fig. 10(a) bounds an essential annulus A in $S \times I \setminus F$, hence $|p| = 1$ since $S \times I$ has no F-spanning annuli, so K is a 2-bridge knot by the argument of [9, pp. 2161–2162].

Case (B): *Both F_0' and F_1' are Moebius bands.*

The only possibility in this case is the one shown (abstractly) in Fig. 10(b): for otherwise, by Lemma (4.1) (iv), the first saddle above the 0-level would join the arc component of $S_0 \cap F$ with itself, necessarily in an orientable fashion, and so $S_r \cap F$ would have two circle components for any level r in between the saddle levels; but then the first saddle below the 1-level must join the circle component of $S_1 \cap F$ with itself, contradicting Lemma (4.1) (iv).

Hence $\widetilde{F}$ is a pair of pants and, since all the critical points of $h|\widetilde{F}$ are saddles, there exists an embedded arc β in $\widetilde{F}$ with one endpoint in $\partial F_0'$ and the other in $\partial F_1'$ which is monotone in $S \times I$ and such that $h|R$ has no critical points for some small regular neighborhood R of β in $\widetilde{F}$. Observe that, for $i = 0, 1$, if $\partial F_i'$ is a $(p_i, 2)$-circle in S_i, then F_i' is isotopic in S^3 to a Moebius band B_i which is a $(1, 2)$ cable of a $(p_i, 1)$-circle t_i in S_i. Therefore the once-punctured Klein bottle $F_0' \cup R \cup F_1'$ can be isotoped into $B_0 \cup R' \cup B_1$ for some monotone subrectangle R' of R. As $F_0' \cup R' \cup F_1'$ is isotopic to F in S^3, it follows that K is a knot of the form $K(t_0^*, t_1^*, R')$.

Case (C): *Both F_0' and F_1' are empty.*

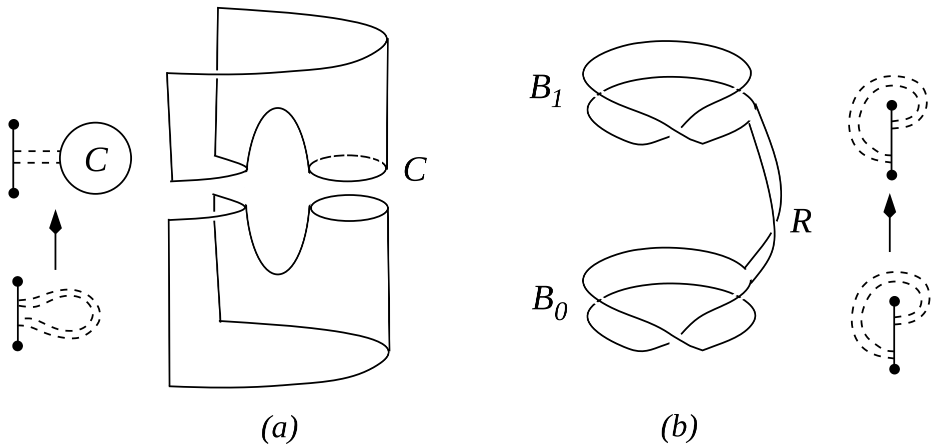


Figure 11. Construction of F via saddles in Case (C).

In this case the saddles, when read from bottom to top and top to bottom, must join the arcs $S_0 \cap F, S_1 \cap F$ with themselves, respectively, both in an orientable fashion or both in a nonorientable fashion; the possible cases are described (abstractly) in Fig.11. In the case of Fig.11(a), if the level circle C has slope (p, q) relative to S_0, then there is an essential annulus in $S \times I \setminus \widetilde{F}$ with boundary slope (p, q) in S_0. Hence $|p| = 1$ or $|q| = 1$ since $S \times I$ has no F-spanning annuli, so K is a 2-bridge knot by the argument of [9, pp. 2161–2162].

In the case of Fig.11(b) let $0 < r_1 < r_2 < 1$ be the saddle levels and, for $i = 1, 2$, let B_i be the Moebius band $F \cap S \times [r_i - \varepsilon, r_i + \varepsilon]$ for a sufficiently small $\varepsilon > 0$. Then $F \cap S \times [r_1 + \varepsilon, r_2 - \varepsilon]$ is a rectangle R, and $B_1 \cup R \cup B_2$ is a once-punctured Klein bottle isotopic to F in S^3. Hence K is a knot of the form $K(t_0^*, t_1^*, R')$, where t_i is the core of the Moebius band B_i in the level r_i and R' is a monotone subrectangle of R.

Case (D): *F_0' is a Moebius band and F_1' is empty.*

Suppose the first saddle below the 1-level joins the arc component of $F \cap S_1$ with itself in an orientable fashion; then the first saddle above the 0-level necessarily joins the arc component of $F \cap S_0$ with itself in a nonorientable fashion. The situation here is similar to that of Case (A): the circle $\partial F_0'$ bounds an annulus A in $S \times I$ which can be isotoped away from $\widetilde{F}$ (see Fig. 10(a), with $C = \partial F_0'$), hence the slope of ∂A in S_1 must be integral and so K is a 2-bridge knot.

Otherwise, the first saddle below the 1-level, say at level $0 < r_1 < 1$, joins the arc component of $F \cap S_1$ with itself in a nonorientable fashion, while the first saddle above the 0-level joins the arc component of $F \cap S_0$ with the circle $\partial F_0'$. This time the situation is similar to that of Cases (B) and the second part of (C): for a small $\varepsilon > 0$, if B_1 is the Moebius band $F \cap S \times [r_1 - \varepsilon, r_1 + \varepsilon]$, then $R = F \cap S \times [0, r_1 - \varepsilon]$ is a rectangle and $F_0' \cup R \cup B_1$ is a once-punctured Klein bottle isotopic to F in S^3, hence K is a knot of the form $K(t_0^*, t_1^*, R')$, where R' is a monotone subrectangle of R and t_0, t_1 can be described as in Cases (B) and (C), respectively. □

Received July 31, 2003

Final version received April 05, 2004

Enrique Ramírez-Losada
CIMAT
Apdo. Postal 402
36000 Guanajuato, Gto.
México
kikis@cimat.mx

Luis G. Valdez-Sánchez
Department of Mathematical Sciences
University of Texas at El Paso
El Paso, TX 79968
USA
valdez@math.utep.edu

References

[1] B. E. Clark, *Crosscaps and knots*, Internat. J. Math. Math. Sci. **1** (1978), 113–123.
[2] M. Eudave Muñoz, *On nonsimple* 3*-manifolds and* 2*-handle addition*, Topology Appl. **55** (1994), 131–152.
[3] ———, *Incompressible surfaces and* $(1,1)$ *knots*, Preprint, 2003.
[4] W. Floyd and A. Hatcher, *Incompressible surfaces in punctured-torus bundles*, Topology Appl. **13** (1982), 263–282.
[5] ———, *The space of incompressible surfaces in a* 2*-bridge link complement*, Trans. Amer. Math. Soc. **305** (1988), 575–599.
[6] H. Goda and M. Teragaito, *Tunnel number one genus one non-simple knots*, Tokyo J. Math. **22** (1999), 99–103.
[7] A. Hatcher and W. Thurston, *Incompressible surfaces in* 2*-bridge knot complements*, Invent. Math. **79** (1985), 225–246.
[8] K. Ichihara, M. Ohtouge, and M. Teragaito, *Boundary slopes of non-orientable Seifert surfaces for knots*, Topology Appl. **122** (2002), 467–478.
[9] H. Matsuda, *Genus one knots which admit* $(1,1)$*-decompositions*, Proc. Amer. Math. Soc. **130** (2002), 2155–2163.
[10] K. Morimoto and M. Sakuma, *On unknotting tunnels for knots*, Math. Ann. **289** (1991), 143–167.
[11] K. Morimoto, M. Sakuma, and Y. Yokota, *Identifying tunnel number one knots*, J. Math. Soc. Japan **48** (1996), 667–688.
[12] E. Ramírez-Losada and L. G. Valdez-Sánchez, *Once-punctured Klein bottles in knot exteriors*, to appear in Topology Appl.
[13] M. Scharlemann, *There are no unexpected tunnel number one knots of genus one*, Trans. Amer. Math. Soc. **356** (2004), 1385–1442.
[14] M. Teragaito, *Crosscap numbers of torus knots*, Topology Appl. **138** (2004), 219–238.
[15] L. G. Valdez-Sánchez, *Once-punctured Klein bottles in knot exteriors with common boundary slopes*, Preprint, 2004.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

TOPOLOGY OF 3-MANIFOLDS AND A CLASS OF GROUPS II

S.K. ROUSHON

ABSTRACT. This is a continuation of an earlier preprint [17] under the same title. These papers grew out of an attempt to find a suitable finite sheeted covering of an aspherical 3-manifold so that the cover either has infinite or trivial first homology group. With this motivation we defined a new class of groups. These groups are in some sense eventually perfect. Here we prove results giving several classes of examples of groups which do (not) belong to this class. Also we prove some basic results on these groups and state two conjectures. A direct application of one of the conjectures to the virtual Betti number conjecture is mentioned. For completeness, here we reproduce parts of [17].

0. Introduction

The main motivation to this paper and [17] came from 3-manifold topology while trying to find a suitable finite sheeted covering of an aspherical 3-manifold so that the cover has either infinite or trivial first integral homology group. In [15] it was proved that $M^3 \times \mathbb{D}^n$ is topologically rigid for $n > 1$ whenever $H_1(M^3, \mathbb{Z})$ is infinite. Also the same result is true when $H_1(M^3, \mathbb{Z})$ is 0. The remaining case is when $H_1(M^3, \mathbb{Z})$ is nontrivial finite. There are induction techniques in surgery theory which can be used to prove topological rigidity of a manifold if certain finite sheeted coverings of the manifold are also topologically rigid. In the case of manifolds with nontrivial finite first integral homology groups, there is a natural finite sheeted cover, namely, the one which corresponds to the commutator subgroup of the fundamental group.

So we start with a closed aspherical 3-manifold M with nontrivial finite first integral homology group and consider the finite sheeted covering M_1 of M corresponding to the commutator subgroup. If $H^1(M_1, \mathbb{Z}) \neq 0$ or $H_1(M_1, \mathbb{Z}) = 0$ then we are done, otherwise we again take the finite sheeted cover of M_1 corresponding to the commutator subgroup and continue. The group theoretic conjecture (Conjecture 0.2) in this article implies that this process stops in the sense that for some i either $H^1(M_i, \mathbb{Z}) \neq 0$ or $H_1(M_i, \mathbb{Z}) = 0$.

Motivated by the above situation we define the following class of groups.

Definition (0.1). An abstract group G is called *adorable* if $G^i/G^{i+1} = 1$ for some i, where $G^i = [G^{i-1}, G^{i-1}]$, the commutator subgroup of G^{i-1}, and $G^0 = G$. The smallest i for which the above property is satisfied is called the *degree of adorability* of G. We denote it by $doa(G)$.

2000 *Mathematics Subject Classification*: Primary: 20F19, 57M99. Secondary: 20E99.

Keywords and phrases: 3-manifolds, discrete subgroup of Lie groups, commutator subgroup, perfect groups, virtual Betti number conjecture, generalized free product, *HNN*-extensions.

Obvious examples of adorable groups are finite groups, perfect groups, simple groups and solvable groups. The second and third class of groups are adorable groups of degree 0. The free products of perfect groups are adorable (in fact perfect). The nontrivial abelian groups and symmetric groups on $n \geq 5$ letters are adorable of degree 1. Another class of adorable groups are $GL(R) = Lim_{n\to\infty} GL_n(R)$. Here R is any ring with unity and $GL_n(R)$ is the multiplicative group of $n \times n$ invertible matrices. These are adorable groups of degree 1. This follows from the Whitehead lemma which says that the commutator subgroup of $GL_n(R)$ is generated by the elementary matrices and the group generated by the elementary matrices is a perfect group. Also $SL_n(\mathbb{C})$, the multiplicative group of $n \times n$ matrices with complex entries, is a perfect group. In fact we will prove that any connected Lie group is adorable as an abstract group. The full braid groups on more than 4 strings are adorable of degree 1.

We observe the following two elementary facts in the next section.

Theorem (1.8). *A group G is adorable if and only if there is a filtration $G_n < G_{n-1} < \cdots < G_1 < G_0 = G$ of G so that G_i is normal in G_{i-1}, G_{i-1}/G_i is abelian for each i, and G_n is a perfect group.*

Theorem (1.10). *Let H be a normal subgroup of an adorable group G. Then H is adorable if one of the following conditions is satisfied.*

- *G/H is solvable.*
- *for some i, G^i/H^i is abelian.*
- *for some i, G^i is simple.*
- *for some i, G^i is perfect and the group G^i/H^{i+1} does not have any proper abelian normal subgroup.*

Also, the braid groups on more than 4 strings provide examples which show that an arbitrary finite index normal subgroup of an adorable group need not be adorable.

In Section 4 the following result about Lie groups is proved.

Theorem (4.9). *Every connected real or complex Lie group is adorable as an abstract group.*

Below we give some examples of nonadorable groups. Proofs of nonadorability of some of these examples are easy. Proofs for the other examples are given in the next sections.

Some examples of groups which are not adorable are nonabelian free groups and fundamental groups of surfaces of genus greater than 1; for the intersection of a monotonically decreasing sequence of characteristic subgroups of a nonabelian free group consists of the trivial element only. The commutator subgroup of $SL_2(\mathbb{Z})$ is the nonabelian free group on 2 generators. Hence $SL_2(\mathbb{Z})$ is not adorable. Also, by Stallings' theorem, if the fundamental group of a compact 3-manifold has finitely generated nonabelian commutator subgroup which is not isomorphic to the Klein bottle group with infinite cyclic abelianization then the group is not adorable. It is known that most of these 3-manifolds support a hyperbolic metric by Thurston's hyperbolization theorem. It is easy to show that the pure braid group is not adorable as there is a surjection of any pure braid group of more than 2 strings onto a nonabelian free group.

From now on, whenever we give examples of nonadorable groups, we will mention its close relationship with nonpositively curved Riemannian manifolds. This will help us state a general conjecture (Conjecture (0.1)).

The next result gives some important classes of examples of nonadorable groups which are generalized free products $G_1 *_H G_2$ or HNN-extensions $K*_H$. We always assume $G_1 \neq H \neq G_2$ and $K \neq H$.

Theorem (2.3). *Let G be a group.*

If $G = G_1 *_H G_2$ is a generalized free product and $G^1 \cap H = (1)$, then one of the following holds.

- *G^1 is perfect.*
- *G^1 is isomorphic to the infinite dihedral group.*
- *G is not adorable.*

If $G = K*_H = \langle K, t \;\; tHt^{-1} = \phi(H)\rangle$ is an HNN-extension and $G^1 \cap H = (1)$, then G is not adorable.

In the second case and in the last possibility of the first case for $i \geq 1$, the rank of G^i/G^{i+1} is ≥ 2.

In Corollary (2.7) we deduce a more general version of Theorem (2.3) and show that if H is n-step G-solvable (see Definition (2.6)) then in the amalgamated free product case either G is adorable of degree at most $n+1$ or is not adorable and in the HNN-extension case it is always nonadorable.

We will give some more examples (Lemma (2.8) and Example (2.9)) of a class of nonadorable generalized free products and examples of compact Haken 3-manifolds with nonadorable fundamental groups.

At this point, recall that if M is a connected, closed oriented 3-manifold and $\pi_2(M, x) \neq 0$, then by the Sphere theorem (see p. 40 of [9]) there is an embedded 2-sphere in M representing a nonzero element of $\pi_2(M, x)$. Hence M can be written as a connected sum of two nonsimply connected 3-manifolds and thus $\pi_1(M, x)$ is a nontrivial free product. In addition, if we assume that $\pi_1(M, x)$ is not perfect and M is not the connected sum of two projective 3-spaces then by Theorem (2.3) $\pi_1(M, x)$ is not adorable. Thus we see that most closed 3-manifolds with $\pi_2(M, x) \neq 0$ have nonadorable fundamental groups.

The next result is about groups with some geometric assumption. Recall that a torsion free Bieberbach groups is the fundamental group of a Riemannian manifold with sectional curvature equal to 0 everywhere.

Corollary (4.3). *A torsion free Bieberbach group is nonadorable unless it is solvable.*

The following theorem deals with groups under some homological hypothesis. This theorem has an interesting application in knot theory and possibly in 3-manifolds in general also.

Theorem (4.4). *Let G be a group satisfying the following properties.*

- *$H_1(G, \mathbb{Z})$ has rank ≥ 3.*
- *$H_2(G^j, \mathbb{Z}) = 0$ for $j \geq 0$.*

Then G is not adorable. Moreover, G^j/G^{j+1} has rank ≥ 3 for each $j \geq 1$.

The Proposition below is a consequence of the above Theorem.

Proposition (4.7). *A knot group is adorable if and only if it has trivial Alexander polynomial.*

In fact in this case the commutator subgroup of the knot group is perfect. All other knot groups are not adorable. On the other hand any knot complement supports a complete nonpositively curved Riemannian metric [13].

After seeing the preprint [17] Tim Cochran informed me that the Proposition (4.7) was also observed by him in corollary 4.8 of [5].

Note that most of the torsion free examples of nonadorable groups we mentioned above act freely and properly discontinuously (except in the case of the braid groups, which is still an open question) on a simply connected complete nonpositively curved Riemannian manifold. Also, we recall that a solvable subgroup of the fundamental group of a nonpositively curved manifold is virtually abelian [20]. There are generalization of these results to the case of locally $CAT(0)$ spaces [1]. Considering these facts we pose the following conjecture.

Conjecture (0.1). *The fundamental groups of generic class of complete nonpositively curved Riemannian manifolds, or, more generally, of generic class of locally $CAT(0)$ metric spaces, are not adorable.*

One can even ask the same question for hyperbolic groups.

Now we state the conjecture we referred to before. Though in [17] this conjecture was stated for any finitely presented torsion free groups, our primary aim was the following particular case.

Conjecture (0.2). *Let G be a finitely presented torsion free group which is isomorphic to the fundamental group of a closed aspherical 3-manifold such that G^i/G^{i+1} is a finite group for all i. Then G is adorable.*

Using Theorem (3.1) in Section 3 it is easy to show that the above conjecture is true for aspherical Seifert fibered spaces. In fact we will show that most Seifert fibered spaces have nonadorable fundamental groups.

Here note that a partial converse of the above conjecture is true for closed 3-manifolds. Before we prove this claim note that the hypothesis of the conjecture implies that each G^i is finitely generated.

Lemma (0.3). *Let G be the fundamental group of a closed 3-manifold, such that for some i, G^i is nontrivial, finitely generated, and perfect. Then for each i, G^i/G^{i+1} is a finite group.*

Proof. Since G^i is a nontrivial perfect group, it is not a surface group. Also since G^i is finitely generated, by theorem 11.1 of [9], G^i is of finite index in G. This proves the Lemma. $\square$

Remark (0.4). After seeing the preprint [17] Peter A. Linnell pointed out to me that certain finite index subgroups of $SL(n,\mathbb{Z})$ for $n \geq 3$ satisfy the hypothesis of conjecture 0.2 of [17], but they are not adorable. These are some noncocompact lattices in $SL(n,\mathbb{R})$ which are residually finite p-groups and satisfy Kazhdan property T. I thank Professor Linnell for the stimulating example. We describe his example in the Appendix. Conjecture (0.2) remains open for the fundamental groups of closed aspherical 3-manifolds and for cocompact discrete subgroups of Lie groups. Considering this situation we state our main problem now.

Main Problem. *Find groups for which the Conjecture (0.2) is true.*

Note that G^i/G^{i+1} is finite for each i if and only if G/G^i is finite for each i. Thus, in other words, the above conjecture says that a nonadorable aspherical 3-manifold group has an infinite solvable quotient. Compare this observation with Proposition (4.1).

Also note that by Theorem (2.3), if the group G in Conjecture (0.2) is not perfect and not isomorphic to $\mathbb{Z}_2 * \mathbb{Z}_2$ then it is irreducible. Thus we can assume that the group G in the Conjecture is irreducible. Recall that a group is *irreducible* if the group is not isomorphic to a free product of two nontrivial groups.

There is another consequence of this conjecture. That is, if Conjecture (0.2) is true then the virtual Betti number conjecture will be true if a modified (half) version of it is true. We mention it below.

Modified virtual Betti number conjecture. *Let M be a closed aspherical 3-manifold such that $H_1(M,\mathbb{Z}) = 0$. Then there is a finite sheeted covering $\tilde{M}$ of M with $H_1(\tilde{M},\mathbb{Z})$ infinite.*

It is easy to see that the Conjecture (0.2) and the Modified virtual Betti number conjecture together implies the virtual Betti number conjecture.

Virtual Betti number conjecture. *Any closed aspherical 3-manifold has a finite sheeted covering with infinite first homology group.*

The virtual Betti number conjecture was raised as a question by John Hempel in question 1.2 of [10].

1. Some elementary facts about adorable groups

Recall that a group is called *perfect* if the commutator subgroup of the group is the whole group.

Lemma (1.1). *Let $f : G \to H$ be a surjective homomorphism with G adorable. Then H is also adorable and $doa(H) \leq doa(G)$.*

Example (1.2). The Artin pure braid group on more than 2 strings is not adorable, for it has a quotient a nonabelian free group. In fact the full braid group on n-strings is not adorable for $n \leq 4$ and adorable of degree 1 otherwise. (see [8]).

Lemma (1.3). *The product $G\times H$ of two groups are adorable if and only if both the groups G and H are adorable. Also if $G \times H$ is adorable then $doa(G \times H) = \max\{doa(G), doa(H)\}$.*

On the contrary, in the case of free product of groups, almost all the time the output is nonadorable. Hence, adorability is mainly a property for irreducible groups. We will consider the case of free product and more generally the generalized free product case in the next section.

Lemma (1.4). *Let G be an adorable group and H a normal subgroup of G. Assume that G^{i_0} is simple for some i_0. Then H is also adorable and $doa(H) \leq doa(G)$.*

Remark (1.5). In the above lemma, instead of assuming the strong hypothesis that G^{i_0} is simple, we can assume only that G^{i_0} is perfect and G^{i_0}/H^{i_0+1} does not have any proper normal abelian subgroup. With this weaker hypothesis the proof follows from the fact that the kernel of the surjective homomorphism $G^{i_0}/H^{i_0+1} \to G^{i_0}/H^{i_0}$ is either trivial or $G^{i_0} = H^{i_0}$. In either case it follows that H is adorable.

LEMMA (1.6). *Let H be a normal subgroup of a group G such that G^i/H^i is abelian for some i. Then G is adorable if and only if H is adorable.*

PROPOSITION (1.7). *Let H be a normal subgroup of a group G such that G/H is solvable. Then H is adorable if and only if so is G.*

Proof. Before we start the proof, we note down some generality. Suppose G has a filtration of the form $G_n < G_{n-1} < \cdots < G_1 < G_0 = G$, where G_i is normal in G_{i-1} and G_{i-1}/G_i is abelian for each i. Since G_{i-1}/G_i is abelian for each i, we have $G'_{i-1} \subset G_i$. Replacing i by $i+1$ we get $G'_i \subset G_{i+1}$. Consequently, $G^i_0 = G^i = \{G'\}^{i-1} \subset {G_1}^{i-1} \subset \{{G_1}'\}^{i-2} \subset {G_2}^{i-2} \subset \cdots \subset G'_{i-1} \subset G_i$. Thus we get $G^n \subset G_n$.

Denote G/H by F. As F is solvable we have $1 \subset F^k \subset \cdots \subset F^1 \subset F^0 = F$, where F^k is abelian. Let $\pi : G \to G/H$ be the quotient map. We have the following sequence of normal subgroups of G:

$$\cdots \subset H^n \subset H^{n-1} \cdots \subset H^1 \subset H \subset \pi^{-1}(F^k) \cdots \subset \pi^{-1}(F^0) = G\,.$$

Note that this sequence of normal subgroups satisfy the same properties as those of the filtration G_i of G above. Hence $G^{k+i} \subset H^{i-1}$. Now if G is adorable then, for some i, G^{k+i} is perfect. We have

$$H^{k+i} \subset G^{k+i} = G^{k+k+i+2} \subset H^{k+i+1}.$$

But we already have $H^{k+i+1} \subset H^{k+i}$. That is, H^{k+i} is perfect, hence H is adorable. Conversely if H is adorable then for some i, H^i is perfect. Note from the above inclusions that $H^i = G^i$ for some large i. Hence G is also adorable. $\square$

THEOREM (1.8). *A group G is adorable if and only if there is a filtration $G_n < G_{n-1} < \cdots < G_1 < G_0 = G$ of G so that G_i is normal in G_{i-1}, G_{i-1}/G_i is abelian for each i, and G_n is a perfect group.*

Proof. We use Proposition (1.7) and induction on n to prove the 'if' part of the Theorem. So assume that there is a filtration of G as in the hypothesis. Then G_n is an adorable subgroup of G with solvable quotient G/G_n. Proposition (1.7) proves this implication. The 'only if' part of the Theorem follows from the definition of adorable groups. $\square$

COROLLARY (1.9). *Let G be a torsion free infinite group and F be a finite quotient of G with kernel H such that H is free abelian and also central in G. Then G is adorable.*

Proof. Recall that equivalence classes of extensions of F by H are in one to one correspondence with $H^2(F, H)$, which is isomorphic to $Hom(F, (\mathbb{R}/\mathbb{Z})^n)$ for n the rank of H (exercise 3, p. 95 of [3]). If F is perfect then $Hom(F, (\mathbb{R}/\mathbb{Z})^n) = 0$ and hence the extension $1 \to H \to G \to F \to 1$ splits. But by hypothesis G is torsion free. Hence F is not perfect. By a similar argument it can be shown that

F^i is perfect for no i unless it is the trivial group. Since F is finite this proves that F is solvable and hence G is adorable, in fact solvable. □

We sum up the above Lemmas and Propositions in the following Theorem.

THEOREM (1.10). *Let H be a normal subgroup of an adorable group G. Then H is adorable if one of the following conditions is satisfied.*
- *G/H is solvable.*
- *for some i, G^i/H^i is abelian.*
- *for some i, G^i is simple.*
- *for some i, G^i is perfect and the group G^i/H^{i+1} does not have any proper abelian normal subgroup.*

Remark (1.11). It is known that any countable group is a subgroup of a countable simple group (see theorem 3.4, chapter IV of [14]). Also, we mentioned before that even finite index normal subgroup of an adorable group need not be adorable. So the above theorem is best possible in this regard.

In the next section we give some more examples of virtually adorable groups which are not adorable.

The following is an analogue of a theorem of Hirsch for poly-cyclic groups. The proofs of Lemmas (A) and (B) in the proof of the theorem are easy and we leave them to the reader.

THEOREM (1.12). *The following are equivalent.*
- *G is a group which admits a filtration $G = G_0 > G_1 > \cdots > G_n$ with the property that each G_{i+1} is normal in G_i with quotient G_i/G_{i+1} cyclic and G_n is a perfect group which satisfies the maximal condition for subgroups.*
- *G is adorable and satisfies the maximal condition for subgroups, i.e., for any sequence $H_1 < H_2 < \cdots$ of subgroups of G there is an i such that $H_i = H_{i+1} = \cdots$.*

Proof. The proof is along the same lines as that of Hirsch's theorem. The main lemma is the following.

LEMMA (A). *Let H_1 and H_2 be two subgroup of a group G and $H_1 \subset H_2$. Let H be a normal subgroup of G with the property that $H \cap H_1 = H \cap H_2$ and the subgroup generated by H and H_1 is equal to the subgroup generated by H and H_2. Then $H_1 = H_2$.*

Let us first prove that the first statement implies the second. By Theorem (1.8) it follows that the first statement implies G is adorable. Now we check the maximal condition by induction on n. As G_n already satisfies the maximal condition, we only need to check that G_{n-1} also satisfies the maximal condition, which follows from the following Lemma and by noting that G_{n-1}/G_n is cyclic.

LEMMA (B). *Let H be a normal subgroup of a group G such that both H and G/H satisfy the maximal condition. Then G also satisfies the maximal condition.*

Proof. Let $K_1 < K_2 < \cdots$ be an increasing sequence of subgroups of G. Consider the two sequences of subgroups $H \cap K_1 < H \cap K_2 < \cdots$ and $\{H, K_1\} < \{H, K_2\} < \cdots$. Here $\{A, B\}$ denotes the subgroup generated by the subgroups A and B. As H and G/H both satisfy the maximal condition, there are integers

k and l so that $H \cap K_k = H \cap K_{k+1} = \cdots$ and $\{H, K_l\} = \{H, K_{l+1}\} = \cdots$. Assume $k \geq l$. Then by Lemma (A) $K_k = K_{k+1} = \cdots$. □

Now we deduce the first statement from the second. As G is adorable, it has a filtration $G = G_0 > G_1 > \cdots > G_n$ with G_n perfect and each quotient abelian. Also G_n satisfies the maximal condition, as it is a subgroup of G and G satisfies the maximal condition. Since G satisfies the maximal condition, each quotient G_i/G_{i+1} is finitely generated. Now a filtration as in (1) can easily be constructed.

This proves the theorem. □

2. Generalized free products and adorable groups

We begin this section with the following result on free product of groups.

Recall that the infinite dihedral group D_∞ is isomorphic to $\mathbb{Z} \rtimes \mathbb{Z}_2 \simeq \mathbb{Z}_2 * \mathbb{Z}_2$.

Proposition (2.1). *The free product G of two nontrivial groups, one of which is not perfect, is either isomorphic to D_∞ or not adorable. Moreover, in the nonadorable case, the rank of the abelian group G^i/G^{i+1} is greater than or equal to 2 for all $i \geq 1$.*

Proof. Let G be the free product of the two nontrivial groups G_1 and G_2, where one of G_1 and G_2 is not perfect. Then, as the abelianization of $G = G_1 * G_2$ is isomorphic to $G_1/G_1^1 \oplus G_2/G_2^1$, G is also not perfect.

By Kurosh Subgroup theorem (see proposition 3.6 of [14]), any subgroup of G is isomorphic to a free product $*_i A_i * F$, where F is a free group and the groups A_i are conjugates of subgroups of either G_1 or G_2. In particular, the commutator subgroup G^1 is isomorphic to $*_i A_i * F$ for some A_i and F. Note that $[G_1, G_2] = \langle g_1 g_2 g_1^{-1} g_2^{-1} \mid g_i \in G_i, i = 1, 2\rangle$ is a subgroup of G^1. Now assume that G is not D_∞. Then $[G_1, G_2]$ is a nonabelian free group and clearly $[G_1, G_2] \cap G_1 = (1) = [G_1, G_2] \cap G_2$. Also $[G_1, G_2]$ is not conjugate to any subgroup of G_1 or G_2. Hence $[G_1, G_2]$ is a subgroup of F, which shows that F is a nontrivial nonabelian free group. Hence the abelianization of G^1 is nontrivial. By a similar argument, using Kurosh Subgroup theorem, we conclude that no G^n is perfect. This proves the first assertion of the Proposition. The second part follows from the fact that the free group F has rank ≥ 2 and a nonabelian free group has derived series consisting of nonabelian free groups. □

Remark (2.2). In Proposition (2.1) we have seen that the free product of any nontrivial group with a nonperfect group is either D_∞ or nonadorable. The natural question that arises here is what happens in the amalgamated free product case of two groups along a nontrivial group, or in the case of an HNN-extension? First, recall that there are examples of simple groups which are amalgamated free product of two nonabelian free groups along a (free) subgroup (see [2]). We give another example. Let $M = \mathbb{S}^3 - N(k)$ be a knot complement of a knot k in the 3-sphere. Assume that the Alexander polynomial of the knot is nontrivial. Then by Proposition (4.7) we know that $\pi_1(M)$ is not adorable. Recall that the first homology of M is generated by a meridian of the torus boundary of M, and the longitude, which is parallel to the knot in $\mathbb{S}^3$, represents the zero in $H_1(M, \mathbb{Z})$. Now glue two copies of M along the boundary by sending the above longitude of one copy to the meridian of the other and vice versa. Then the resulting manifold N has fundamental group isomorphic to the amalgamated free product $\pi_1(M) *_{\mathbb{Z}\times\mathbb{Z}} \pi_1(M)$, and an application of the Mayer-Vietoris sequence for integral homology shows that N has trivial first homology. That is, N has perfect fundamental group. Another example in this connection is the fundamental group of a torus knot complement in $\mathbb{S}^3$. This group is of the form $G = \mathbb{Z} *_{\mathbb{Z}} \mathbb{Z}$. If the knot is of type (p, q) then the two inclusions of $\mathbb{Z}$ in $\mathbb{Z}$ in the above amalgamated free product are defined by multiplication by p and q respectively. But G is not adorable as it has nonabelian free commutator subgroup. In the following theorem we consider a more general situation.

From now on, whenever we consider a generalized free product $G = G_1 *_H G_2$ or an HNN-extension $G = K*_H$, unless otherwise stated, we will always assume that $G_1 \neq H \neq G_2$ and $K \neq H$.

Theorem (2.3). *Let G be a group.*

If $G = G_1 *_H G_2$ is a generalized free product and $G^1 \cap H = (1)$, then one of the following holds.

- *G^1 is perfect.*
- *G^1 is isomorphic to the infinite dihedral group D_∞.*
- *G is not adorable.*

If $G = K*_H = \langle K, t \mid tHt^{-1} = \phi(H)\rangle$ is an HNN-extension and $G^1 \cap H = (1)$, then G is not adorable.

In the second case and in the last possibility of the first case for $i \geq 1$, the rank of G^i/G^{i+1} is ≥ 2.

Note that the assumption $G^1 \cap H = (1)$ implies that H is abelian.

To prove the theorem we need to recall the bipolar structure on generalized free products and the characterization of generalized free products by the existence of a bipolar structure on the group by Stallings.

Definition (2.4). (definition, p. 207 of [14]) A *bipolar structure* on a group G is a partition of G into five disjoint subsets $H, EE, EE^*, E^*E, E^*E^*$ satisfying the following axioms. (The letters X, Y, Z will stand for the letters E or E^* with the convention that $(X^*)^* = X$, etc.)

- H is a subgroup of G.
- If $h \in H$ and $g \in XY$, then $hg \in XY$.
- If $g \in XY$, then $g^{-1} \in YX$. (*Inverse axiom*)
- If $g \in XY$ and $f \in Y^*Z$, then $gf \in XZ$. (*Product axiom*)
- If $g \in G$, there is an integer $N(g)$ such that, if there exist $g_1, \ldots, g_n \in G$ and $X_0, \ldots, X_n$ with $g_i \in X^*_{i-1}X_i$ and $g = g_1 \cdots g_n$, then $n \leq N(g)$. (*Boundedness axiom*)
- $EE^* \neq \emptyset$. (*Nontriviality axiom*)

It can be shown that every amalgamated free product or HNN-extension has a bipolar structure (p. 207-208 of [14]). The following theorem of Stallings shows that the converse is also true.

Theorem (2.5). (Theorem 6.5 of [14]) *A group G has a bipolar structure if and only if G is either a nontrivial free product with amalgamation (possibly an ordinary free product) or an HNN-extension.*

Proof of Theorem (2.3). First note that the first 5 properties in the above definition are hereditary, that is, any subgroup F of G has a partition by subsets satisfying these properties. The induced partition of F is obtained by taking the intersections of $H, EE, \ldots$ with F. But $EE^* \cap F$ could be empty. We replace F by the commutator subgroup G^1 of G. We would like to check the sixth property (that is, the nontriviality axiom) for this induced partition on G^1.

We consider the amalgamated free product case first. Recall that if we write $g \in G - H$ in the form $g = c_1 \cdots c_n$, where no $c_i \in H$ and each c_i is in one of the factors G_1 or G_2 and successive c_i, c_{i+1} come from different factors, then $g \in EE^*$ if and only if $c_1 \in G_1$ and $c_n \in G_2$. Such a word is called *cyclically reduced*. Thus EE^* consists of all cyclically reduced words. Let $g_1 \in G_1 - H$ and $g_2 \in G_2 - H$; then $g_1 g_2 g_1^{-1} g_2^{-1}$ is a cyclically reduced word and is contained in $EE^* \cap G^1$. Hence the induced partition on G^1 defines a bipolar structure on G^1 with amalgamating subgroup $G^1 \cap H = (1)$. Hence G^1 is a free product of two nontrivial groups. Using Proposition (2.1) we complete the proof in this case.

When G is an HNN-extension we have a similar situation. We have to check that $EE^* \cap G^1 \neq \emptyset$. Recall from (p. 208 of [14]) that if we write $g \in G - H$ in the reduced form $g = h_0 t^{\epsilon_1} h_1 \cdots t^{\epsilon_n} h_n$ (where $\epsilon_i = \pm 1$ and $h_i \in K$ for each i) then $g \in EE^*$ if and only if $h_0 \in K - H$, or $h_0 \in H$ and $\epsilon = +1$, and $h_n \in H$ and $\epsilon_n = +1$. Now let $h_0 \in K - H$ and $h_1 \in H$, then $h_0(h_1 t^{-1}) h_0^{-1} (h_1 t^{-1})^{-1} = (h_0 h_1) t^{-1} h_0^{-1} t h_1^{-1} \in EE^* \cap G^1$. Hence the induced partition on G^1 gives a bipolar structure on G^1. Since $G^1 \cap H = (1)$ we get that G^1 is a free product of a nontrivial group with the infinite cyclic group. Hence Proposition (2.1) applies again. $\square$

We introduce below a stronger version of the notion of solvability which depends both on the group and the group where it is embedded.

Definition (2.6). A subgroup H of a group G is called *G-solvable* (or *subgroup solvable*) if $G^n \cap H = (1)$ for some n. If in addition $G^{n-1} \cap H \neq (1)$ then H is called *n-step G-solvable* (or *n-step subgroup solvable*).

Note that if H is G-solvable then H is solvable. Also if G is solvable then any subgroup of G is G-solvable.

Now we can state a Corollary of Theorem (2.3). The proof is easily deduced from the proof of Theorem (2.3) and is left to the reader.

COROLLARY (2.7). *Let G be a group.*

If $G = G_1 *_H G_2$ is a generalized free product and H is n-step G-solvable, then one of the following holds.

- *G is adorable of degree n and not solvable.*
- *$G^n \simeq D_\infty$.*
- *G is not adorable.*

If $G = K*_H = \langle K, t \quad tHt^{-1} = \phi(H)\rangle$ is an HNN-extension and H is G-solvable, then G is not adorable.

In the second case and in the last possibility of the first case for $i \geq 1$, the rank of G^i/G^{i+1} is ≥ 2.

The following Lemma considers some more generalized free product cases.

LEMMA (2.8). *Let $G_1 *_H G_2$ be a generalized free product with H abelian and contained in the center of both G_1 and G_2. Also assume that one of G_1/H or G_2/H is not perfect. Then $G_1 *_H G_2$ is either solvable or not adorable.*

Proof. Using normal form of elements of $G_1 *_H G_2$ it is easy to show that the center of $G_1 *_H G_2$ is H. This implies that we have a surjective homomorphism $G_1 *_H G_2 \to (G_1 *_H G_2)/H = G_1/H * G_2/H$. By Proposition (2.1), $G_1/H * G_2/H$ is either the infinite dihedral group or not adorable and hence $G_1 *_H G_2$ is either solvable or not adorable by Lemma (1.1). □

Example (2.9). Using Lemma (2.8) we now give a large class of examples of compact Haken 3-manifolds with nonadorable fundamental groups. Let M and N be two compact orientable Seifert fibered 3-manifolds with nonempty boundary and orientable base orbifold. Such examples of M and N are torus knot complements in $\mathbb{S}^3$. Let ∂M and ∂N be the boundary components of M and N respectively. Note that both ∂M and ∂N are tori. Let $\gamma_1 \subset \partial M$ and $\gamma_2 \subset \partial N$ be simple closed curves which are parallel to some regular fiber of M and N, respectively. Recall that both γ_1 and γ_2 represent central elements of $\pi_1(M)$ and $\pi_1(N)$ respectively. Now choose an annulus neighborhood A_1 of γ_1 in ∂M and A_2 of γ_2 in ∂N and glue M and N identifying A_1 with A_2 by a diffeomorphism which sends γ_1 to γ_2. Let P be the resulting manifold. Then P is a compact Haken 3-manifold with tori boundary and, by the Seifert-van Kampen theorem, $\pi_1(P)$ satisfies the hypothesis of Lemma (2.8) and hence is either solvable or not adorable. Here note that the manifold P itself is Seifert fibered. In the next section we will show that in fact an infinite group which is the fundamental group of a compact Seifert fibered 3-manifold is nonadorable except for some few cases.

3. Adorability and 3-manifolds

Seifert fibered spaces are a fundamental and very important class of 3-manifolds. Conjecturally (due to Thurston) any 3-manifold is build from Seifert fibered spaces and hyperbolic 3-manifolds. Results of Jaco-Shalen, Johannson and Thurston say that this is in fact true for any Haken 3-manifold.

THEOREM (3.1). *Let M^3 be a compact Seifert fibered 3-manifold. Then one of the following four cases occur.*

- *$(\pi_1(M))^i$ is finite for some $i \leq 2$.*
- *$\pi_1(M)$ is solvable.*
- *$\pi_1(M)$ is not adorable and $(\pi_1(M))^i/(\pi_1(M))^{i+1}$ has rank greater than 1 for all i greater than some i_0.*
- *$\pi_1(M)$ is perfect.*

Proof. First we recall some well known group theoretic information about the fundamental group of Seifert fibered spaces. If B is the base orbifold of M then there is a surjective homomorphism $\pi_1(M) \to \pi_1^{orb}(B)$, where $\pi_1^{orb}(B)$ is the orbifold fundamental group of B. Recall that $\pi_1^{orb}(B)$ is a Fuchsian group. Also recall that the above surjective homomorphism is part of the following exact sequence:

$$1 \to C \to \pi_1(M) \to \pi_1^{orb}(B) \to 1.$$

Here C is the cyclic normal subgroup of $\pi_1(M)$ generated by a regular fiber of the Seifert fibration of M. Also if $\pi_1(M)$ is infinite then C is an infinite cyclic subgroup of $\pi_1(M)$.

Some examples of Seifert fibered 3-manifolds with finite fundamental group are lens spaces and the Poincare sphere. So, from now on we assume $\pi_1(M)$ is infinite. Then we get the exact sequence:

$$1 \to \mathbb{Z} \to \pi_1(M) \to \pi_1^{orb}(B) \to 1.$$

There are now two cases to consider.

Case 1. $\pi_1^{orb}(B)$ is finite. By lemma 2.5 of [7] $\pi_1(M)$ has a finite normal subgroup G with quotient isomorphic either to $\mathbb{Z}$ or to D_∞. Since D_∞ is solvable $(\pi_1(M))^i$ is finite for some $i \leq 2$.

Case 2. $\pi_1^{orb}(B)$ is infinite and not a perfect group. Then by Theorem 1.5 of [18] there is a torsion free normal subgroup H of $\pi_1^{orb}(B)$ so that $\pi_1^{orb}(B)/H$ is a finite solvable group. Hence by Proposition (1.7) $\pi_1^{orb}(B)$ is adorable if and only if so is H. Since H is of finite index in $\pi_1^{orb}(B)$ by a result of Hoare, Karrass and Solitar (see proposition 7.4, Chapter III of [14]), H is again a Fuchsian group. But a torsion free Fuchsian group is the fundamental group of a compact surface (evident from the presentation of such groups). Hence H is either $\mathbb{Z}$, $\mathbb{Z} \times \mathbb{Z}$, $\mathbb{Z} \rtimes \mathbb{Z}$, or nonadorable. Thus by Proposition (1.7) $\pi_1^{orb}(B)$ is either solvable or nonadorable. If $\pi_1^{orb}(B)$ is solvable then from the above exact sequence it follows that $\pi_1(M)$ is also solvable. On the other hand Lemma (1.1) shows $\pi_1(M)$ is nonadorable whenever $\pi_1^{orb}(B)$ is.

Next, consider the case when $\pi_1^{orb}(B)$ is a perfect group. Let $x_1, x_2, \ldots, x_n$ be the cone points on B with indices $p_1, p_2, \ldots, p_n$ greater than or equal to 2. By (theorem 1.5 of [18]) $\pi_1^{orb}(B)$ is perfect if and only if $B = \mathbb{S}^2$ and the indices $p_1, p_2, \ldots, p_n$ are pairwise coprime. It is well known that in this situation M is an integral homology 3-sphere and hence $\pi_1(M)$ is also perfect. This proves the theorem. $\square$

Notice that the proof of the above theorem is not very illuminating in the sense that it does not show the cases when the groups are nonadorable or solvable. Below we show that in fact in most cases the fundamental group of a compact Seifert fibered space is nonadorable. For simplicity of presentation we consider Seifert fibered spaces whose base orbifold B is orientable and has only cone singularities. Note that the proof of the Theorem deals with both orientable and nonorientable bases B and for any kind of singularities. First let us consider the case when M has nonempty boundary. Since B also has nonempty boundary, $\pi_1^{orb}(B)$ is a free product of cyclic groups [9] and hence, by Proposition (2.1), $\pi_1^{orb}(B)$ is either the infinite dihedral group or is nonadorable if it is a nontrivial free product. Hence either $\pi_1(M)$ is solvable (when $\pi_1^{orb}(B)$ is dihedral or cyclic) or (by Lemma (1.1)) $\pi_1(M)$ is not adorable.

If M is closed then we have the same situation as above except that $\pi_1^{orb}(B)$ has the following form:

$$\pi_1^{orb}(B) = \langle a_1, \ldots, a_g, b_1, \ldots, b_g, x_1, \ldots, x_n \mid x_1^{j_1} = \cdots = x_n^{j_n} = 1;$$

$$\Pi_{j=1}^{g}[a_j, b_j]x_1 \cdots x_n = 1\rangle,$$

where $x_1,\dots x_n$ represent loops around the cone points of B. We will consider the case $g=0$ at the end of the proof. If $g\geq 1$ then adding the extra relations $a_1=1$ we get that $\pi_1^{orb}(B)$ has the following homomorphic image

$$\langle a_2,\dots,a_g,b_1,\dots,b_g,x_1,\dots,x_n \mid x_1^{j_1}=\cdots=x_n^{j_n}=1;$$
$$\Pi_{j=2}^{g}[a_j,b_j]x_1\cdots x_n=1\rangle\,.$$

If there is no cone point on B and $g=1$ then M is an $\mathbb{S}^1$-bundle over the torus, and hence has solvable fundamental group. Otherwise, the last group is a free product of the infinite cyclic group (generated by b_1) and another group, and hence not adorable by Proposition (2.1). Thus $\pi_1^{orb}(B)$ is also not adorable by Lemma (1.1). Consequently so is $\pi_1(M)$.

Now we consider the case when $g=0$. There are two further cases to consider.

Case A. $\pi_1^{orb}(B)$ is finite. This case occurs when B has at most 3 cone points, and if exactly 3 cone points with indices n_1,n_2,n_3 then $\frac{1}{n_1}+\frac{1}{n_2}+\frac{1}{n_3}>1$ (see theorem 12.2, of [9]). We have already discussed this case in **Case 1** in the proof of the theorem.

Case B. $\pi_1^{orb}(B)$ is infinite. In this case there are the following two possibilities (see theorem 12.2 of [9]). (a) B has more than 3 cone points. (b) B has 3 cone points with indices j_1,j_2,j_3 so that $\frac{1}{j_1}+\frac{1}{j_2}+\frac{1}{j_3}\leq 1$.

For (a) we need the following easily verified remark.

Remark (3.2). *If B is a sphere with* 3 *cone points then* $|\pi_1^{orb}(B)|\geq 3$.

Now recall that in (a) $\pi_1^{orb}(B)$ has the following presentation.

$$\langle x_1,\dots,x_n \mid x_1^{j_1}=\cdots=x_n^{j_n}=1; x_1\cdots x_n=1\rangle$$

where $n\geq 4$. Now assume $n\geq 6$ and add the relation $x_1x_2x_3=1$ in the above presentation. Then $\pi_1^{orb}(B)$ surjects onto the free product of

$$\langle x_1,x_2,x_3 \mid x_1^{j_1}=x_2^{j_2}=x_3^{j_3}=1; x_1x_2x_3=1\rangle$$

and

$$\langle x_4,\dots,x_n \mid x_4^{j_4}=\cdots=x_n^{j_n}=1; x_4\cdots x_n=1\rangle.$$

By Proposition (2.1) and Remark (3.2) it follows that $\pi_1^{orb}(B)$ is either perfect or not adorable, and hence so is $\pi_1(M)$. In the case $n=5$, if there is a pair of indices j_k and j_l so that $(j_k,j_l)\geq 3$ then it is easy to show that $\pi_1(M)$ is nonadorable. We leave the remaining cases to the reader.

In (b), when $\frac{1}{j_1}+\frac{1}{j_2}+\frac{1}{j_3}=1$, $\pi_1^{orb}(B)$ is a discrete group of isometries of the Euclidean plane. Recall that a torsion free discrete group of isometries of the Euclidean plane is isomorphic to $\mathbb{Z}$ or $\mathbb{Z}\times\mathbb{Z}$ or $\mathbb{Z}\rtimes\mathbb{Z}$ and hence, by the result of Sah we mentioned above, $\pi_1^{orb}(B)$ is either perfect or solvable. On the other hand, if $\frac{1}{j_1}+\frac{1}{j_2}+\frac{1}{j_3}<1$ then $\pi_1^{orb}(B)$ is a discrete groups of isometries of the hyperbolic plane. Since a group of isometries of the hyperbolic plane does not contain a free abelian group on more than one generator, it follows by the result of Sah that in this case $\pi_1^{orb}(B)$ is either perfect, a finite solvable extension of $\mathbb{Z}$, or nonadorable. Hence $\pi_1(M)$ is either solvable, perfect, or nonadorable.

Remark (3.3). Recall that a Fuchsian group is a discrete subgroup of $PSL(2,\mathbb{R})$, and that it is either a free product of cyclic groups or is isomorphic to a group of the form $\pi_1^{orb}(B)$. In the free product case, except for the infinite dihedral group, all other free products are nonadorable. In the remaining cases we have already seen in the proof of Theorem (3.1) that a Fuchsian group is either finite, perfect, solvable, or nonadorable, and in most cases it is nonadorable. It is not known to me if a similar situation occurs for discrete subgroups of $PSL(2, \mathbb{C})$. Such information would be very useful to get some hold on the virtual Betti number conjecture for hyperbolic 3-manifolds.

4. (Non)adorability under homological or geometric hypothesis

In Section 2, under some group theoretic hypothesis, we showed when a generalized free product or an HNN-extension produces a nonadorable group.

This section deals with some homological or geometric (or topological) hypothesis on a group which ensures that the group is nonadorable.

PROPOSITION (4.1). *Let M^3 be a compact 3-manifold with the property that there is an exact sequence of groups $1 \to H \to \pi_1(M) \to F \to 1$ such that H is finitely generated nonabelian but not the fundamental group of the Klein bottle and F is an infinite solvable group. Then $\pi_1(M)$ is not adorable.*

Proof. By theorem 11.1 of [9] it follows that H is the fundamental group of a compact surface. Also as H is nonabelian and not the Klein bottle group, it is not adorable. The Proposition now follows from Proposition (1.7). □

PROPOSITION (4.2). *Let G be a torsion free group and H a free nonabelian (or abelian) normal subgroup of G with quotient F a nontrivial finite (or finite perfect) group. Then G is not adorable.*

Proof. If H is nonabelian then, by Stallings' theorem, G itself is free and hence not adorable. So assume H is free abelian. Since in this case F is a perfect group, the restriction of the quotient map $G \to F$ to G^i is again surjective for each i with $H \cap G^i$ as kernel. And since G is infinite and torsion free, $H \cap G^i$ is nontrivial free abelian for all i. This shows that each G^i is again a Bieberbach group. Note that if $H^1(G^i, \mathbb{Z}) = 0$ then G^i is centerless, and it is known that centerless Bieberbach groups are meta-abelian with nontrivial abelian holonomy group and hence solvable [11]. But since each G^i surjects onto a nontrivial perfect group it cannot be solvable. Hence $H^1(G^i, \mathbb{Z}) \neq 0$ for each i. This proves the Proposition. □

The conclusion of the above Proposition remains valid if we assume that F is nonsolvable adorable.

By Bieberbach's theorem [4], we have the following Corollary.

COROLLARY (4.3). *The fundamental group of a closed flat Riemannian manifold is nonadorable unless it is solvable.*

So far we have given examples of nonadorable groups which are fundamental groups of known class of manifolds or of manifolds with some strong Riemannian structure. The following Theorem gives a general class of examples of nonadorable groups under some homological conditions.

THEOREM (4.4). *Let G be a group satisfying the following properties*
- $H_1(G,\mathbb{Z})$ *has rank* ≥ 3.
- $H_2(G^j,\mathbb{Z}) = 0$ *for* $j \geq 0$.

Then G is not adorable. Moreover, G^j/G^{j+1} has rank ≥ 3 for each $j \geq 1$.

Proof. Consider the short exact sequence.

$$1 \to G^1 \to G \to G/G^1 \to 1\,.$$

We use the Hochschild-Serre spectral sequence (p. 171 of [3]) of the above exact sequence. The E^2-term of the spectral sequence is $E^2_{pq} = H_p(G/G^1, H_q(G^1,\mathbb{Z}))$. Here $\mathbb{Z}$ is considered as a trivial G-module. This spectral sequence gives rise to the following five term exact sequence:

$$H_2(G,\mathbb{Z}) \to E^2_{20} \to E^2_{01} \to H_1(G,\mathbb{Z}) \to E^2_{10} \to 0.$$

Using (2) we get

$$0 \to H_2(G/G^1, H_0(G^1,\mathbb{Z})) \to H_0(G/G^1, H_1(G^1,\mathbb{Z})) \to H_1(G,\mathbb{Z})$$
$$\to H_1(G/G^1, H_0(G^1,\mathbb{Z})) \to 0.$$

As $\mathbb{Z}$ is a trivial G-module we get

$$0 \to H_2(G/G^1,\mathbb{Z}) \to H_0(G/G^1, H_1(G^1,\mathbb{Z})) \to H_1(G,\mathbb{Z}) \to$$
$$H_1(G/G^1,\mathbb{Z}) \to 0.$$

Note that the homomorphism between the last two nonzero terms in the above exact sequence is an isomorphism. Also, the second nonzero term on the left is isomorphic to the co-invariant $H_1(G^1,\mathbb{Z})_{G/G^1}$, and hence we have the following:

$$H_2(G/G^1,\mathbb{Z}) \simeq H_1(G^1,\mathbb{Z})_{G/G^1}.$$

Since G/G^1 has rank ≥ 3 we get that $H_2(G/G^1,\mathbb{Z})$ has rank greater or equal to 3. This follows from the following lemma.

LEMMA (4.5). *Let A be an abelian group. Then the rank of $H_2(A,\mathbb{Z})$ is $rkA(rkA-1)/2$ if rkA is finite, and infinite otherwise.*

Proof. If A is finitely generated then, from the formula $H_2(A,\mathbb{Z}) \simeq \bigwedge^2 A$, it follows that the rank of $H_2(A,\mathbb{Z})$ is $rkA(rkA-1)/2$. In the case where A is countable and infinitely generated, there are finitely generated subgroups A_n of A such that A is the direct limit of A_n. Now, as homology of group commutes with direct limits, the proof follows using the previous case. Similar arguments apply when A is uncountable. □

To complete the proof of the theorem, note that there is a surjective homomorphism $H_1(G^1,\mathbb{Z}) \to H_1(G^1,\mathbb{Z})_{G/G^1}$. Thus we have proved that $H_1(G^1,\mathbb{Z})$ also has rank ≥ 3. Finally, replacing G by G^n and G^1 by G^{n+1} and using induction on n, the proof is completed. □

There are two important consequences of Theorem (4.4). First we recall some definitions from [19].

Let R be a nontrivial commutative ring with unity. The class $E(R)$ consists of groups G for which the trivial G-module R has a RG-projective resolution

$$\cdots \to P_2 \to P_1 \to P_0 \to R \to 0$$

such that the map $\mathbf{1}_R \otimes \partial_2 : R \otimes_{RG} P_2 \to R \otimes_{RG} P_1$ is injective. Note that if a group belongs to $E(R)$ then $H_2(G, R) = 0$. Also this condition is sufficient to belong to $E(R)$ for groups of cohomological dimension less or equal to 2. By definition G lies in E if it belongs to $E(R)$ for all R. A characterization of E-groups is that a group G is an E-group if and only if G belongs to $E(\mathbb{Z})$ and G/G^1 is torsion free (see lemma 2.3 of [19]).

Corollary (4.6). *Let G be an E-group and rank of $H_1(G, \mathbb{Z})$ is ≥ 2. Then G is not adorable.*

Proof. By Theorem A of [19] it follows that G satisfies the second condition of Theorem (4.4). Hence we get that $H_1(G^2, \mathbb{Z})$ has rank ≥ 1 and hence in particular G^2 is not perfect. On the other hand, an E-group has derived length $0, 1, 2,$ or infinity (remark after (theorem A of [19]). Thus G is not adorable. □

In the following Proposition we give an application of Theorem (4.4) for knot groups.

Proposition (4.7). *Let $H = \pi_1(\mathbb{S}^3 - k)$, where k is a nontrivial knot in the 3-sphere with nontrivial Alexander polynomial. Then H is not adorable. Moreover, if the rank of H^1/H^2 is greater than or equal to 3 then the same is true for H^j/H^{j+1} for all $j \geq 2$.*

In fact, a stronger version of the Proposition follows, since by [19] the successive quotients of the derived series of G are torsion free. Thus we get that the successive quotients of the derived series are nontrivial and torsion free.

Proof of Proposition (4.7). First recall that the second condition of Theorem (4.4) follows from (theorem A of [19]). On the other hand, the commutator subgroup of a knot group is perfect if and only if the knot has trivial Alexander polynomial. So assume that H^1 is not perfect. If H^1 is finitely generated then in fact it is nonabelian free and hence H is not adorable. If rank of H^1/H^2 is ≥ 3 then the proof follows from the above Theorem. So assume that rank of H^1/H^2 is ≤ 2.

Recall that the rank of the abelian group H^1/H^2 is equal to the degree of the Alexander polynomial of the knot (see theorem 1.1 of [6]). Thus if the rank of H^1/H^2 is 1 then the Alexander polynomial has degree 1, which is impossible as the Alexander polynomial of a knot always has even degree. Next if rank of H^1/H^2 is 2 then H is not adorable by Corollary (4.6) and noting that knot groups are E-groups. □

Definition (4.8). A Lie group is called *adorable* if it is adorable as an abstract group.

Theorem (4.9). *Every connected (real or complex) Lie group is adorable.*

Proof. Let G be a Lie group and consider its derived series:

$$\cdots \subset G^n \subset G^{n-1} \cdots \subset G^1 \subset G^0 = G.$$

Note that each G^i is a normal subgroup of G. Define $G_i = \overline{G^i}$. Then we have a sequence of normal subgroups

$$\cdots \subset G_n \subset G_{n-1} \cdots \subset G_1 \subset G_0 = G$$

so that G_i is a closed Lie subgroup of G and G_i/G_{i+1} is abelian for each i. Suppose that, for some i, dim $G_i = 0$, i.e., G_i is a closed discrete normal subgroup of G. We claim G_i is abelian. For, fix $g_i \in G_i$ and consider the continuous map $G \to G_i$ given by $g \mapsto g g_i g^{-1}$. As G is connected and G_i is discrete, the image of this map is the singleton $\{g_i\}$. That is, g_i commutes with all $g \in G$ and hence G_i is abelian.

As $G^i \subset G_i$, G^i is also abelian. Thus G is solvable and hence adorable.

Next assume no G_i is discrete. Then, as G is finite dimensional and the G_i's are Lie subgroups of G, there is an i_0 so that $G_j = G_{j+1}$ for all $j \geq i_0$ and dim $G_{i_0} \geq 1$. We need the following Lemma to complete the proof of the Theorem.

LEMMA (4.10). *Let G be a (real or complex) Lie group such that $\overline{G^1} = G$. Then $G^2 = G^1$, that is, G^1 is a perfect group.*

Proof. The proof of the lemma follows from theorems XII.3.1 and XVI.2.1 of [12]. □

We have $G^{i_0} \subset G_{i_0}$ and hence

$$G_{i_0} = G_{i_0+1} = \overline{G^{i_0+1}} \subset \overline{G^1_{i_0}} \subset \overline{G_{i_0}} = G_{i_0}.$$

This implies $\overline{G^1_{i_0}} = G_{i_0}$. Now from the above Lemma we get G_{i_0} is adorable. Thus G_{i_0} is a normal adorable subgroup of G_{i_0-1} with quotient G_{i_0-1}/G_{i_0} abelian and hence, by Proposition (1.7), G_{i_0-1} is also adorable. By induction, it follows that G is adorable. □

5. Appendix

In this section we describe the counterexample given by Peter A. Linnell to conjecture 0.2 of [17].

Example (5.1). (P.A. Linnell) Let $n \geq 3$ and p be an odd prime. Let K be the kernel of the homomorphism $SL(n, \mathbb{Z}) \to SL(n, \mathbb{Z}/p\mathbb{Z})$ which is induced by the homomorphism $\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$. When $p = 2$, let K be the kernel of $SL(n, \mathbb{Z}) \to SL(n, \mathbb{Z}/4\mathbb{Z})$. Now we have the following three facts about K.

• K is a residually finite p-group. Hence we get that K^{i+1} is a proper subgroup of K^i for each i.

• K satisfies Kazhdan property T. Hence K^i/K^{i+1} is a finite group for each i.

• K is finitely presented and torsion free.

Thus K is not adorable. But by the second and the third fact above, K satisfies the hypothesis of conjecture 0.2 of [17].

A notable fact is that K is a noncocompact discrete subgroup of $SL(n, \mathbb{R})$. It would be very interesting to prove Conjecture (0.2) for cocompact discrete subgroup of Lie groups.

6. Problems

In this section we state some problems for a further study on adorable groups. We also give the motivations behind each problem and mention known results related to the problem.

Problem (6.1). Study the Main Problem for some particular class of groups, for example for cocompact discrete subgroups of Lie groups or for groups which are fundamental groups of closed nonpositively curved Riemannian manifolds.

Problem (6.1) is related to the particular case of the virtual Betti number conjecture for hyperbolic 3-manifolds. We have already seen that a discrete subgroup of $PSL(2,\mathbb{R})$ is either finite, solvable, perfect, or nonadorable. In fact, it is possible to describe when each of these possibilities occurs. A similar result about discrete subgroups of $PSL(2,\mathbb{C})$ would be very important. A more precise problem is the following.

Problem (6.2). Given a positive integer n, does there exist a discrete (torsion free) subgroup of $PSL(2,\mathbb{C})$ which is adorable of degree n?

Problem (6.3). Find all 3-manifolds with adorable fundamental group.

Some examples of such 3-manifolds are integral homology 3-spheres and knot complement of knots with trivial Alexander polynomial. In Theorem (3.1) we have seen that most Seifert fibered spaces have nonadorable fundamental group, and also we have shown when the fundamental group is adorable.

Problem (6.4). Prove that most groups are not adorable.

A possible approach to study Problem (6.4) is by the same method which was used to show that most groups are hyperbolic.

A small and first step towards Conjecture (0.2) is the following.

Problem (6.5). Show that Conjecture (0.2) is true for the fundamental groups of compact Haken 3-manifolds.

We have already mentioned that it is true for Seifert fibered spaces. Note that if the fundamental group of a compact Haken 3-manifold satisfies the hypothesis of Conjecture (0.2) then the manifold has to be closed.

Received June 19, 2003

Final version received March 03, 2004

School of Mathematics
Tata Institute
Homi Bhabha Road
Mumbai 400 005
India
roushon@math.tifr.res.in
http://www.math.tifr.res.in/~ roushon

References

[1] M.R. Bridson and A. Haefliger, Metric spaces of nonpositive curvature, Springer-Verlag, Berlin, Heidelberg, 1999.

[2] M. Burger and S. Mozes, *Finitely presented simple groups and product of trees*, C. R. Acad. Sci. Paris Sér. I Math. **324** (1997), 747–752.

[3] Kenneth S. Brown, Cohomology of groups, Springer-Verlag, New York Heidelberg Berlin, 1982.

[4] Leonard S. Charlap, Bieberbach groups and flat manifolds, Springer-Verlag, New York Berlin Heidelberg London Paris Tokyo, 1986.

[5] Tim Cochran, *Noncommutative knot theory*, Algebr. Geom. Top. **4** (2004), 347–398.
[6] R.H. Crowell, *The group G'/G'' of a knot group G*, Duke Math. Journal, **30** (1963), 349–354.
[7] F.T. Farrell and L.E. Jones, *The lower algebraic K-theory of virtually infinite cyclic groups*, K-theory, **9** (1995), 13–30.
[8] E.A. Gorin and V.Ja. Lin, *Algebraic equations with continuous coefficients, and certain questions of the algebraic theory of braids*, Mat. Sb., **78** (1969), 579–610.
[9] John Hempel, 3-manifolds, Ann. of Math. Stud. **86**, Princeton Univ. Press, Princeton NJ, 1976.
[10] John Hempel, *Homology of coverings*, Pacific J. of Math., **112** (1984), 83–113.
[11] H. Hiller and C. Sah, *Holonomy of flat manifolds with $b_1 = 0$*, Quart. J. Math. Oxford, **37** (1986), 177–187.
[12] G. Hochschild, The structure of Lie groups, Holden-Day, Inc., 728 Montgomery Street, San Francisco, California, 1965.
[13] B. Leeb, *3-manifolds with(out) metrices of nonpositive curvature*, Invent. Math., **122** (1995), 277–289.
[14] Roger C. Lyndon and Paul E. Schupp, Combinatorial group theory, Springer-Verlag, Berlin Heidelberg New York, 1977.
[15] Sayed K. Roushon, *L-theory of 3-manifolds with non-vanishing first Betti number*, Internat. Math. Res. Notices, **3** (2000), 107–113.
[16] Sayed K. Roushon, *Vanishing structure set of Haken 3-manifolds*, Math. Ann., **318** (2000), 609–620.
[17] Sayed K. Roushon, *Topology of 3-manifolds and a class of groups*, preprint, arXiv:math.GT/0209121.
[18] C.-H. Sah, *Groups related to compact Riemann surfaces*, Acta. Math., **123** (1969), 13–42.
[19] Ralph Strebel, *Homological methods applied to the derived series of groups*, Comment. Math. Helv., **49** (1974), 302–332.
[20] S.T. Yau, *On the fundamental group of compact manifolds of nonpositive curvature*, Ann. of Math., (2) **93** (1971), 579–585.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

LUSTERNIK-SCHNIRELMANN CATEGORY, HOPF DUALITY, AND ISOLATED INVARIANT SETS

JOSÉ M.R. SANJURJO

This paper is dedicated to Professor Francisco Javier González Acuña on the occasion of his sixtieth birthday

ABSTRACT. We establish an inequality relating the Lusternik–Schnirelmann coefficient of the unstable manifold of an isolated invariant set of a flow and the coefficients of a Morse decomposition of this set. We also establish some homological relations between the Morse sets of a particular Morse decomposition of an isolated invariant set. This result introduces Hopf duality properties in the context of continuous dynamical systems.

1. Introduction

This paper can be partially considered as a continuation of the article [15], that was devoted to the study of properties of the Lusternik-Schnirelmann category in the context of dynamical systems. As we remarked in the introduction of that paper, there are several different definitions of this coefficient, although most of them agree in the important case of ANR's (Absolute Neighborhod Retracts). The interested reader can consult the review articles [5] by R.H.Fox and [9], [10] by I.M. James for general information about this topological invariant.

The definition of Lusternik-Schnirelmann category of a compactum used in this paper, as well as in [15], is that introduced by K. Borsuk in [1]. It was also remarked in [15] that Borsuk gave this definition in the context of his shape theory, which is a branch of Geometric Topology that has been used in the study of the global properties of dynamical systems. The Lusternik-Schnirelman category of a (metric) compactum X in the sense of Borsuk is defined in the following way: Suppose that X is embedded in an ANR, M (this is no loss of generality since all compacta can be embedded in ANR's, for instance in the Hilbert cube or in Euclidean spaces in the finite-dimensional case). We denote by $\eta(X)$ the number defined as follows (see [15]):

If $X = \emptyset$ then $\eta(X) = 0$

If $X \neq \emptyset$ and if there exist natural numbers n such that:

(1) for every neighborhood U of X in M there exist compacta $X_1, \dots, X_n$ contractible in U and such that $X = X_1 \cup \dots \cup X_n$,

then $\eta(X)$ denotes the smallest of all such numbers n.

If $X \neq \emptyset$ and if no natural number n satisfies (1) then $\eta(X) = \infty$.

2000 *Mathematics Subject Classification*: 37B30, 55M30, 55P55 .

Keywords and phrases: Lusternik-Schnirelmann category, Hopf duality, Morse decompositions, shape .

Partially supported by DGES .

The definition of $\eta(X)$ does not depend on M or on the particular embedding that we consider. The coefficient $\eta(X)$ is the Lusternik-Schnirelmann category of X in the sense of Borsuk and it is a homotopy invariant or, more generally, a shape invariant. Borsuk's definition is different from the one considered, for instance, in [5], but both of them agree in the case of compact ANR's. The coefficient $\eta(X)$ is originally defined only for compacta but, by the homotopy invariance mentioned below, it makes sense to consider it also for spaces with the homotopy type of compacta.

The main result proved in [15] was an inequality relating the Lusternik-Schnirelmann category of the unstable manifold of an isolated invariant compactum of a flow and the categories of the Morse sets of a Morse decomposition of that compactum. There was, however, an important restriction for the validity of the result: the isolated invariant set should be regular. In the present paper we give an example showing that this requirement is essential and that the inequality does not hold for more general isolated invariant sets. On the other hand, we prove that if we consider the unstable manifold endowed with its *intrinsic topology* (as defined by Robbin and Salamon [14]) then the restriction can be removed and the inequality is true in its full generality.

Another aim of this note is to provide a dynamical framework to express duality properties that are studied in Topology under the name of *Hopf duality.* This kind of duality refers to $(n-1)$-manifolds, W, embedded in the n-sphere S^n, and establishes homological relations between the two n-manifolds with boundary into which S^n is decomposed by W (see Steenrod and Epstein [17]). We consider here the situation of flows defined in a locally compact metric space X possessing an attractor, M, which is an n-manifold satisfying some specific conditions. The attractor is endowed with a Morse decomposition $\{M_0, M_1, M_2\}$ where M_0 is an $(n-1)$-submanifold of M decomposing M into two manifolds with common boundary M_0. We present in the paper some duality properties of the homology and cohomology Conley indices of the Morse sets. We also study the more general situation in which M is required only to be an isolated invariant set (not necessarily an attractor) and we get some homological properties of the Morse sets. The most general result in this direction, stated in Corollary 3, is presented in terms of the unstable manifold with its intrinsic topology and formulated in the language of Čech homology. We remark that the use of the intrinsic topology is also essential in this result. We use in our proofs some of the basic theory of Dynamical Systems, stability, attractors, isolated invariant sets, and Conley index as presented by Bhatia and Szego [2], Conley [3], or Rybakowski [13], and also some Algebraic Topology of manifolds, in particular the classical Poincaré and Lefschetz duality theorems. The books by Dold [4] and Hatcher [8] are good references for these subjects. There are a number of recent papers in the literature where other topological properties of attractors and Morse decompositions are studied (see, for instance, [6], [7], [11]).

We say that a Morse decomposition $\{M_0, M_1, M_2\}$ is connected if all the Morse sets are connected. The rest of the terminology is the standard one in the literature.

The results in this paper were obtained while the author was visiting the University of Manchester (U.K.). The author is grateful to the members of the Department of Mathematics of that University, and in particular to Nigel Ray, for their hospitality, and to his colleagues at the Complutense and Polytechnical Universities of Madrid, A. Giraldo, J.M. Montesinos-Amilibia, M.A. Morón and F. Ruiz del Portal for useful conversations.

2. Lusternik-Schnirelmann category and the intrinsic topology

We shall consider a flow $\phi\colon X \times \mathbb{R} \to X$ defined on a locally compact metric space, X. We use the notation $\omega(x)$ (respectively $\omega^*(x)$) to denote the positive (resp. negative) limit set of the point x, i.e., the set of points $y \in X$ such that there exists a sequence $t_n \to +\infty$ (resp. $t_n \to -\infty$) and $xt_n \to y$. A Morse decomposition of an isolated invariant compactum K is a collection $D = \{M_1, \dots, M_k\}$ of disjoint closed invariant subsets of K such that for every $x \in K$ there are integers i and j, with $i \le j$, such that $\omega(x) \subset M_i$, $\omega^*(x) \subset M_j$ and, if $i = j$, $x \in M_i = M_j$. See [6] for information concerning Morse decompositions. The unstable manifold $W^u(K)$ of an isolated invariant compactum is the set $\{x \in X \mid \omega^*(x) \subset K\}$.

In [15] it was proved that if K is a regular isolated invariant set (i.e. if K admits an isolating neighborhood N such that the orbits which leave N never return) and $\{M_1, \dots, M_k\}$ is a Morse decomposition of K then the following inequality holds:

$$\eta(W^u(K)) \le \eta(M_1) + \cdots + \eta(M_k).$$

However, the following example shows that the condition of regularitity is essential in that result.

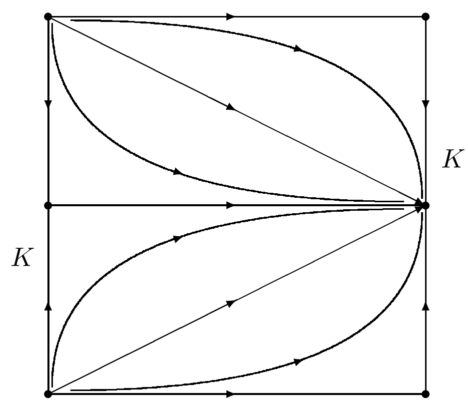


Figure 1. A flow on a torus X with an isolated invariant set K which is a copy of S^1.

In this example we have a flow defined on a torus X, the isolated invariant set K is a copy of S^1 with an attractor-repeller decomposition consisting of two stationary points and the unstable manifold $W^u(K) = X$. In this case we have $\eta(W^u(K)) = 3$ but $\eta(M_1) + \eta(M_2) = 2$.

On the other hand, Robbin and Salamon introduced a new topology on the unstable manifold of K which has proved to be very useful in the definition and study of properties related to the shape index. We briefly recall the definition of this topology:

Consider an index pair (N, L) for K (see [3] for a definition) and the associated inverse system $((N/L)_s, p_{st})$, where $(N/L)_s = N/L$ for every $s \in \mathbb{R}_+$ and, if

$s \leq t$, then $p_{st} : (N/L)_t \to (N/L)_s$ is defined by

$$p_{st}(x) = \begin{cases} x(t-s) & \text{if } x[0,t-s] \subset N - L \\ * & \text{otherwise} \end{cases}$$

(the point $*$ above corresponds to the equivalence class of L).

If we take the inverse limit $Z = \lim\left((N/L)_s, p_{st}\right)$ and denote by $\star$ the point in Z all of whose coordinates are the base point $* \in N/L$, then there is a natural map $h : Z - \{\star\} \to W^u(K)$ defined in the following way: If $\mathbf{x} = (x_s) \in Z - \{\star\}$ take $t \in \mathbb{R}_+$ such that $x_t \in N - L$. Then $h(\mathbf{x}) = x_t t$. The map h does not depend on the choice of t and it is a continuous bijection. The topology on $W^u(K)$ which makes h a homeomorphism is called the intrinsic topology. The corresponding topological space is designated by W^i. Some properties of this space have been studied in [16].

We shall prove in our first result that, if we consider the intrinsic topology on the unstable manifold, then there are no restrictions for the validity of the former result establishing a comparison between the Lusternik-Schnirelmann category of the unstable manifold of the invariant set and that of its Morse decomposition. In the sequel we shall always use proper index pairs for K, i.e. index pairs (N, L) such that every point x in the exit set L immediately leaves the neighborhood N (i.e. $x[0,t] \nsubseteq N$ for every $t > 0$). By a result of McCord [12], proper index pairs always exist.

Theorem (2.1). *Suppose that K is an isolated invariant set of the flow $\phi : X \times \mathbb{R} \to X$ and $\{M_1, \ldots, M_k\}$ is a Morse decomposition of K. Then the unstable manifold W^i of K, endowed with its intrinsic topology, has compact homotopy type and its Lusternik-Schnirelmann coefficient η satisfies the following inequality*

$$\eta(W^i) \leq \eta(M_1) + \cdots + \eta(M_k).$$

Proof. The theorem can be reduced to the case when the Morse decomposition consists of two sets, M_1 and M_2, together with the use of an induction argument. The case $k = 1$ can be handled in practice as a particular case of $k = 2$ with $M_1 = K$ and $M_2 = \emptyset$. The argument is as follows: If the theorem is true for $k - 1$ Morse sets then to prove the general case consider the attractor-repeller decomposition of K given by $M_1^* = \{\, x \in K \mid \omega(x) \cup \omega^*(x) \subset \cup_{i=1}^{k-1} M_i\}$ and M_k. If the theorem is true for two Morse sets then $\eta(W^i) \leq \eta(M_1^*) + \eta(M_k)$. If we consider now the flow restricted to M_1^*, then $\{M_1, \ldots, M_{k-1}\}$ is a Morse decomposition of M_1^*, and the $(k-1)$-dimensional version of the the theorem applied to $X = K = M_1^*$ establishes the general result. Hence, we restrict ourselves to an attractor-repeller decomposition of K, $\{M_1, M_2\}$. Let (N, L) be a proper index pair for K in X. We can assume that N is embedded in a space $M \in ANR$. Consider the asymptotic negative set $N^- = \{x \in N \mid x(-\infty, 0] \subset N\}$. Let U be an arbitrary neighborhood of N^- in M. If $\eta(M_1) = l_1$ and $\eta(M_2) = l_2$ then, using elementary properties of ANRs, it is possible to see that there are closed neighborhoods (in K) W_1 and W_2 of M_1 and M_2, respectively, that can be represented as unions $W_1 = \bigcup_{i=1}^{l_1} W_1^i$, $W_2 = \bigcup_{i=1}^{l_2} W_2^i$ of compact sets which are contractible in U. Moreover, since $\{M_1, M_2\}$ is an attractor-repeller decomposition of K, the use of Lyapunov functions allows us to show that K can be decomposed into two closed sets K_1 and K_2 which are deformable in K into subsets of W_1 and W_2 respectively and, hence, representable as $K_1 = \bigcup_{i=1}^{l_1} K_1^i$, $K_2 = \bigcup_{i=1}^{l_2} K_2^i$ with K_1^i and K_2^i contractible in U. Now, using again properties of ANRs, we can find a neighborhood $\hat{N}$ of K in X wich admits a decomposition

$$\hat{N} = \bigcup_{i=1}^{l_1+l_2} N_i$$

with N_i contractible in U. It has been proved in [15] that N^- is deformable to a subset of $\hat{N}$ and, hence, N^- is also representable as the union of l_1+l_2 compacta contractible in U; as a consequence, $\eta(N^-) \leq l_1 + l_2$.

Consider now the unstable manifold W^i endowed with the intrinsic topology. Define $n^- = N^- \cap L$. For every point $x \in W^i - N^-$ we select the only point in n^- contained in the trajectory of x. This point can be expressed as xt_x with $t_x < 0$. We claim that the map $\alpha : W^i \to \mathbb{R}$ such that $\alpha(x) = t_x$ if $x \in W^i - N^-$ and $\alpha(x) = 0$ otherwise is continuous. To see this, suppose that we have a sequence of points $x_n \to x$ in W^i. We consider the only non-trivial case, when $x \notin N^- - n^-$. This convergence in the intrinsic topology is equivalent to the fact that there is an $s < 0$ such that: a) $x_n(s)$ and $x(s)$ belong to $N^- - n^-$ and b) $x_n(s) \to x(s)$ with the extrinsic topology inherited from X. Take now the positive numbers s_{x_n} and s_x such that $x_n(s + s_{x_n})$ and $x(s + s_x)$ lie in n^-. The sequence s_{x_n} is bounded since, otherwise, we would have intervals of trajectories $[x_n s, x_n(s + s_{x_n})]$ contained in N, with s_{x_n} arbitrarily large, and this would imply that the trajectory of $x(s)$ would be contained in N and N would not be isolating. Hence, there is a subsequence of s_{x_n} which converges to a number s_0. We can assume, without loss of generality, that $s_{x_n} \to s_0$. Thus $x_n(s + s_{x_n}) \to x(s + s_0)$. Since n^- is compact, $x(s + s_0) \in n^-$ and, hence, $s_0 = s_x$. As a consequence, $t_{x_n} = s_{x_n} + s \to t_{x_0} = s_{x_0} + s$ and α is continuous. We consider now the homotopy

$$\Phi : W^i \times [0,1] \to W^i$$

defined by $\Phi(x,s) = x(st_x)$. This map is continuous since it can be considered as a composition of the map $W^i \times [0,1] \to W^i \times \mathbb{R}$ given by $(x,s) \to (x, st_x)$, which is continuous, and the flow restricted to W^i, which is also continuous with the intrinsic topology. The homotopy Φ realizes a strong deformation retraction of W^i into N^- and, hence, W^i has compact homotopy type. Moreover, we have already seen that $\eta(N^-) \leq l_1 + l_2$. Hence $\eta(W^i) \leq \eta(M_1) + \eta(M_2)$, which proves the case $k = 2$. The general case is a consequence of the argument established before. □

3. Duality and Morse sets

We study in this section some properties of flows that are related to the topological situation of Hopf duality. This situation arises when $(n-1)$-manifolds, W, embedded in the n-sphere S^n, induce homological relations between the two n-manifolds with boundary into which S^n is decomposed by W (see Steenrod and Epstein [17]). We present here a much more general situation applicable to a Morse decomposition $\{M_0, M_1, M_2\}$ of a manifold M which is an isolated invariant set of a flow $\phi : X \times \mathbb{R} \to X$, where M_0 is an $(n-1)$-submanifold of M decomposing M into two manifolds with common boundary M_0, and M_1

and M_2 are general Morse sets (not necessarily manifolds). We recall that connected Morse decomposition means a decomposition where all the Morse sets are connected.

THEOREM (3.1). *Let $\phi : X \times \mathbb{R} \to X$ be a flow defined on a locally compact metric space X. Let $M \subset X$ be an orientable, compact, connected n-dimensional manifold which is an attractor of ϕ. Suppose that $H^k(M) = H^{k+1}(M) = \{0\}$ for a given index k. Let $\{M_0, M_1, M_2\}$ be a connected Morse decomposition of M, where M_0 is an $(n-1)$-submanifold of M, decomposing M into two manifolds with common boundary M_0. We then have the following relations involving the homological and cohomological Conley indices: 1) $CH^{k+1}(M_1) = CH_{n-k}(M_2)$ and 2) $CH_{n-k-1}(M_0) = CH^{k+1}(M_1) \oplus CH^{k+1}(M_2)$.*

If we only assume that M is an isolated invariant set of ϕ (not necessarily an attractor) then, with the same hypotheses as above, we have the following relations involving Čech homology and cohomology of the Morse sets: 1) $\check{H}^k(M_1) = \check{H}_{n-k-1}(M_2)$ and 2) $\check{H}^k(M_0) = \check{H}^k(M_1) \oplus \check{H}^k(M_2)$.

Proof. We shall only prove the statement in the theorem concerning the Conley indices since the proof of the second part, which refers to Čech homology, involves the same kind of ideas and can be left to the reader.

Since M_0 separates M, the Morse sets M_2 and M_1 are contained in different components of $M - M_0$. The orbits $\gamma(x) \subset M - M_2$ with $\omega^*(x) \subset M_2$ have their ω-limit contained in M_0 and, as a consequence, the set $\mathcal{M}_{02} = \{x \in M \mid \omega^*(x) \subset M_2\} \cup M_0$ is an isolated invariant set of X and, in fact, an attractor of the flow restricted to M (and, hence, an attractor of the flow ϕ). Moreover, (M_0, M_2) is an attractor-repeller decomposition of $\mathcal{M}_{02}$. Let U be a positively invariant compact neighborhood of $\mathcal{M}_{02}$ in X and consider another positively invariant compact neighborhood U_0 of M_0 not meeting M_2 and such that U_0 is contained in the interior of U. Then (U, U_0) is an index pair for the Morse set M_2. We select a nested sequence (U^n, U_0^n) of similar index pairs with $\cap U^n = \mathcal{M}_{02}$ and $\cap U_0^n = M_0$. By using the properties of the flow, it is possible to define a sequence of maps $r_n : (U^1, U_0^1) \to (U^n, U_0^n)$ such that $j_n r_n \simeq i_1$ and $r_n j_n \simeq i_n$, where $j_n : (U^n, U_0^n) \to (U^1, U_0^1)$ is the inclusion and $i_n : (U^n, U_0^n) \to (U^n, U_0^n)$ is the identity. The homotopies considered here are those of pairs. It follows that $j_{nn-1} r_n \simeq r_{n-1}$, where $j_{nn-1} : (U^n, U_0^n) \to (U^{n-1}, U_0^{n-1})$ is the inclusion, and then we have that the induced homomorphisms between Čech homology groups $(r_n)_* : \check{H}_*(U^1, U_0^1) \to \check{H}_*(U^n, U_0^n)$ are isomorphisms and, by continuity of Čech homology, we get that $\check{H}_*(U^1, U_0^1)$ is isomorphic to $\check{H}_*(\mathcal{M}_{02}, M_0)$. We deduce from this that the homological Conley index $CH_*(M_2)$ is $\check{H}_*(\mathcal{M}_{02}, M_0)$.

The hypotheses in the theorem imply that $\mathcal{M}_{02}$ is a compact orientable n-manifold with boundary M_0. Then, using Lefschetz duality, we get that $H_{n-k}(\mathcal{M}_{02}, M_0) = H^k(\mathcal{M}_{02})$. Moreover, the pair $(M_1, \mathcal{M}_{02})$ is a repeller-attractor decomposition of M and, considering the long exact sequence of cohomological Conley indices associated to such decomposition,

$$\cdots \to CH^k(M) \to CH^k(\mathcal{M}_{02}) \to CH^{k+1}(M_1) \to CH^{k+1}(M) \to \cdots,$$

and the fact that $\{0\} = H^k(M) = CH^k(M)$ and $\{0\} = H^{k+1}(M) = CH^{k+1}(M)$ (since M is an attractor), we get that $CH^{k+1}(M_1) = CH^k(\mathcal{M}_{02})$. On the other hand we have that $CH^k(\mathcal{M}_{02}) = H^k(\mathcal{M}_{02})$ for the same reason as before and we conclude from this that $CH_{n-k}(M_2) = CH^{k+1}(M_1)$.

We consider now the triad $(M, \mathcal{M}_{02}, \mathcal{M}_{01})$, where $\mathcal{M}_{01} = \{x \in M \mid \omega^*(x) \subset M_1\} \cup M_0$. $\mathcal{M}_{01}$ is, as $\mathcal{M}_{02}$, a compact orientable n-manifold with boundary, and $\mathcal{M}_{02} \cap \mathcal{M}_{01} = M_0$. According to Dold [4], p. 291, this triad is excisive and, hence, it has an associated Mayer-Vietoris sequence in homology

$$\cdots \to H_{n-k}(M) \to H_{n-k-1}(M_0) \to$$
$$H_{n-k-1}(\mathcal{M}_{02}) \oplus H_{n-k-1}(\mathcal{M}_{01}) \to H_{n-k-1}(M) \to \cdots$$

By the Poincaré duality theorem $H_{n-k}(M) = H_{n-k-1}(M) = \{0\}$ and, since M_0 is an attractor, we have that

$$CH_{n-k-1}(M_0) = H_{n-k-1}(M_0) = H_{n-k-1}(\mathcal{M}_{01}) \oplus H_{n-k-1}(\mathcal{M}_{02}).$$

Now, arguments involving Lefschetz duality, similar to those used before, show that $H_{n-k-1}(\mathcal{M}_{01}) = CH^{k+1}(M_1)$ and $H_{n-k-1}(\mathcal{M}_{02}) = CH^{k+1}(M_2)$. This establishes the second equality in the statement concerning the Conley indices and completes the proof. $\square$

The second part of Theorem 2 admits a different, more general, version, in which the homological hypothesis is placed on the unstable manifold of M with its intrinsic topology. This can be obtained as the following consequence of the theorem:

Corollary (3.2). *Let $\phi : X \times \mathbb{R} \to X$ be a flow defined on a locally compact metric space X. Let $M \subset X$ be an orientable, compact, connected n-dimensional manifold which is an isolated invariant set of ϕ. Suppose that $\check{H}^k(W^i(M)) = \check{H}^{k+1}(W^i(M)) = \{0\}$ for a given index k. Let $\{M_0, M_1, M_2\}$ be a connected Morse decomposition of M, where M_0 is an $(n-1)$-submanifold of M, decomposing M into two manifolds with common boundary M_0. Then, we have the following relations involving Čech homology and cohomology of the Morse sets: 1) $\check{H}^k(M_1) = \check{H}_{n-k-1}(M_2)$ and 2) $\check{H}^k(M_0) = \check{H}^k(M_1) \oplus \check{H}^k(M_2)$.*

Proof. The isolated invariant set K can be represented as the intersection of a nested sequence of isolating neighborhoods N_i such that every N_i has an exit set L_i with (N_i, L_i) a proper index pair. In the proof of Theorem 1 it has been shown that the unstable manifold W^i with its intrinsic topology is homotopically equivalent to N_i^- for every i. Since K is also the intersection of the nested sequence of the N_i^- we have, by the continuity property of Čech cohomology, that $\check{H}^k(K) = \check{H}^k(W^i(M)) = \{0\}$ and $\check{H}^{k+1}(K) = \check{H}^{k+1}(W^i(M)) = \{0\}$. Hence the corollary is a consequence of Theorem 2. $\square$

Received June 30, 2003

Final version received March 10, 2004

Facultad de Matemáticas
Universidad Complutense
28040 Madrid
Spain
jose_ sanjurjo@mat.ucm.es

References

[1] K. Borsuk, *On the Lusternik-Schnirelmann category in the theory of shape*, Fund. Math. **99** (1978), 35–42

[2] N. P. Bhatia and G.P. Szegö, Stability theory of dynamical systems, Springer-Verlag, Berlin-Heidelberg-New York, 1970.

[3] C. Conley, *Isolated invariant sets and the Morse index*, CBMS Regional Conference Series in Mathematics **38**, Amer. Math. Soc. Providence, RI, 1978.

[4] A. Dold, Lectures on Algebraic Topology, Springer-Verlag, Berlin-Heidelberg-New York 1972.

[5] R.H. Fox, *On the Lusternik-Schnirelmann category*, Ann. Math. **42** (1941), 333–370.

[6] A. Giraldo and J.M.R. Sanjurjo, *On the global structure of invariant regions of flows with asymptotically stable attractors*, Mathematische Zeitschrift **232** (1999), 739–746.

[7] M. Gobbino, *Topological properties of attractors for dynamical systems*, Topology **40** (2000), 279–298.

[8] A. Hatcher, Algebraic Topology, Cambridge University Press, Cambridge 2001.

[9] I.M. James, *On category in the sense of Lusternik-Schnirelmann*, Topology **17** (1978), 331–348

[10] I.M. James, Lusternik-Schnirelmann category. In Handbook of Algebraic Topology (Elsevier, 1995), 1293–1310.

[11] L. Kapitanski and I. Rodnianski, *Shape and Morse theory of attractors*, Comm. Pure Appl. Math. **53** (2000), 218–242.

[12] C. McCord, *Mappings and homological properties in the Conley index theory*, Ergodic Th. and Dynam. Sys. **8*** (1988), 175–198.

[13] K.P. Rybakowsky, The homotopy index and partial differential equations, Universitext, Springer-Verlag, Berlin-Heidelberg-New York 1987.

[14] J.W. Robbin and D. Salamon, *Dynamical systems, shape theory and the Conley index*, Ergodic Th. and Dynam. Sys. **8*** (1988), 375–393.

[15] J.M.R. Sanjurjo, *Lusternik-Schnirelmann category and Morse decompositions*, Mathematika **47** (2000), 299–305.

[16] J.M.R. Sanjurjo, *Morse equations and unstable manifolds of isolated invariant sets*, Nonlinearity **16** (2003) 1435–1448.

[17] N.E. Steenrod and D.B.A. Epstein, Cohomology operations, Princeton Univ. Press, Princeton, NJ, 1962.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

THE JØRGENSEN NUMBER OF THE WHITEHEAD LINK GROUP

HIROKI SATO

Dedicated to Professor Francisco Javier Gonzalez Acuña on his sixtieth birthday

ABSTRACT. In this paper we consider the Jørgensen number of the Whitehead link group. The result is as follows: The Jørgensen number of the Whitehead link group is two. Furthermore, we will represent the point corresponding to the Whitehead link group by using the coordinates introduced in Sato [12].

1. Introduction

(1.1) In 1976 Jørgensen obtained the following important theorem, called Jørgensen's inequality, which gives a necessary condition for a non-elementary Möbius transformation group $G = \langle A, B \rangle$ to be discrete.

THEOREM (A) (Jørgensen [1]). *Suppose that the Möbius transformations A and B generate a non-elementary discrete group. Then*

$$J(A, B) := |\operatorname{tr}^2(A) - 4| + |\operatorname{tr}(ABA^{-1}B^{-1}) - 2| \geq 1\,.$$

The lower bound 1 is best possible.

Definition (1.1.1). Let A and B be Möbius transformations. The *Jørgensen number* $J(A, B)$ of the ordered pair (A, B) is defined by

$$J(A, B) := |\operatorname{tr}^2(A) - 4| + |\operatorname{tr}(ABA^{-1}B^{-1}) - 2|\,.$$

We denote by Möb the set of all orientation-preserving Möbius transformations. We recall that Möb ($= PSL(2, \mathbb{C})$) acts on the upper half space H^3 of $\mathbb{R}^3$ as the group of conformal isometries of hyperbolic 3-space. A subgroup G of Möb is said to be *elementary* if there exists a finite G-orbit in $\hat{\mathbb{R}}^3$.

Definition (1.1.2). Let G be a non-elementary two-generator subgroup of Möb. The *Jørgensen number* $J(G)$ for G is defined by $J(G) := \inf\{J(A, B) \mid A$ and B generate $G\}$.

Definition (1.1.3). A non-elementary two-generator subgroup G of Möb is a *Jørgensen group* if G is a discrete group with $J(G) = 1$.

Jørgensen and Kiikka showed the following for Jørgensen groups.

2000 *Mathematics Subject Classification*: Primary 30F40; Secondary 20H10, 57M25.

Keywords and phrases: Jørgensen's inequality, Jørgensen number, Jørgensen group, the Whitehead link group.

Partly supported by the Grants-in-Aid for Co-operative Research as well as Scientific Research, the Ministry of Education, Science, Sports and Culture, Japan.

Theorem (B) (Jørgensen-Kiikka [2]). *Let $\langle A, B\rangle$ be a non-elementary discrete group with $J(A,B) = 1$. Then either A is elliptic of order at least seven or A is parabolic.*

If $\langle A, B\rangle$ is a Jørgensen group such that A is parabolic and $J(A,B) = 1$, then we call it a *Jørgensen group of parabolic type.* There is an infinite number of Jørgensen groups (see Jørgensen-Lascurain-Pignataro [3], Li-Oichi-Sato [6], [7], [8], Sato [12]). The following familiar groups are all Jørgensen groups of parabolic type: The modular group, the Picard group (Jørgensen-Lascurain-Pignataro [3], Sato [13], Sato-Yamada [14]), the figure-eight knot group [12], "the Gehring-Maskit group" [12], where "the Gehring-Maskit group" is the group studied in Maskit [10].

The link group of a link L is the fundamental group of the complement space $S^3 - L$. Since, by Mostow's rigidity theorem, the link group of a hyperbolic link has essentially one faithful discrete $PSL(2,\mathbb{C})$-representation, the link group is naturally identified with a Kleinian (discrete) group. In this paper we call the link group of the Whitehead link the *Whitehead link group*, that is, the Whitehead link group is the discrete faithful representation of the fundamental group of the Whitehead link. Now this gives rise to the following problems.

Problem (1.1.4). Is the Whitehead link group a Jørgensen group?

Problem (1.1.5). If the Whitehead link group is not a Jørgensen group, find the Jørgensen number of the group.

In this paper we will give the answers to the problems, that is, we have the following theorem and corollary.

Theorem (1.1.6). *The Jørgensen number of the Whitehead link group is two.*

Corollary (1.1.7). *The Whitehead link group is not a Jørgensen group.*

(1.2) Throughout this paper we will always write elements of Möb as matrices with determinant 1. Let $\langle A, B\rangle$ be a marked two-generator group such that A is parabolic. Then we can normalize A and B as follows:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B := B_{\sigma,\mu} = \begin{pmatrix} \mu\sigma & \mu^2\sigma - 1/\sigma \\ \sigma & \mu\sigma \end{pmatrix},$$

where $\sigma \in \mathbb{C} \setminus \{0\}$ and $\mu \in \mathbb{C}$. We denote by $G_{\sigma,\mu}$ the marked group generated by A and $B_{\sigma,\mu}$: $G_{\sigma,\mu} = \langle A, B_{\sigma,\mu}\rangle$. We say that $(\sigma,\mu) \in (\mathbb{C} \setminus \{0\}) \times \mathbb{C}$ is the point representing a marked group $G_{\sigma,\mu}$ and that $G_{\sigma,\mu}$ is the marked group corresponding to a point (σ,μ).

In Li-Oichi-Sato [6], [7], [8] and Sato [12], we considered the case of $\mu = ik$ $(k \in \mathbb{R})$. Namely, we considered the marked two-generator group $G_{\sigma,ik} = \langle A, B_{\sigma,ik}\rangle$ generated by

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B_{\sigma,ik} = \begin{pmatrix} ik\sigma & -k^2\sigma - 1/\sigma \\ \sigma & ik\sigma \end{pmatrix},$$

where $\sigma \in \mathbb{C} \setminus \{0\}$ and $k \in \mathbb{R}$.

Now we have the following conjecture.

Conjecture (1.2.1). *For any Jørgensen group G of parabolic type there exists a marked group $G_{\sigma,ik}$ $(\sigma \in \mathbb{C} \setminus \{0\}, k \in \mathbb{R})$ such that $G_{\sigma,ik}$ is conjugate to G.*

If this conjecture is true, then it is sufficient to consider the case of $\mu = ik$ in order to find all Jørgensen groups of parabolic type. We found all Jørgensen groups $G_{\sigma,ik}$ of parabolic type (Li-Oichi-Sato [6], [7], [8] and Sato [12]).

(1.3) Let C be the following cylinder: $C = \{(\sigma, ik) \mid |\sigma| = 1,\ k \in \mathbb{R}\}$.

Theorem (C) (Sato [12]). *If a marked two-generator group $G_{\sigma,ik} = \langle A, B_{\sigma,ik}\rangle$ $(\sigma \in \mathbb{C} \setminus \{0\}, k \in \mathbb{R})$ is a Jørgensen group such that $J(A, B_{\sigma,ik}) = 1$, then the point (σ, ik) representing $G_{\sigma,ik}$ lies on the cylinder C.*

If we set $\sigma = -ire^{i\theta}$, which is used in [12], then we have the following theorem.

Theorem (D) (Jørgensen-Lascurain-Pignataro [3], Sato [12], [13], Sato-Yamada [14]). (i) *The group $G_{\sigma,ik}$ with $\sigma = -ie^{\pi i/2}$ and $k = 0$ is conjugate to the modular group.*

(ii) *The group $G_{\sigma,ik}$ with $\sigma = -ie^{\pi i/2}$ and $k = 1/2$ is conjugate to the Picard group.*

(iii) *The group $G_{\sigma,ik}$ with $\sigma = -ie^{\pi i/6}$ and $k = \sqrt{3}/2$ is conjugate to the figure-eight knot group.*

(iv) *The group $G_{\sigma,ik}$ with $\sigma = -i$ and $k = \sqrt{3}/2$ is conjugate to the "Gehring - Maskit group".*

Now this gives rise to the following problem.

Problem (1.3.1). *Represent the Whitehead link group by using the coordinates (σ, ik) introduced in §1.2, that is, by $G_{\sigma,ik}$ $(\sigma \in \mathbb{C} \setminus \{0\}, k \in \mathbb{R})$.*

For this problem we have the following theorem.

Theorem (1.3.2). *The Whitehead link group is conjugate to the marked two-generator group $G_{\sigma,ik}$ where $\sigma = \sqrt{2}e^{3\pi i/4}$ and $k = -1/2$.*

Finally we note that this paper is closely connected with the following problem:

Problem (1.3.3). Let r be a real number with $r \geq 1$. When is there a discrete group whose Jørgensen number is equal to r?

The methods of the proofs of Propositions 2 and 3 are applicable to this problem for every natural number $r \geq 2$ (Li-Oichi [5]).

The author expresses gratitude to Professor M. Sakuma for his suggestions about knot theory. Thanks are due to the referees for their careful reading and valuable suggestions.

2. Proof of Theorem (1.1.6).

In this section we will give a proof of Theorem (1.1.6).

Theorem (E) (cf. Wielenberg [15], Krushkal', Apanasov and Gusevskiĭ [4]). *The Whitehead link group G_W has the following presentation:*

$$G_W = \langle A, B \mid (A^{-1}BAB^{-1})(ABA^{-1}B^{-1})(AB^{-1}A^{-1}B)(A^{-1}B^{-1}AB) = I\rangle,$$

where

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 1-i & 1 \end{pmatrix},$$

and I is the identity matrix.

Proposition (2.1). *Let G_W be the Whitehead link group generated by A and B in Theorem (E). Then an element X of G_W has the following form:*

$$X = \begin{pmatrix} 1+(1-i)a & b \\ (1-i)c & 1+(1-i)d \end{pmatrix}.$$

where $a, b, c, d \in \mathbb{Z} + i\mathbb{Z}$, $(1-i)(a+d-bc) - 2iad = 0$.

Proof. We will prove this proposition by induction. Let

$$X = \begin{pmatrix} 1+(1-i)a & b \\ (1-i)c & 1+(1-i)d \end{pmatrix}.$$

be an element of G_W, where $a, b, c, d \in \mathbb{Z} + i\mathbb{Z}$, $(1-i)(a+d-bc) - 2iad = 0$.

For the matrix A, we can take $a = c = d = 0$, $b = 1$. For the matrix B, we can take $a = b = d = 0$, $c = 1$.

Let

$$Y = \begin{pmatrix} 1+(1-i)\alpha & \beta \\ (1-i)\gamma & 1+(1-i)\delta \end{pmatrix}.$$

be an element of G_W, where $\alpha, \beta, \gamma, \delta \in \mathbb{Z} + i\mathbb{Z}$, $(1-i)(\alpha+\delta-\beta\gamma) - 2i\alpha\delta = 0$.

If we set

$$XY = \begin{pmatrix} x & y \\ z & u \end{pmatrix},$$

then

$$\begin{aligned}
x &= 1 + (1-i)\{a + \alpha + b\gamma + (1-i)a\alpha\} \\
y &= (b+\beta) + (1-i)(\alpha\beta + b\delta) \\
z &= (1-i)\{c + \gamma + (1-i)(c\alpha + d\gamma)\} \\
u &= 1 + (1-i)\{d + \delta + c\beta + (1-i)d\delta\}.
\end{aligned}$$

Thus $XY \in G_W$. Our proof is now complete. □

Proposition (2.2). *Let G_W be the Whitehead link group in Theorem (E). Let $\langle X, Y\rangle$ be a non-elementary subgroup generated by X and Y, where $X, Y \in G_W$. Then the Jørgensen number of (X, Y) is greater than or equal to two: $J(X, Y) \geq 2$.*

Proof. Let

$$X = \begin{pmatrix} 1+(1-i)a & b \\ (1-i)c & 1+(1-i)d \end{pmatrix}, \qquad Y = \begin{pmatrix} 1+(1-i)\alpha & \beta \\ (1-i)\gamma & 1+(1-i)\delta \end{pmatrix},$$

where $a, b, c, d \in \mathbb{Z} + i\mathbb{Z}$, $(1-i)(a+d-bc) - 2iad = 0$ and $\alpha, \beta, \gamma, \delta \in \mathbb{Z} + i\mathbb{Z}$, $(1-i)(\alpha+\delta-\beta\gamma) - 2i\alpha\delta = 0$.

We set

$$XYX^{-1}Y^{-1} = \begin{pmatrix} x & y \\ z & u \end{pmatrix}.$$

Then by straightforward calculations we have

$$\begin{aligned} x &= 1+(1-i)\{a+\alpha+d+\delta+2b\gamma-(b+\beta)(c+\gamma)\}+ \\ &\quad (1-i)^2m_2+(1-i)^3m_3+(1-i)^4m_4 \end{aligned}$$

for some $m_2, m_3, m_4 \in \mathbb{Z}+i\mathbb{Z}$, and

$$\begin{aligned} u &= 1+(1-i)\{a+\alpha+d+\delta+2c\beta-(b+\beta)(c+\gamma)\}+ \\ &\quad (1-i)^2n_2+(1-i)^3n_3+(1-i)^4n_4 \end{aligned}$$

for some $n_2, n_3, n_4 \in \mathbb{Z}+i\mathbb{Z}$.

Thus

$$\begin{aligned} \operatorname{tr}(XYX^{-1}Y^{-1}) &= 2+2(1-i)\{a+\alpha+d+\delta+b\gamma+c\beta-(c+\gamma)(b+\beta)\} \\ &\quad +(1-i)^2(m_2+n_2)+(1-i)^3(m_3+n_3)+(1-i)^4(m_4+n_4)\,. \end{aligned}$$

Hence we have

$$\begin{aligned} |\operatorname{tr}(XYX^{-1}Y^{-1})-2| &= 2|\{a+d+\alpha+\delta+b\gamma+c\beta-(c+\gamma)(b+\beta)\}(1-i)- \\ &\quad i(m_2+n_2)-(1+i)(m_3+n_3)-2(m_4+n_4)|\,. \end{aligned}$$

We set

$$\begin{aligned} K &= \{a+d+\alpha+\delta+b\gamma+c\beta-(c+\gamma)(b+\beta)\}(1-i)-i(m_2+n_2)- \\ &\quad (1+i)(m_3+n_3)-2(m_4+n_4). \end{aligned}$$

Then $K \in \mathbb{Z}+i\mathbb{Z}$ and

$$|\operatorname{tr}(XYX^{-1}Y^{-1})-2| = 2|K|.$$

Since $\langle X, Y\rangle$ is non-elementary, we have $K \neq 0$. Hence

$$|\operatorname{tr}(XYX^{-1}Y^{-1})-2| \geq 2.$$

Thus $J(X,Y) \geq 2$, since $|\operatorname{tr}^2(X)-4| \geq 0$. □

Proposition (2.3). *Let A, B be the matrices in Theorem (E). Set $C = AB$. Then A and C generate the Whitehead link group G_W and $J(A,C) = 2$.*

Proof. We can see by easy calculations that

$$C = AB = \begin{pmatrix} 2-i & 1 \\ 1-i & 1 \end{pmatrix},$$

and so we have $J(A,C) = |1-i|^2 = 2$. □

Proof of Theorem (1.1.6). Theorem (1.1.6) follows from Propositions (2.2) and (2.3). □

Corollary (1.1.7) follows from Theorem (1.1.6) and Theorem (C).

Let

$$T = \begin{pmatrix} 1 & -1/2 \\ 0 & 1 \end{pmatrix}.$$

We set $A^* = TAT^{-1}$ and $C^* = TCT^{-1}$. Then we have

$$A^* = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad C^* = \begin{pmatrix} (3-i)/2 & (5-i)/4 \\ 1-i & (3-i)/2 \end{pmatrix}.$$

By easy calculations we have the following.

Proposition (2.4). *Let A^* and C^* be the above matrices. Then the group $\langle A^*, C^*\rangle$ generated by A^* and C^* is conjugate to the Whitehead link group G_W. The point representing the marked group $\langle A^*, C^*\rangle$ is $(\sigma, \mu) = (1-i, (2+i)/2)$.*

Remark (2.5). The marked group $\langle A^*, C^*\rangle$ is not a group of type $G_{\sigma,ik}$ ($\sigma \in \mathbb{C}\setminus\{0\}, k \in \mathbb{R}$).

3. Proof of Theorem (1.3.2)

In this section we will prove Theorem (1.3.2). We use Poincaré's polyhedron theorem for the proof of Proposition (3.2). See Maskit [9] for Poincaré's polyhedron theorem and the terminology associated with the theorem, for example, a side pairing transformation, a cycle transformation at the edge.

Let P be the regular ideal octahedron in Ratcliffe [11], p.454, Figure 10.3.14]. We denote the sides of P by $S_A, S_B, S_C, S_D, S_{A'}, S_{B'}, S_{C'}$ and $S_{D'}$, which are the sides A, B, C, D, A', B', C' and D', respectively, in Figure 10.3.14 in the Ratcliffe's book. Let a, b and c be the same sides of P as in Figure 10.3.14. Furthermore, the vertices v_j $(j = 1, 2, \cdots, 6)$ are defined as follows: Let v_1 be the common vertex of the sides S_A, S_B, S_C and S_D. For simplicity we represent v_1 by $v_1 = [S_A, S_B, S_C, S_D]$. Similarly we set $v_2 = [S_A, S_D, S_{C'}, S_{D'}]$, $v_3 = [S_A, S_{D'}, S_{A'}, S_B]$, $v_4 = [S_B, S_{A'}, S_{B'}, S_C]$, $v_5 = [S_D, S_{C'}, S_{B'}, S_C]$ and $v_6 = [S_{A'}, S_{B'}, S_{C'}, S_{D'}]$.

Let f_A, f_B, f_C and f_D be the side pairing transformations of S_A to $S_{A'}$, of S_B to $S_{B'}$, of S_C to $S_{C'}$, and of S_D to $S_{D'}$, respectively.

Proposition (3.1) (cf. Ratcliffe [11]). *Let f_A, f_B, f_C and f_D be the side pairing transformations defined in the above. Then f_A, f_B, f_C and f_D generate the Whitehead link group $G_{W,R}$ in the sense of Ratcliffe.*

Let A, B, C and D the matrices corresponding to the Möbius transformations f_A, f_B, f_C and f_D, respectively. We use the same notation $G_{W,R}$ for the group generated by A, B, C and D.

Proposition (3.2). *Let*

$$D = \begin{pmatrix} (1-i)/2 & (1-i)/2 \\ (1-i)/2 & (3+i)/2 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} (1-i)/2 & (1-i)/2 \\ (-1+i)/2 & (1+3i)/2 \end{pmatrix}.$$

Then the Whitehead link group $G_{W,R}$ in Proposition (3.1) has the following presentation:

$$G_{W,R} = \langle D, C \mid DC^{-2}DCD^{-1}C^{-1}D^{-1}C^2D^{-1}C^{-1}DC = I\rangle.$$

Proof. Let P be the regular ideal octahedron in Ratcliffe [11], p.454. Let the sides S_A, S_B, S_C, S_D; $S_{A'}, S_{B'}, S_{C'}, S_{D'}$, the edges a, b, c and the vertices v_j $(j = 1, 2, \cdots, 6)$ be as defined in the above. We normalize $v_1 = \infty$, $v_3 = 1$, and $v_6 = 0$. Then we have $v_2 = -i$, $v_4 = i$ and $v_5 = -1$.

Let f_A be the side pairing transformation of S_A to $S_{A'}$ with $f_A(v_1) = v_3$, $f_A(v_2) = v_4$, $f_A(v_3) = v_6$. Then we have $f_A(z) = (z-1)/(z+(2i-1))$.

Let f_B be the side pairing transformation of S_B to $S_{B'}$ with $f_B(v_1) = v_5$, $f_B(v_3) = v_6$, $f_B(v_4) = v_4$. Then we have $f_B(z) = (-z+1)/(z-(2i+1))$.

Let f_C be the side pairing transformation of S_C to $S_{C'}$ with $f_C(v_1) = v_5$, $f_C(v_5) = v_6,\ f_C(v_4) = v_2$. Then we have $f_C(z) = (z+1)/(-z+(2i-1))$.

Let f_D be the side pairing transformation of S_D to $S_{D'}$ with $f_D(v_1) = v_3,\ f_D(v_2) = v_2,\ f_D(v_5) = v_6$. Then we have $f_D(z) = (z+1)/(z+(2i+1))$.

By considering the cycle transformations at the edges a, b and c we have the relations $f_B^{-1} f_C f_D^{-1} f_A = I,\ f_C^{-1} f_D^{-1} f_C f_B = I$ and $f_A f_D^{-1} f_A^{-1} f_B = I$, respectively, where I is the identity mapping. From these relations we obtain the following relation:

$$f_D f_C^{-2} f_D f_C f_D^{-1} f_C^{-1} f_D^{-1} f_C^2 f_D^{-1} f_C^{-1} f_D f_C = I.$$

Let D and C the matrices corresponding to the Möbius transformations f_D and f_C, respectively. Then we have the presentation of the Whitehead link group $G_{W,R}$ as follows:

$$G_{W,R} = \langle D, C \mid DC^{-2}DCD^{-1}C^{-1}D^{-1}C^2D^{-1}C^{-1}DC = I\rangle.$$

□

Proposition (3.3). *Let*

$$G^*_{W,R} = \langle D^*, C^* \mid D^*(C^*)^{-2}D^*C^*(D^*)^{-1}(C^*)^{-1}$$
$$(D^*)^{-1}(C^*)^2(D^*)^{-1}(C^*)^{-1}D^*C^* = I\rangle,$$

where

$$D^* = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad C^* = \begin{pmatrix} (1+i)/2 & (3+i)/4 \\ -(1-i) & (1+i)/2 \end{pmatrix}.$$

Then (i) $G^*_{W,R}$ *is conjugate to the Whitehead link group* $G_{W,R}$ *in Proposition (3.2).*

(ii) $J(D^*, C^*) = 2$.

Proof. (i) We set

$$R = \begin{pmatrix} 2^{1/4}e^{-3\pi i/8}a & 2^{1/4}e^{-3\pi i/8}b \\ 2^{-1/4}e^{3\pi i/8} & 2^{-1/4}e^{3\pi i/8}i \end{pmatrix},$$

where $a = -(1+3i)/4$ and $b = -(1+i)/4$.

Let D and C be the matrices as in Proposition (3.2). Then

$$RDR^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad RCR^{-1} = \begin{pmatrix} (1+i)/2 & (3+i)/4 \\ -(1-i) & (1+i)/2 \end{pmatrix}.$$

Thus $D^* = RDR^{-1}$ and $C^* = RCR^{-1}$. Hence $G^*_{W,R}$ is conjugate to $G_{W,R}$.

(ii) By easy calculations we have that

$$D^*C^*(D^*)^{-1}(C^*)^{-1} = \begin{pmatrix} -2i & 2-i/2 \\ -2i & 2 \end{pmatrix}.$$

Hence we have $J(D^*, C^*) = 2$. □

We easily see the following.

Proposition (3.4). *Let* D^* *and* C^* *be the matrices in Proposition (3.3). Then the point representing the marked group* $\langle D^*, C^*\rangle$ *is* $(\sigma, ik) = (\sqrt{2}e^{3\pi i/4}, -i/2)$.

Proof of Theorem (1.3.2). Theorem (1.3.2) follows from Propositions (3.3) and (3.4). □

Received June 20, 2003

Final version received February 16, 2004

Department of Mathematics
Faculty of Science
Shizuoka University
836 Ohya, Shizuoka 422-8529
Japan
smhsato@ipc.shizuoka.ac.jp

References

[1] T. Jørgensen, *On discrete groups of Möbius transformations*, Amer. J. Math. **98** (1976), 739–749.
[2] T. Jørgensen and M. Kiikka, *Some extreme discrete groups*, Ann. Acad. Sci. Fenn. **1** (1975), 245–248.
[3] T. Jørgensen, A. Lascurain and T. Pignataro, *Translation extensions of the classical modular group*, Complex Variable **19** (1992), 205–209.
[4] S. L. Krushkal', B. N. Apanasov and N. A. Gusevskiĭ, Kleinian groups and uniformization in examples and problems, Trans. Math. Monographs **62**, Amer. Math. Soc., Providence, RI, 1986.
[5] C. Li and M. Oichi, *Jørgensen numbers and discrete groups*, in preparation.
[6] C. Li, M. Oichi and H. Sato, *Jørgensen groups of parabolic type I* (*Finite type*), RIMS Koukyuroku **1293**, Kyoto Univ. (2002), 65–77.
[7] C. Li, M. Oichi and H. Sato, *Jørgensen groups of parabolic type II* (Countably infinite case), to appear in Osaka J. Math. (2004).
[8] C. Li, M. Oichi and H. Sato, *Jørgensen groups of parabolic type III* (Uncountably infinite case), to appear in Kodai Math. J. (2005).
[9] B. Maskit, Kleinian Groups, Springer-Verlag, New York-Berlin-Heidelberg, 1987.
[10] B. Maskit, Some special 2-generator Kleinian groups, Proc. Amer. Math. Soc. **106** (1989), 175–186.
[11] J. R. Ratcliffe, Foundations of hyperbolic manifolds, Grad. Texts in Math. **149**, Springer-Verlag, New York 1994.
[12] H. Sato, *One-parameter families of extreme groups for Jørgensen's inequality*, Contemporary Math. **256** (The First Ahlfors - Bers Colloquium), edited by I. Kra and B. Maskit) (2000), 271–287.
[13] H. Sato, *Jørgensen groups and the Picard group*, Proc. of The Third ISAAC International Conference, Academic Scientific Publ. (2003), 149-158.
[14] H. Sato and R. Yamada, *Some extreme Kleinian groups for Jørgensen's inequality*, Rep. Fac. Sci. Shizuoka Univ. **27** (1993), 1–8.
[15] N. J. Wielenberg, The structure of certain subgroups of the Picard group, Math. Proc. Cambridge Philos. Soc. **84** (1978), 427–438.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

AUTOMORPHISMS OF THE 3-SPHERE THAT PRESERVE A GENUS TWO HEEGAARD SPLITTING

MARTIN SCHARLEMANN

Abstract. An updated proof of a 1933 theorem of Goeritz, exhibiting a finite set of generators for the group of automorphisms of S^3 that preserve a genus two Heegaard splitting. The group is analyzed via its action on a certain connected 2-complex.

1. Introduction

In 1933 Goeritz [Go] described a set of automorphisms of the standard unknotted genus two handlebody in S^3, each of which extends to all of S^3. He further observed that any such automorphism is a product of elements of this finite set. Stated somewhat differently, Goeritz identified a finite set of generators for the group $\mathcal{H}$, defined as isotopy classes of orientation-preserving homeomorphisms of the 3-sphere that leave a genus two Heegaard splitting invariant. Goeritz' theorem was generalized to Heegaard splittings of arbitrarily high genus by Powell [Po], but the proof contains a serious gap.[1] So a foundational question remains unresolved: Is the group of automorphisms of the standard genus g Heegaard splitting of S^3 finitely generated and, if so, what is a natural set of generators. The finite set of elements that Powell proposes as generators remains a very plausible set.

Since the gap in Powell's proof has escaped attention for 25 years, Goeritz' original theorem might itself be worth a second look. In addition, his argument is difficult for the modern reader to follow, is published in a fairly inaccessible journal and is a bit old-fashioned in its outlook. In view of the use that has been made of it in recent work on tunnel number one knots (cf [ST], [Sc]) it seems worthwhile to present an updated proof, in hopes also that it might be relevant to the open analogous problem for Heegaard splittings of higher genus.

The purpose of this note is to present such a proof, one influenced by the idea of thin position. One way to describe the outcome of this investigation is this: there is a natural 2-complex Γ (which deformation retracts to a graph) on which $\mathcal{H}$ acts transitively. One can write down an explicit finite presentation for the stabilizer $\mathcal{H}_P$ of a vertex $v_P \in \Gamma$ and observe that the stabilizer acts transitively on the edges of Γ incident to v_P. In particular, if we add to $\mathcal{H}_P$ any element δ

2000 *Mathematics Subject Classification*: 57M40.

Keywords and phrases: Heegaard splitting, reducing sphere, curve complex.

Research supported in part by an NSF grant.

[1] On p. 210, Case 2 the argument requires that, among the chambers into which $\phi^{-1}(s_k)$ divides the handle, there are two adjacent ones that each contain pieces of G^h_k. There is no apparent reason why this should be true.

of $\mathcal{H}$ that takes v_P to some adjacent vertex then the subgroup generated by $\mathcal{H}_P$ and δ is exactly the subgroup that preserves the component in which v_P lies. This in fact is all of $\mathcal{H}$, because it turns out that Γ is connected. The proof that Γ is connected can be viewed as the core argument in this paper.

2. The complex Γ and its vertex stabilizers

We outline the general setting, referring the reader to [Po, Section 1] for details. Let V denote the standard unknotted genus-two handlebody in S^3, with closed complement W also a genus two handlebody. Let $\mathcal{H}$ denote the group of orientation-preserving homeomorphisms of S^3 that preserve V. Regard two as equivalent if there is an isotopy from one to the other via isotopies that preserve V. Any orientation preserving homeomorphism of S^3 is isotopic to the identity, so an element $h : (S^3, V) \to (S^3, V)$ of $\mathcal{H}$ is isotopic, as a homeomorphism of S^3, to the identity. This gives an alternate view of $\mathcal{H}$: an element of $\mathcal{H}$ corresponds to an isotopy of S^3 from the identity to a homeomorphism that preserves V setwise.

For $T = \partial V = \partial W$, $S^3 = V \cup_T W$ is a genus two Heegaard splitting of S^3. In the language of Heegaard splittings, a *reducing sphere* $P \subset S^3$ is a sphere that intersects T transversally in a single essential circle and so intersects each handlebody in a single essential disk. Since P is separating in S^3, $P \cap T$ is a separating curve in T, which we will denote c. A straightforward innermost disk argument shows that P is determined up to isotopy rel T by the circle c.

Suppose Q is another reducing sphere, with the circles c and $Q \cap T$ isotoped to intersect transversally and minimally in T. Then the number of points of intersection $|P \cap T \cap Q|$ is denoted $P \cdot Q$. Clearly $P \cdot Q = 0$ if and only if P and Q are isotopic since the only separating essential curve in either punctured torus component of $T - c$ is boundary parallel. Since reducing spheres are separating, $P \cdot Q$ is always even. An elementary argument (see [ST, Lemma 2.5]) shows that $P \cdot Q \neq 2$ and in some sense characterizes (up to multiple half-Dehn twists about c) all spheres Q so that $P \cdot Q = 4$. See Figure 1

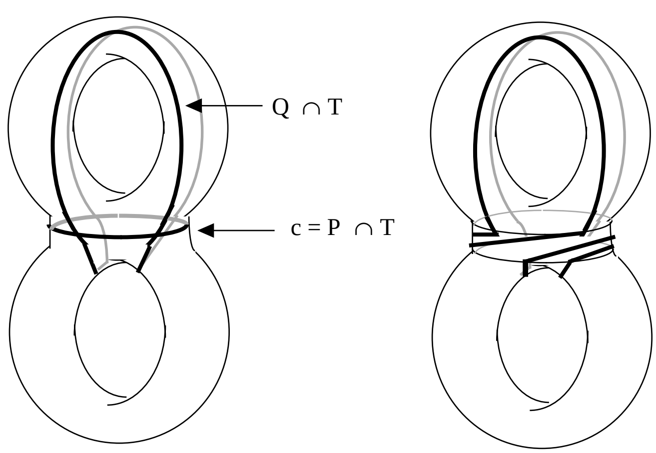


Figure 1

This suggests a useful simplicial complex: Let Γ be the complex in which each vertex represents an isotopy class of reducing spheres and a collection $P_0, \ldots, P_n$ of reducing spheres bounds an n-simplex if and only if $P_i \cdot P_j = 4$ for all $0 \leq i \neq j \leq n$. In fact it follows easily from the characterization in [ST, Lemma 2.5] that $n \leq 2$. Figure 2 illustrates a collection of three reducing spheres whose corresponding vertices in Γ span a 2-simplex in Γ. (An alternate view, in which V appears as *(pair of pants)* $\times I$, is shown in Figure 3.) Thus Γ is a 2-complex.

Each edge of Γ lies on a single 2-simplex. This is perhaps best seen in Figure 3: The curve $P \cap T$ is uniquely defined by the curves $Q \cap T$ and $R \cap T$ shown. (For example, if the curve $P \cap T$ is altered by Dehn twists around the outside boundary of the pair of pants, it becomes a curve that is non-trivial in $\pi_1(V)$, so it can't bound a disk in V.) So the 2-complex Γ deformation retracts naturally to a graph, in which each 2-simplex in Γ is replaced by the cone on its three vertices.

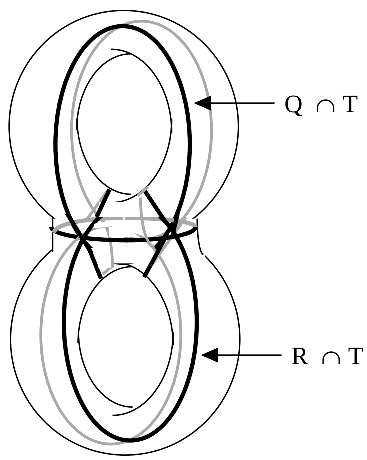


Figure 2

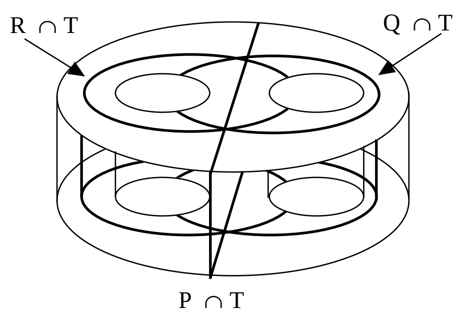


Figure 3

A reducing sphere P divides S^3 into two 3-balls $B_{\pm}$ and T intersects each 3-ball in a standard unknotted punctured torus, unique up to isotopy rel boundary.

It follows that for any two reducing spheres P and Q there is an orientation preserving homeomorphism of S^3, preserving V as a set, that carries P to Q. Thus $\mathcal{H}$ acts transitively on the vertices of Γ.

We now explicitly give a presentation of the group that stabilizes a vertex of Γ. As above, let P be a reducing sphere for the Heegaard splitting $S^3 = V \cup_T W$ and suppose $h : (S^3, V) \to (S^3, V)$ is an orientation preserving homeomorphism that leaves P invariant. That is, suppose h represents an element in $\mathcal{H}$ that stabilizes the vertex in Γ corresponding to P.

First assume that h preserves the orientation of P. Let $T_\pm = T \cap B_\pm$ denote the two punctured torus components of $T - P$; since h preserves orientation of both S^3 and P we have $h(T_+) = T_+$ and $h(T_-) = T_-$. Up to isotopy there is a unique non-separating curve $\mu_\pm \subset T_\pm$ that bounds a disk in V and a unique non-separating curve $\lambda_\pm$ that bounds a disk in W and we may choose these curves so that $\mu_\pm \cap \lambda_\pm$ is a single point. Hence, up to equivalence in $\mathcal{H}$, we may with little difficulty assume that each wedge of circles $\gamma_\pm = \mu_\pm \cup \lambda_\pm$ is mapped to itself by h and, on each $\gamma_\pm$, the homeomorphisms $h|\mu_\pm : \mu_\pm \to \mu_\pm$ and $h|\lambda_\pm : \lambda_\pm \to \lambda_\pm$ are either simultaneously orientation preserving (in which case we can take them both to be the identity) or simultaneously orientation reversing (in which case we can take them each to be reflections that preserve their intersection point). Having identified h on $\gamma_\pm$ we observe that $T - (\gamma_+ \cup \gamma_-)$ is an annulus A, and any end-preserving homeomorphism $A \to A$ is determined up to isotopy and Dehn twists around its core by $h|\partial A$. The upshot of this discussion is the following description:

Lemma (2.1). *: Let $\mathcal{H}_P^+$ be the subgroup of $\mathcal{H}$ represented by homeomorphisms that restrict to orientation-preserving homeomorphisms of P. Then*

$$\mathcal{H}_P^+ \cong Z_2 + Z$$

with generators given by the automorphisms α and β shown in Figure 4.

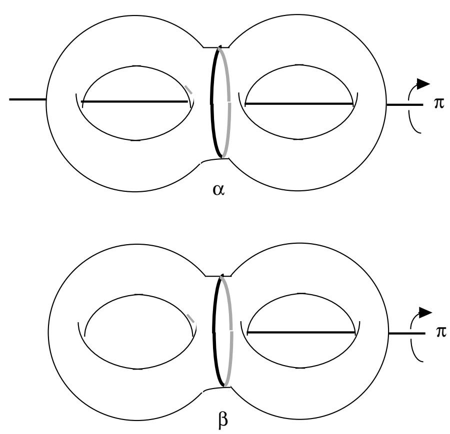


Figure 4

The situation is only slightly more complicated if we drop the requirement that $h|P$ be orientation preserving since the order two element $\gamma \in \mathcal{H}$ shown in Figure 5 preserves P but reverses its orientation.

LEMMA (2.2). *Let $\mathcal{H}_P$ be the subgroup of $\mathcal{H}$ represented by homeomorphisms that preserve P. Then $\mathcal{H}_P$ is an extension of $\mathcal{H}_P^+$ by Z_2, via the relations $\gamma\alpha\gamma = \alpha$ and $\gamma\beta\gamma = \alpha\beta$.*

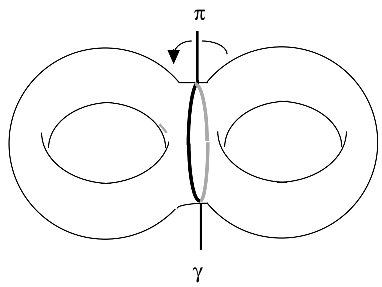


Figure 5

Finally, observe that if Q and Q' are reducing sphere so that $P \cdot Q = 4$ and $P \cdot Q' = 4$ then for some $n \in \mathbb{Z}$, either β^n or $\beta^n\gamma$ carries Q to Q'. (See discussion of Figure 1 above.) Interpreting this in terms of the action of $\mathcal{H}$ on the complex Γ we have:

COROLLARY (2.3). *Let $\mathcal{H}_P$ be the subgroup of $\mathcal{H}$ that stabilizes the vertex $v_P \in \Gamma$ corresponding to P. Then $\mathcal{H}_P$ is transitive on the edges of Γ incident to v_P.*

3. Intersection of reducing spheres

Suppose T_0 is an oriented punctured torus containing oriented simple closed curves μ, λ that intersect in a single point. For α an essential embedded arc in T_0 define the *slope* $\sigma(\alpha) \in \mathbb{Q} \cup \{\infty\}$ of the arc α as follows: Orient α and let $p = \alpha \cdot \mu$ and $q = \alpha \cdot \lambda$ be the algebraic intersection numbers of the corresponding homology classes. Then $\sigma(\alpha) = p/q$. Reversing the orientation of α has no effect on the slope, since it changes the sign of both p and q. An alternate description of the (unsigned) slope is this: minimize by an isotopy in T_0 the numbers $p = \alpha \cap \mu$ and $q = \alpha \cap \lambda$; then $|\sigma(\alpha)| = p/q$. If $\beta \subset T_0$ is another essential arc, with slope r/s define their distance $\Delta(\alpha, \beta) = |ps - qr| \in \mathbb{N}$. It is easy to see that if the arcs α and β are disjoint then $\Delta(\alpha, \beta) \leq 1$. Any embedded collection of arcs in T_+ constitutes at most three parallel families of arcs, with slopes of any pair of disjoint non-parallel arcs at a distance of one.

We now apply this terminology in the setting given above: P is a reducing sphere for $V \cup_T W$, the closed 3-ball components of $S^3 - P$ are $B_\pm$, the punctured tori $T \cap B_\pm$ are denoted $T_\pm$ and Q is a reducing sphere for $V \cup_T W$ that is not isotopic to P and has been isotoped so as to minimize $|P \cap Q \cap T| = P \cdot Q$. It will be convenient to imagine P as a level sphere of a standard height function on S^3, with B_+ above P and B_- below P. When we use the terms *above* and *below* in what follows, we will be refering to such a height function.

In each of $T_\pm$ there are closed non-separating curves $\mu_\pm, \lambda_\pm$ bounding respectively disks in V and disks in W and for each pair, $\mu_\pm \cap \lambda_\pm$ is a single point. We will consider the collection of arcs $Q \cap T_\pm$ and their slopes with respect to $\mu_\pm, \lambda_\pm$. Fix at the outset some orientations, e.g. orient T (hence $T_\pm$) as ∂V and orient $\mu_\pm, \lambda_\pm$ so that the algebraic intersection number $\mu_\pm \cdot \lambda_\pm = 1$. (The exact choice of orientations is not critical.)

Lemma (3.1). *There is some arc in either $Q \cap T_+$ or in $Q \cap T_-$ that is of slope ∞ and another such arc is of slope* 0.

Proof. An outermost disk cut off by the disk $P \cap V$ from the disk $Q \cap V$ is a meridian disk D of the solid torus $V \cap B_+$ or $V \cap B_-$. Then the arc $D \cap T$ must be of slope 0. A symmetric argument on the disks $P \cap W, Q \cap W$ gives an arc of slope ∞. □

Lemma (3.2). *Suppose that an arc α_+ of $Q \cap T_+$ has slope ∞ (resp.* 0*) and that there is an arc α_- of slope* 0 *(resp. ∞) in T_- that is disjoint from Q. Then there is a reducing sphere R so that $P \cdot R = 4$ and $R \cdot Q < P \cdot Q$.*

The same hypothesis, but with T_+ and T_- reversed, leads to the same conclusion.

Proof. Since α_- is merely required to be disjoint from Q, with no loss we may assume that the ends of $\alpha_\pm$ on the circle $c = P \cap T$ are disjoint. Say that the arcs $\alpha_\pm$ *cross* if the ends of α_+ and α_- alternate around c; that is, if the ends of α_+ lie on different arc components of $c - \alpha_-$.

Claim. Some pair of arcs that satisfy the hypotheses for $\alpha_\pm$ cross.

Proof of claim. Assume, on the contrary, that no such pair of arcs crosses. Then among arcs of $Q \cap T_\pm$ satisfying the conditions for $\alpha_\pm$ choose the pair whose ends are closest to each other on the circle c. The ends of $\alpha_\pm$ divide c into four arcs, one of them, denoted β_+, is bounded by the ends of α_+ and the other, denoted β_-, by the ends of α_-. Let $c_\pm = |Q \cap interior(\beta_\pm)|$.

$T_+ - \eta(\alpha_+)$ is an annulus A; denote the boundary component that contains $\beta_\pm$ by $\partial_\pm A$. Then $|\partial_+ A \cap Q| = c_+$ and $|\partial_- A \cap Q| \geq c_-$. (The inequality reflects the fact that Q may also intersect the two intervals $c - \beta_\pm$.) No arc of $Q \cap A$ can have both ends on $\partial_- A$, else it would have been parallel to α_+ in T_+, and yet closer to α_-. We conclude that $c_+ \geq c_-$. Arguing symmetrically on $T_- - \eta(\alpha_-)$, we obtain $c_- \geq c_+ + 2$, the extra 2 arising from the ends of α_+. The two inequalities conflict, a contradiction proving the claim.

Following the claim, we assume that $\alpha_\pm$ cross. Let $\rho \subset T$ be the circle obtained by banding the circle c to itself along the two arcs $\alpha_\pm \subset T_\pm$. It is a single circle because $\alpha_\pm$ cross. Moreover, it's easy to see that ρ is an essential circle in T (there are essential curves in T on both sides of ρ) and that ρ bounds disks both in V and W. So ρ is the intersection with T of a reducing sphere R. Moreover, $R \cdot P = |\rho \cap P| = 4$ and $R \cdot Q \leq |\rho \cap Q| = |c \cap Q| - 2 = P \cdot Q - 2$ since the ends of α_+ no longer count. □

Proposition (3.3). *There is a reducing sphere R so that $P \cdot R = 4$ and $R \cdot Q < P \cdot Q$.*

Proof. If there are two arcs of $(Q \cap T) - c$, one of slope 0 and one of slope ∞, one lying in T_+ and the other lying in T_-, the result follows from Lemma (3.2). Following Lemma (3.1) we know that there are arcs of slope both 0 and ∞. Thus we are done unless both these arcs lie in T_- say, and each arc of $Q \cap T_+$ has finite, non-zero slope. Moreover, if all arcs of $Q \cap T_+$ have slope 1 (or slope -1) then a curve of slope 0 in T_+ will be disjoint from $Q \cap T_+$ and again we would be done by Lemma (3.2). If σ, τ are slopes of arcs in $Q \cap T_+$, then, because $|\Delta(\sigma, \tau)| \leq 1$, the inequality $0 < |\sigma| < 1$ would imply that $|\tau| \leq 1$ and that σ and τ have the same sign. Finally, a curve that has slope σ, will have slope $1/\sigma$ if the roles of V and W are reversed. Following these considerations, we may as well restrict to the following case:

- Both slopes 0 and ∞ arise among the arcs of $Q \cap T_-$ and
- all arcs of $Q \cap T_+$ have slope σ with $0 < \sigma \leq 1$ and not all have slope 1.

Now consider a sphere $P^+ \subset B_+$ that intersects the solid torus $V \cap B_+$ in two meridian disks, and so intersects W in an annulus. Again isotope the curve $Q \cap T$ so that it intersects the two meridian circles $P^+ \cap T$ minimally. Any arc of $Q \cap T_+$ must intersect P^+, else the arc would be of slope 0. In particular, there is an essential non-separating disk $F \subset W$ so that $\partial F \subset T_+$ (i. e. ∂F is a longitude of the solid torus $V \cap B_+$) so that $F \cap P^+$ is a single spanning arc of the annulus $P^+ \cap W$ and so that the arc of $\partial F - P^+$ lying below P^+ (i. e. in the pair of pants component of $T_+ - P^+$ adjacent to c) is disjoint from Q. See Figure 6.

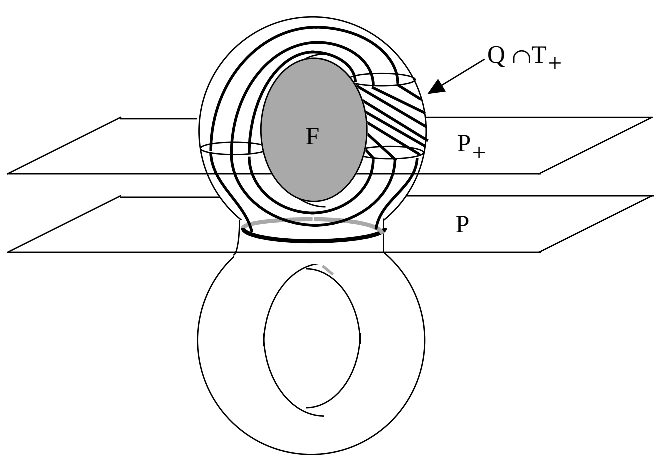


Figure 6

We now examine outermost disks cut off from the disk $Q \cap W$ by the annulus $P^+ \cap W$. Let E be any such disk. Let $V^{\pm}$ be the closed components of $V - P^+$, with V^+ the 1-handle lying above P^+ and V^- the solid torus lying below P^+.

Claim 1. The outermost arc $\epsilon = \partial E \cap P^+$ spans the annulus $P^+ \cap W$.

Proof of Claim 1. This is obvious if E lies above P^+, since all arcs of $Q \cap T$ above P^+ span the 1-handle V^+. If E lies below P^+ the argument is a bit more subtle. Note that V^- is a solid torus with two disks d_1, d_2 in ∂V^- attached to P^+. A simple counting argument (the d_i are parallel in V^+) shows that any arc of $Q \cap (\partial V^- - P^+)$ that has both ends on the same disk d_i is essential in the torus ∂V^-. So an outermost disk $D \subset V^-$ cut off from the disk $Q \cap V$ by the meridian disks d_i must be a meridian disk of the solid torus V^-, and so the arc $\partial D \cap T$ has both ends on d_1, say. The same counting argument shows that some essential arc in $Q \cap V^- \cap T$ must have both its ends on d_2 and so is a meridinal arc for V^- there as well. If the ends of ϵ were both on the same d_i, then $\partial E - \epsilon$ would be a longitudinal arc disjoint from the meridinal arc with ends at the other disk $d_j, j \neq i$. But a longtiudinal arc and a meridinal arc based at different points must necessarily intersect. Hence the ends of ϵ each lie on a different disk d_i, proving Claim 1.

Claim 2. All the outermost disks cut off from $Q \cap W$ by P^+ must lie on the same side of P^+.

Proof of claim 2. Suppose, on the contrary, that the outermost disks $E^\pm$ are cut off, with E^- lying in the component of $S^3 - P^+$ that lies below P^+ and E^+ lying in the component that lies above P^+. Following Claim 1, both arcs $\epsilon^\pm = E^\pm \cap P^+$ span the annulus $P^+ \cap W$.

Since the arc $E^- \cap T$ is disjoint from ∂F it follows from a simple innermost disk, outermost arc argument, that all of E^- can be made disjoint from F; in particular the spanning arcs ϵ^- and $F \cap P^+$ are disjoint. Since the spanning arc ϵ^+ is disjoint from the spanning arc ϵ^- which in turn is disjoint from the spanning arc $F \cap P^+$, ϵ^+ can be isotoped off of $F \cap P^+$ without moving ϵ^-. (See Figure 7.) Then again an innermost disk, outermost arc argument allows us to isotope all of E^+ off of F. Now consider any arc component γ of $(Q \cap T_+) - P^+$. If γ lies below P^+ then it is disjoint from ∂F, by construction; if γ lies above P^+ then since it is disjoint from E^+, it intersects ∂F at most once. In particular, any arc of $Q \cap T^+$ intersects a component of $P^+ \cap T_+$ at least as often as it intersects ∂F, hence its slope has absolute value ≥ 1. This contradicts the second property itemized above, and so proves claim 2).

Claim 3. All the outermost disks cut off from $Q \cap W$ by P^+ must lie above P^+.

Proof of claim 3. Following claim 2) the alternative would be that they all lie below (in B_-). We show how this leads to a contradiction. Consider the disk $Q \cap W$ and how it is cut up by the annulus $P^+ \cap W$. A standard innermost disk argument ensures that all closed curves of intersection can be removed. There is at least one (disk) component E_0 of $(Q \cap W) - P^+$ that is "second outermost", i. e. it is adjacent to some $n \geq 2$ other components of $(Q \cap W) - P^+$, all but at most one of them outermost. See Figure 8. Since E^0 is adjacent to an outermost component, all of which we are assuming lie below P^+, E_0 must lie above P^+. By Claim 1), all the outermost arcs of intersection of P^+ with the disk $Q \cap W$ must span the annulus $W \cap P^+$, so it follows that each of the n arc components of $\partial E_0 \cap T$ spans the 1-handle V^+. In particular, the union of the disk E_0 with the

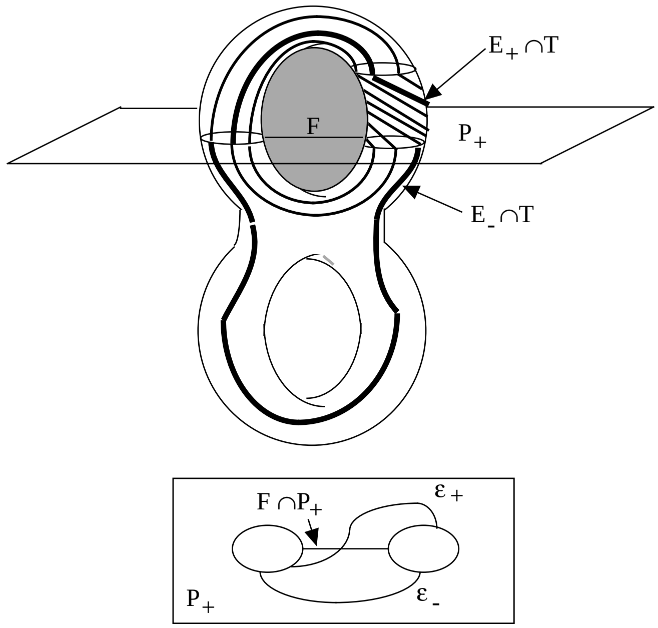


Figure 7

punctured solid torus $P^+ \cup V^+$ is the spine of a Lens space in S^3, a contradiction proving Claim 3).

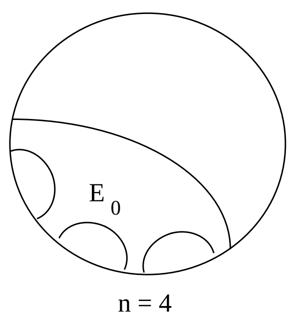


Figure 8

Following Claim 3), consider a sphere P^s that passes through the saddle point of T_+ that lies below P^+. We can assume (see Figure 9) that P^s intersects Q transversally and that every arc of $Q \cap T^+$ that lies above P^s spans the 1-handle V^+. According to claim 3) applied to a plane just slightly higher than P^s, P^s (and so also a plane P^{s-} lying just below P^s) cuts off a disk E^+ from $Q \cap W$ that lies above the plane. Let $\alpha \subset (P^{s-} \cap T)$ be an arc parallel in $P^{s-} \cap W$ to the arc $E^+ \cap P^{s-}$, so the union λ of α and the arc $\gamma = E^+ \cap T$ is a longitude lying above P^{s-} (indeed λ is a meridian of W). It's easy to isotope the ends of

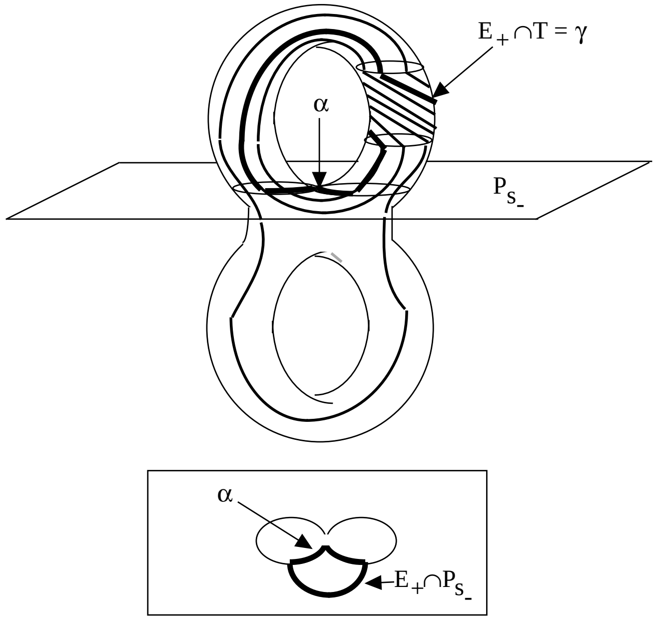


Figure 9

γ closer together in α until no arc of $(Q \cap T) - P^{s_-}$ lying above P^{s_-} has more than one end on λ. It then follows just as in the proof of Claim 2) that any arc component of $Q \cap T_+$ intersects a meridian of V^+ at least as often as it intersects λ and so has slope ≥ 1, a contradiction that completes the proof. $\square$

Corollary (3.4). *The* 2*-complex* Γ *is connected.*

Proof. Let w be a fixed vertex of Γ, with associated reducing sphere Q. Let Γ_0 be any component of Γ. Choose a reducing sphere P among those represented by vertices in Γ_0 so that $P \cdot Q$ is minimized. Unless $P = Q$, Proposition (3.3) provides a reducing sphere R which is represented by a vertex in Γ_0 (indeed one adjacent to the vertex representing P) but for which $R \cdot Q < P \cdot Q$. From the contradiction we conclude then that indeed $P = Q$, so $w \in \Gamma_0$. $\square$

Corollary (3.4) is essentially [ST, Proposition 2.6]. There we used Goeritz' theorem to prove the proposition; here we have proven the proposition from first principles and now observe that it proves Goeritz' theorem.

4. A finite set of generators

Theorem (4.1). *Suppose* $\delta \in \mathcal{H}$ *is any element with the property that* $P \cdot \delta(P) = 4$. *Then the group* $\mathcal{H}$ *is generated by* $\alpha, \beta, \gamma, \delta$.

Proof. Choose any $h \in \mathcal{H}$ and let $Q = h(P)$. If $Q = P$ then by Lemma (2.2), h is in the group generated by α, β and γ. Otherwise, following Corollary (3.4), there is a sequence of reducing spheres $P = P_0, P_1, \ldots, P_n = Q$ so that $P_{i-1} \cdot P_i = 4, i = 1, \ldots, n$. The proof will be by induction on the length n of this sequence – the case $n = 1$ follows from Corollary (2.3). In particular, there is a word ω in the generators $\alpha, \beta, \gamma, \delta$ so that $\omega(P_1) = P$. Apply ω to every sphere in the shorter sequence $P_1, \ldots, P_n = Q$ and obtain a sequence $P = \omega(P_1), \omega(P_2), \ldots, \omega(Q) = \omega(h(P))$. Then by inductive hypothesis, ωh is in the group generated by $\alpha, \beta, \gamma, \delta$, hence so is h. $\square$

There are several natural choices for δ. For example, if we think of V as a ball with two 1-handles attached, the two 1-handles separated by the reducing sphere P, then a slide of an end of one of the 1-handles over the other around a longitudinal curve will suffice for δ. This is the genus two version of Powell's move D_θ ([Po, Figure 4]). Another possibility is to choose an order two element for δ, an element that is conjugate in $\mathcal{H}$ to γ: note from Figure 2 that $Q \cdot \gamma(Q) = 4$.

A bit more imaginative is the automorphism shown in Figure 10 which is of order three and corresponds to rotating one of the two-simplices of Γ around its center. The figure is meant to evoke a more symmetric version of Figure 2: it depicts a thrice punctured sphere with three essential arcs, each pair intersecting in two points. Thicken the figure (i.e. cross with an interval). Then the thrice punctured sphere becomes a genus two handlebody V and each arc becomes a disk. Each disk is the intersection with V of a reducing sphere, and the three reducing spheres are represented by the corners of a single two-simplex σ in Γ. Rotation of the figure by $2\pi/3$ along the axis shown cyclically permutes the three arcs, and so cyclically permutes the three reducing spheres. Hence it also rotates the corresponding 2-simplex σ in Γ.

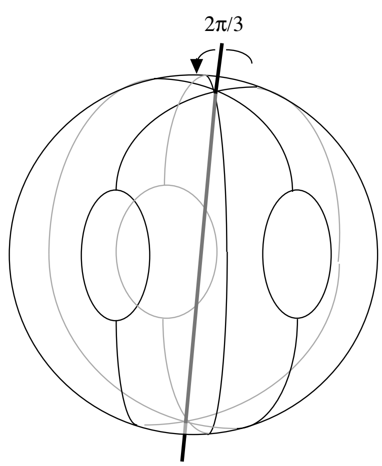


Figure 10

Received July 07, 2003

Final version received May 06, 2004

Mathematics Department
University of California
Santa Barbara, CA 93106
USA
mgscharl@math.ucsb.edu

References

[Go] L. Goeritz, *Die Abbildungen der Brezelfläche und der Volbrezel vom Gesschlect 2*, Abh. Math. Sem. Univ. Hamburg **9** (1933), 244–259.

[Po] J. Powell, *Homeomorphisms of S^3 leaving a Heegaard surface invariant*, Trans. Amer. Math. Soc. **257** (1980), 193–216.

[Sc] M. Scharlemann, *There are no unexpected tunnel number one knots of genus one*, Trans. Amer. Math. Soc. **356** (2004), 1385–1442

[ST] M. Scharlemann, A. Thompson, *Unknotting tunnels and Seifert surfaces*, Proc. London Math. Soc. **87** (2003), 523–544.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

ON INFINITELY PERIODIC KNOTS

MAT TIMM AND OLLIE NANYES

ABSTRACT. Fox asked the following question: "if for every $g \geq 2$, there is a periodic transformation T_g of period g of the 3-sphere S^3 such that $T_g(K) = K$, what kind of knot can K be?" Flapan showed that if K were smooth and if the fixed point set was S^1 and disjoint from K, that K had to be the unknot. In this paper we show that there are non-trivial wild knots of this type that admit periods of all orders, and that all such knots must have an uncountable number of wild points.

1. Introduction

In [3], Fox asked the following question (Question 6, [3]) "if for every $g \geq 2$, there is a periodic transformation T_g of period g of the 3-sphere S^3 ($T_g\colon S^3 \to S^3$) such that $T_g(K) = K$ (but not necessarily $T_g(p) = p$ for any $p \in K$), what kind of knot can K be?". Fox then posed Question 7, which asked, "Given a non-trivial knot K, which periods g does it permit?" He observed that that the answer may depend on the fixed point set F of T_g, whether T is orientation reversing or preserving, and on $K \cap F$.

A knot K is said to be infinitely periodic if, for each $g \in N$, there exists a periodic homeomorphism $T_g : S^3 \to S^3$ such that $T_g(K) = K$. The fixed point set of a homeomorphism $T : S^3 \to S^3$ will be denoted by $F = F(T) = \{p \in S^3 : T(p) = p\}$. Note that for any such T, $F(T)$ is homeomorphic to one of $\emptyset$, S^0, S^1, S^2 or S^3 [3]. We will adopt Flapan's notation [2]: a knot K will be said to be (a, b)-periodic if there is a T with $F(T)$ homeomorphic to a such that $F(T) \cap K$ is homeomorphic to b. K will be said to be infinitely (a, b) periodic if for each $g \geq 2$ there is a T of type (a, b) with period g. In our paper, we will be interested in infinitely $(S^1, \emptyset)$ periodic knots.

In [7], Seifert showed that any smooth torus knot is infinitely periodic with $F = \emptyset$. In [2], Flapan showed that the torus knots are the only smooth, infinitely periodic knots K with $F = \emptyset$ and that the only smooth $(S^1, \emptyset)$ periodic knot which admits infinitely many periods is the unknot.

Our main result shows that the smooth condition on K is essential to Flapan's argument; we will show how to construct an infinitely $(S^1, \emptyset)$ periodic wild knot (Theorem (2.1)). These knots are contained in certain solid tori, and their complements in these solid tori are connected open 3-manifolds with one boundary component that have non-trivial cyclic self-covers of all orders. In section 3 we show that there are an uncountable number of inequivalent $(S^1, \emptyset)$ periodic wild

2000 *Mathematics Subject Classification*: 57M25, 57M30, 57S25.
Keywords and phrases: knots, peroidic knots, wild knots, group actions on S^3.

knots. We will also show that all such knots must have an uncountable number of wild points (Theorem (4.1).

The method used to construct the knots of interest in fact illustrates a general method that can be used to construct many spaces with cyclic self-covers of all orders. For related work, see [1].

2. Construction of an infinitely periodic knot

Henceforth, when we say that a knot K is infinitely periodic we mean infinitely $(S^1, \emptyset)$ periodic. Note that we do not require that each T_g have the same fixed point set F for all g.

We start our construction of an infinitely periodic knot K as follows: view S^3 as $R^3 \cup \{\infty\}$. We will use cylindrical coordinates for $R^3 = \{(r,\theta,z) : r \geq 0, 0 \leq \theta < 2\pi, z \in R\}$. Consider a solid torus $V \subset R^3$, which we will parametrize as $S^1 \times D^2 = \{(\theta,\rho,\phi) : 0 \leq \theta < 2\pi, 0 \leq \rho \leq 2, 0 \leq \phi < 2\pi\}$. (ρ,ϕ) represents a polar coordinate system of a meridional disk of V; the units used for ρ are not the standard distance units in R^3. For example, $(0,2,\pi) \in V$ has R^3 coordinates $(\frac{3}{2},0,0)$ and $(\frac{\pi}{2},1,\frac{3\pi}{4}) \in V$ has R^3 coordinates $(\frac{2-\sqrt{2}}{2},\frac{\pi}{2},\frac{\sqrt{2}}{2})$. The centerline of V will be identified with the unit circle in the $z=0$ plane of R^3, with θ being used for both $S^1 \subset V$ and R^3. $V \cap \{z=0\}$ will be identified with the annulus $\{(r,\theta,0) : \frac{1}{2} \leq r \leq \frac{3}{2}, 0 \leq \theta < 2\pi\}$.

Let C denote the standard "middle thirds" Cantor set in the interval $[0,1]$, and C^* its image under $\pi : [0,2] \to V$, $\pi(x) = \{(\pi x, 0, -)\}$. Note that $\theta = \pi x$. Let D denote the set of intervals that are deleted from $[0,1]$ to form C, together with $(1,2)$. That is,

$$D = \{(1,2)\} \cup \{(\frac{1}{3},\frac{2}{3}),(\frac{1}{9},\frac{2}{9}),(\frac{7}{9},\frac{8}{9}),(\frac{1}{27},\frac{2}{27}),(\frac{7}{27},\frac{8}{27}),(\frac{19}{27},\frac{20}{27}),(\frac{25}{27},\frac{26}{27}).....\}.$$

Let $D_{i,j} \in D$ be defined to be the j'th deleted interval of length $\frac{1}{3^i}$. We let $D_{0,0}$ denote the open interval $(1,2)$. Let $D^*_{i,j}$ denote the image of $D_{i,j}$ under π and $D^*_{0,0}$ denote the semi-circular segment $((\pi,0,-),(2\pi,0,-))$ and D^* denote the collection of the $D^*_{i,j}$. For each $D_{i,j} = (\frac{p}{3^i},\frac{p+1}{3^i})$ ($i \geq 0, j \geq 0$, where $j \leq 2^{i-1}$ for $i \geq 1$), associate an open "curvilinear double cone" $A_{i,j} \subset V$ (see Figure 1), where

$A_{i,j} = \{(\pi x,\rho,\phi) | \frac{p}{3^i} < x < \frac{p+1}{3^i}, 0 \leq \rho < a(x)$ and $a(x) = \frac{1}{2}(\min\{x - \frac{p}{3^i}, \frac{p+1}{3^i} - x\})\}$.

We will let $C_{i,j}$ denote the open "curvilinear double cone" $C_{i,j} = \{(\pi x,\rho,\phi) | \frac{p}{3^i} < x < \frac{p+1}{3^i}, 0 \leq \rho < b(x)$ and $b(x) = \frac{3}{2}(\min\{x - \frac{p}{3^i}, \frac{p+1}{3^i} - x\})\}$. We are using this "nested cone" construction to facilitate our contstruction of the periodic homeomorphisms T; in particular the region between the two cones will serve as a "buffer" for the expansion/contraction part of each T.

Consider a ball pair (B,k), where B is a 3-ball and k is a properly embedded arc (possibly wild at its endpoints) with $k \cap \partial B = \{c\} \cup \{d\}$. Assume that k is knotted in B (rel (c,d)). That is, k together with an arc in ∂B is a non-trivial knot in $B \subset S^3$. Choose a specific knot type for k (e.g., k could be the trefoil as indicated in Figure 1, or the reader's favorite knot). We call k a *pattern arc* for K. For each $A_{i,j}$ there is a smooth embedding $\psi_{i,j} : (B,k) \to V$ where $\psi_{i,j}(c) = \pi(\frac{p}{3^i})$, $\psi_{i,j}(d) = \pi(\frac{p+1}{3^i})$, $\psi_{i,j}(\mathrm{int}(B,k)) = A_{i,j}$ and, for all i,j, all of the

$\psi_{i,j}(B,k) \subset A_{i,j}$ are homeomorphic by orientation preserving homeomorphisms. That is, we glue in the "same" k into each $A_{i,j}$

Let K be the simple closed curve which consists of $C^* \cup \{\cup_{i,j}\{\psi_{i,j}(k)\}, (i \geq 0, j \geq 0$, where $j \leq 2^{i-1}$for $i \geq 1)\}$.

That is, K is the knot formed by replacing the deleted intervals with copies of the arc k, where the copies of k are properly embedded in the $C_{i,j}$. Clearly, K is a wild, non-trivial knot. Figure 1 shows some of the stages of K.

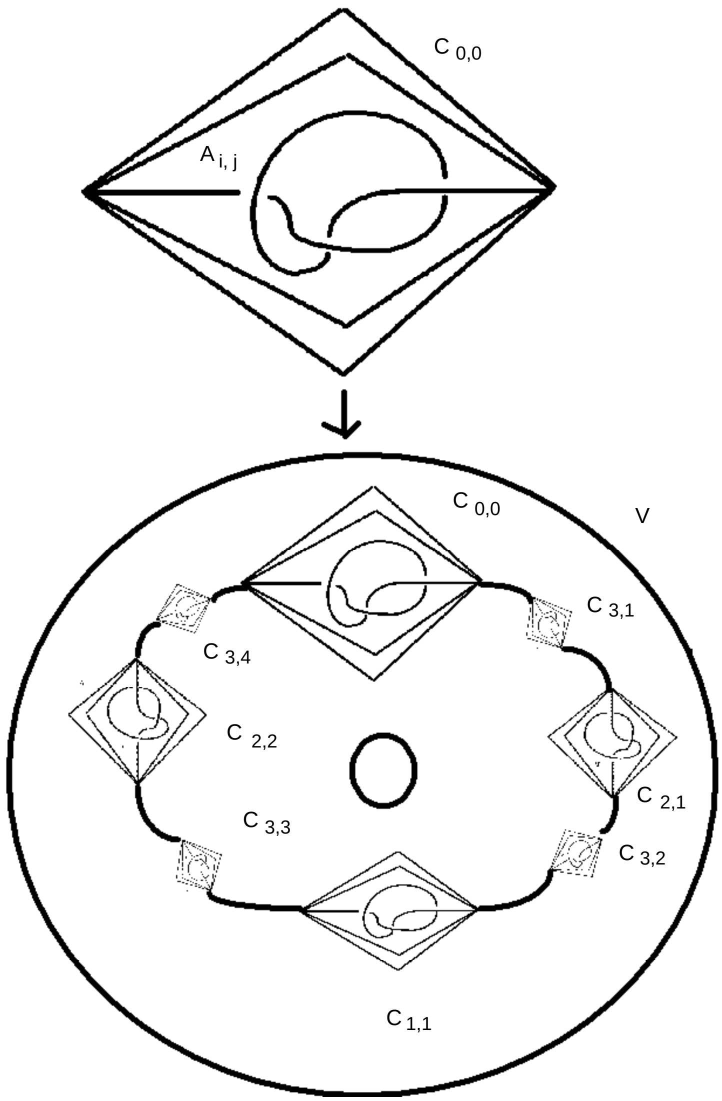


Figure 1

THEOREM (2.1). *K is infinitely $(S^1, \emptyset)$-periodic.*

Proof. First note that for each $n \in \{1, 2, 3, \dots\}$ there exists a homeomorphism f_n of the centerline circle $S^1 \subset V$ (recall $S^1 = \{(\theta, 0, 0) \in V, \theta \in [0, 2\pi)\}$) of period n that maps D^* to itself. We describe this map as follows:

$$f_1(\theta) = \theta$$

$$f_2(\theta) = \begin{cases} \theta + \frac{2\pi}{3}, & \theta \in [0, \frac{\pi}{3}] \\ 3\theta (\operatorname{mod} 2\pi), & \theta \in (\frac{\pi}{3}, \frac{2\pi}{3}] \\ (\theta + \frac{4\pi}{3})(\operatorname{mod} 2\pi), & \theta \in (\frac{2\pi}{3}, \pi] \\ \frac{1}{3}\theta, & \theta \in (\pi, 2\pi) \end{cases}$$

$$f_3(\theta) = \begin{cases} \theta + \frac{2\pi}{3^2}, & \theta \in [0, \frac{\pi}{3^2}] \\ 3\theta (\operatorname{mod} 2\pi), & \theta \in (\frac{\pi}{3^2}, \frac{2\pi}{3}] \\ \frac{(\theta + \frac{4\pi}{3})(\operatorname{mod} 2\pi)}{3}, & \theta \in (\frac{2\pi}{3}, \pi] \\ \frac{1}{3^2}\theta, & \theta \in (\pi, 2\pi) \end{cases}$$

$$f_n(\theta) = \begin{cases} \theta + \frac{2\pi}{3^{(n-1)}}, & \theta \in [0, \frac{\pi}{3^{(n-1)}}] \\ 3\theta (\operatorname{mod} 2\pi), & \theta \in (\frac{\pi}{3^{(n-1)}}, \frac{2\pi}{3}] \\ \frac{(\theta + \frac{4\pi}{3})(\operatorname{mod} 2\pi)}{3^{n-2}}, & \theta \in (\frac{2\pi}{3}, \pi] \\ \frac{1}{3^{n-1}}\theta, & \theta \in (\pi, 2\pi) \end{cases}$$

It will be helpful to define a function

$$\kappa : [0, 2\pi) \times Z^+ \to \{1, 3, 3^{-1}, 3^{-2}, ...3^{-m}, ..\}$$

by

$$\kappa(\theta, n) = \begin{cases} 1, & \theta \in [0, \frac{\pi}{3^{(n-1)}}] \\ 3, & \theta \in (\frac{\pi}{3^{(n-1)}}, \frac{2\pi}{3}] \\ \frac{1}{3^{n-2}}, & \theta \in (\frac{2\pi}{3}, \pi] \\ \frac{1}{3^{n-1}}, & \theta \in (\pi, 2\pi) \end{cases} \qquad (n > 1)$$

$\kappa(\theta, n) = 1$ for $n = 1$.

Next, we can define a homeomorphism $g_{(a,b,\kappa)}$ of the interval $[0, 2]$ (where $0 < a < b < 2, \kappa a < b$) which takes $[0, a]$ to $[0, \kappa a]$ and is the identity on $[b, 2]$ as follows:

$$g_{(a,b,\kappa)}(\rho) = \begin{cases} \kappa\rho, & 0 \le \rho \le a \\ \rho(\frac{(1-\kappa)\rho + \kappa b - a}{b-a}), & a < \rho < b \\ \rho, & b \le \rho \le 2\,. \end{cases}$$

Now, we define a homeomorphism T_{*n} of period n on V which maps $D^* \subset S^1 = \{(\theta, 0, -)\}$ to itself:

$$T_{*n}(\theta,\, \rho, \phi) = \begin{cases} (f_n(\theta), g_{(a(x), b(x), \kappa(\theta, n))}(\rho), \phi); & \theta \in D^* \\ (f_n(\theta), \rho, \phi); & \theta \notin D^* \end{cases} \quad \text{(recall: } \theta = \pi x)$$

The effect of T_{*n} for $n \ge 2$ is to perform an appropriate expansion or contraction in the x coordinate, and expansion or squeezing in the ρ coordinate. Note that for $(\theta, \rho, \phi) \in \partial V$, $T_{*n}(\theta, \rho, \phi) = (f_n(\theta),\, \rho, \phi)$. Hence it is easy to extend

T_{*n} to a homeomorphism T_n of R^3which has fixed point set $F = \{(0, -, z)|z \in R\}$ and has period n where $T_n|_V = T_{*n}$. Note that $T_n(K) = K$ setwise. Theorem (2.1) is proved. □

3. Construction of an uncountable number of mutually inequivalent $(S^1, \emptyset)$-periodic knots

In [5], McPherson shows how to construct an uncountable number of mutually inequivalent Fox-Artin arcs which have one wild endpoint of penetration index three. We will use these arcs as our pattern arc k to demonstrate that there are an uncountable number of mutually inequivalent $(S^1, \emptyset)$-periodic knots.

Let k be an embedded arc in S^3 with endpoints p, q. Assume that k is tame at all of its points except for possibly p. Let $E_1, E_2, \ldots$ be a system of tame closed 3-balls where $\cap_{i=1}^{\infty} E_i = p$ such that for all $i \geq 1, E_{i+1} \subset \text{int}(E_i)$ and $k \cap (E_i - \text{int}(E_{i+1}))$ consists of exactly three arcs: α_i which runs between ∂E_i and $\partial E_{i+1}, \beta_i$ which runs between two points of ∂E_i, and γ_i which runs between two points of ∂E_{i+1}. If β_i and γ_i are unsplittable in each $E_i - \text{int}(E_{i+1})$ (in the sense that if one turns β_i and γ_i into closed loops $\overline{\beta}_i, \overline{\gamma}_i$ by adding arcs along ∂E_i and ∂E_{i+1} respectively then $\overline{\beta}_i \cup \overline{\gamma}_i$ is an unsplittable link in S^3) then k is said to be a *Fox-Artin arc of penetration index 3* (the penetration index comes from the three arcs in each $E_i - \text{int}(E_{i+1})$). Figure 2 shows an example of a Fox-Artin arc. In [4] McPherson shows that all Fox-Artin arcs are wild (and therefore non-trivial) and in [5] he shows that there are an uncountable number of Fox-Artin arcs of penetration index 3 which have inequivalent "local types" at wild point p. It follows that there are an uncountable number of mutually inequivalent Fox-Artin arcs.

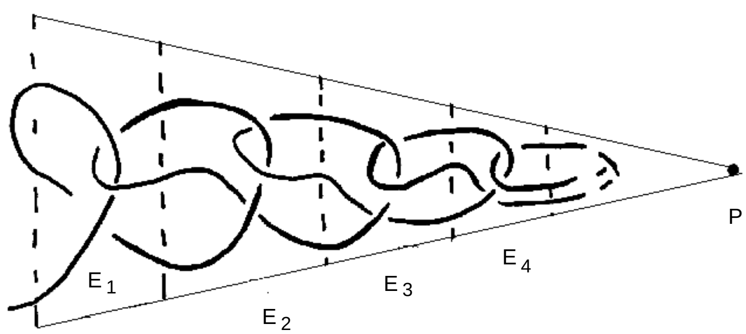


Figure 2.

THEOREM (3.1). *Let K_1 and K_2 be two knots which are constructed in the manner of Section 2, with pattern arcs k_1 and k_2 respectively. If K_1 is equivalent to K_2 then k_1 is equivalent to k_2.*

THEOREM (3.2). *There exists an uncountable number of mutually inequivalent $(S^1, \emptyset)$-periodic knots.*

Proof of Theorem (3.2). Follows directly from Theorem (3.1). □

Proof of Theorem (3.1). First some notation (refer to Section 2): C_i^* refers to the image of the Cantor set in knot K_i, $k_{m,n}^i$ refers to $\psi_{m,n}(k_i)$ in K_i, p^i is the "wild endpoint" of k_i, q^i the tame endpoint of k_i, $p_{m,n}^i$ is $\psi_{m,n}(p_i)$ and $q_{m,n}^i$ is $\psi_{m,n}(q_i)$. Let $h : S^3 \to S^3$ be a homeomorphism such that $h(K_1) = K_2$. We need a technical lemma.

Lemma (3.3). *For all $m, n \in \{0, 1, 2, \ldots\}$, $h(k_{m,n}^1) = k_{r,s}^2$ for some $r, s \in \{0, 1, \ldots\}$.*

Proof of the lemma. For each $q_{m,n}^1$ there exists a disk $Q_{m,n}$ with a product neighborhood $[-1, 1] \times Q_{m,n}$ such that $(\{0\} \times Q_{m,n}) \cap K_1 = q_{m,n}^1$ and $([0, 1] \times Q_{m,n}) \cap k_{m,n}^1$ consists of tame arcs (tamely embedded in $([0, 1] \times Q_{m,n})$), one of which has $q_{m,n}^1$ as an endpoint. We say that these $q_{m,n}^1$ are *tame from one side*. Similarly, for each $p_{m,n}^1$ there exists a disk $P_{m,n}$ with a product neighborhood $[-1, 1] \times P_{m,n}$ such that $(\{0\} \times P_{m,n}) \cap K_1 = p_{m,n}^1$ and $([-1, 0) \times P_{m,n}) \cap k_{m,n}^1$ consists of tame open arcs (tamely embedded in $([-1, 0) \times P_{m,n})$), one of which has $p_{m,n}^1$ as an endpoint (and, of course, is wild when $p_{m,n}^1$ is added in). We say that these $p_{m,n}^1$ are *almost tame from one side*.

If $x \in K^1$ is (almost) tame from one side, then $h(x)$ is (almost) tame from one side. If $y \in C_2^*$ then every open neighborhood contains wild points of K^2 on both sides of y. Therefore $y \neq h(x)$, which implies that y cannot be the image of any point of any $k_{m,n}^1$ because all of the points of $k_{m,n}^1$ are either tame, almost tame from one side, or tame from one side. Therefore $h(q_{m,n}^1) = q_{r,s}^2$ and $h(p_{m,n}^1) = p_{u,v}^2$. If $r \neq u$ or $s \neq v$, any subarc of K^2 running from $q_{r,s}^2$ to $p_{u,v}^2$ must pass through an infinite number of wild points. Therefore $(r, s) = (u, v)$. Hence h takes $k_{m,n}^1$ to $k_{u,v}^2$ which implies that k^1 and k^2 are equivalent. Lemma (3.3) and Theorem (3.1) are proved. □

4. Characterization of infinitely $(S^1, \emptyset)$-periodic knots

We now give a necessary condition for a knot K to be infinitely $(S^1, \emptyset)$-periodic.

Theorem (4.1). *If K is a non-trivial infinitely $(S^1, \emptyset)$-periodic knot, then K has an uncountable number of wild points.*

The proof of Theorem (4.1) will follow after some lemmas and propositions. First, we establish some notation. If A is a set, A' will denote the limit points of A. A point $x \in K$ is said to be tame if there exists a closed p.l. 3-ball B such that $x \in \text{int}(B)$ and $(B,\, B \cap K)$ is a standard ball pair. If x is not a tame point of K, then x is called wild. We denote the set of wild points of K by W. Note that Flapan's work implies that, for the infinitely periodic knots in which we are interested, $W \neq \emptyset$. Also note that W is a compact set in the standard subspace topology. We can start by assuming that K has tame points, else the theorem follows trivially.

For each $p \in N$, let $T_p : S^3 \to S^3$ be a given fixed periodic homeomorphism of period p acting freely in K. Let $S = \{T_p : p \in N\}$. By T_p^k we mean the composition of k copies of T_p. Of course, T_p^p is the identity map. We let

G be the group generated by S. Note that it consists of all finite compositions of the homeomorphisms T_p^k. We denote the orbit of $x \in K$ by $O(x) = \{y |\, y = T(x), T \in G\}$.

PROPOSITION (4.2). *For all $x \in K$, $O(x)$ is infinite.*

Proof. Suppose there is an $x \in K$ with $O(x)$ a finite set. Say, $O(x) = \{x = x_1, \ldots, x_k\}$. Let $p = k! + 1$. Let $T_p \in S$ be the given homeomorphism of period p. We see that T_p permutes the points of $O(x)$ without fixed points since $F \cap K = \emptyset$. That is, the restriction $(T_p|O(x))$ can be thought of as a permutation $\sigma \in \mathrm{Sym}(k)$, the group of permutations of k symbols. So the order of σ divides $k!$. Thus $\sigma^{k!} = (T_p|O(x))^{k!} = T_p^{k!}|O(x)$, which is the identity in $O(x)$. But then $T_p^{k!}$ would be the identity, which contradicts the fact that the order of T_p is $k! + 1$. The proposition is proved. □

Since homeomorphisms take wild points to wild points, if $x \in W$, then $O(x) \subset W$. Thus we have now established that the set of wild points is at least countably infinite. Suppose $x \in W$. Since W is compact and $O(x) \subset W$ is infinite, $O(x)' \subset W$ is not empty. We will establish that we can assume, with no loss of generality, that $O(x)$ contains none of its limit points.

LEMMA (4.3). *If $x \in W$ and if $O(x) \cap (O(x))' \neq \emptyset$, then W is uncountable.*

Proof. Suppose there exists $y \in O(x) \cap (O(x))'$. Then there is some $T \in G$ such that $y = T(x)$. Suppose U is open and $x \in U$. Then the open set $T(U)$ contains an infinite number of points of $O(x)$. Therefore, since T is a homeomorphism, U also contains an infinite number of points of $O(x)$. Hence $O(x) \subset O(x)'$. Hence $\overline{O(x)} \subset O(x)' \subset (\overline{O(x)})'$. That is, all points of the closure of $O(x)$ are limit points of $\overline{O(x)}$. Since $\overline{O(x)}$ is compact, $\overline{O(x)}$ is uncountable (see, e.g., Theorem 6.5, page 176 of [6]). But $\overline{O(x)} \subset W$, therefore W is uncountable. The lemma is proved. □

We need one more lemma:

LEMMA (4.4). *If $y_a \in O(y)'$, then $O(y_a)' \subset O(y)'$.*

Proof. It suffices to show that $O(y_a) \subset O(y)'$. Let $x \in O(y_a)$ and U be an open set which contains x. Then $x = T(y_a)$ for some $T \in G$. $T^{-1}(U)$ is open and contains y_a. Hence $T^{-1}(U)$ contains an infinite number of points of $O(y)$. Therefore U contains an infinite number of points of $O(y)$. The lemma is proved. □

Proof of Theorem (4.1). We will prove Theorem (4.1) by showing that no one to one map from the integers to W can be onto. From Lemma (4.3), we will assume that $O(x)$ contains none of its limit points. Let $f : N \to W$ be a one to one map and let $f(n) = y_n$. We will use induction to show the following: given subsets $V_1, V_2, \ldots, V_k$ which are open in K, have tame frontier (frontier in K) and contain $y_1, y_2, \ldots, y_k$, respectively, where for all $1 \leq i < j \leq k$, either $\overline{V}_i \cap \overline{V_j} = \emptyset$ or $V_i = V_j$, we can find another open set V_{k+1} containing y_{k+1} such that $V_{k+1} = V_j$ for some $1 \leq j \leq k$ or $\overline{V_{k+1}} \cap \overline{V}_j = \emptyset$ for all $1 \leq j \leq k$. Furthermore $K - (V_1 \cup V_2 \cup \ldots V_k \cup V_{k+1})$ contains an infinite number of wild points, namely $O(y_p)'$, for some $p \leq k + 1$.

Consider y_1. By Lemma (4.3), there is an open set $V_1 \subset K$ which separates y_1 from all other points of $O(y_1)$. It follows that V_1 contains no limit points of $O(y_1)$. We can assume that $O(y_1)'$ is countable, since, if $O(y_1)'$ were uncountable, $O(y_1) \subset W$ and W is closed; it follows that $O(y_1)'$ would be an uncountable subset of W, which would prove the theorem. Note that we can assume with no loss of generality that $\mathrm{Fr}(V_1) \subset K$ is tame. For, if $\mathrm{Fr}(V_1)$ were wild, we could attempt to find a smaller open interval (open in K) which contains y_1 whose endpoints are tame; if such an interval cannot be found then W must be uncountable. Let $M_1 = K - V_1$. M_1 is a compact set which contains $O(y_1)'$.

Now we proceed by induction. Assume by hypothesis of induction that we have open sets $V_1, V_2, \ldots, V_k$ containing $y_1, y_2, \ldots, y_k$ respectively, and for all $1 \le i < j \le k$, either $\overline{V}_i \cap \overline{V}_j = \emptyset$ or $V_i = V_j$ and $\mathrm{Fr}(V_i)$ is tame. We also have $O(y_p)' \subset K - (V_1 \cup V_2 \cup \cdots \cup V_k) = M_k$ for some p, $1 \le p \le k$. Consider y_{k+1}. If $y_{k+1} \in \overline{V}_1 \cup \overline{V}_2 \cup \cdots \cup \overline{V}_k$, say, $y_{k+1} \in \overline{V}_r$, then $y_{k+1} \in V_r$ (recall: $\mathrm{Fr}(V_r)$ is tame). Set $V_{k+1} = V_r$.

Otherwise, $y_{k+1} \notin \overline{V}_1 \cup \overline{V}_2 \cup \cdots \cup \overline{V}_k$ and $\mathrm{Fr}(V_i)$ is tame for each $1 \le i \le k$. There are two cases to consider.

Case 1. If $y_{k+1} \notin O(y_p)'$ then $y_{k+1} \in K - (O(y_p)' \cup (\overline{V_1} \cup \overline{V_2} \cup \cdots \cup \overline{V_k}))$ which is an open set in K. Use regularity to find an open V_{k+1}such that $y_{k+1} \in V_{k+1}$ and $\overline{V}_{k+1} \cap (O(y_p)' \cup (\overline{V_1} \cup \overline{V_2} \cup \cdots \cup \overline{V_k})) = \emptyset$. We can assume with no loss of generality that $\mathrm{Fr}(V_{k+1})$ is tame. Furthermore, we can define $M_{k+1} = K - (V_1 \cup V_2 \cup \cdots \cup V_{k+1})$. Note that M_{k+1} is compact and contains $O(y_p)'$.

Case 2. If $y_{k+1} \in O(y_p)'$, then by Lemma (4.4) $O(y_{k+1})' \subset O(y_p)'$ which implies that $O(y_{k+1})' \subset O(y_p)' \subset K - (V_1 \cup V_2 \cup \cdots \cup V_k)$. But, since $y_{k+1} \notin O(y_{k+1})'$, y_{k+1} belongs to the open set $K - (O(y_{k+1})' \cup (\overline{V_1} \cup \overline{V_2} \cup \cdots \cup \overline{V_k}))$. As in case 1, this open set contains an open neighborhood V_{k+1} of y_{k+1} in K such that $\overline{V}_{k+1} \cap \overline{V}_i = \emptyset$ for all $1 \le i \le k$. Again we can assume that $\mathrm{Fr}(V_{k+1})$ is tame and that the compact set $M_{k+1} = K - (V_1 \cup V_2 \cup \cdots \cup V_{k+1})$ equals $(K - (V_1 \cup \cdots \cup V_k)) \cap (K - V_{k+1}))$ which contains $O(y_{k+1})'$.

Therefore we have obtained $M_{k+1} = K - (V_1 \cup V_2 \cup \cdots \cup V_k \cup V_{k+1})$ and have shown that $M_{k+1} \cap W$ is nonempty.

Now consider the nested compact sets $(M_1 \cap W) \supset (M_2 \cap W) \supset \ldots (M_k \cap W) \supset \ldots$. These sets are all compact and the collection $\{M_k \cap W \mid k \in N\}$ has the finite intersection property. Hence $\cap_{i=1}^{\infty}(M_i \cap W) \neq \emptyset$ and contains no point of the sequence $\{y_i\}$. Hence $\cap_{i=1}^{\infty}(M_i \cap W)$ is not in the range of f. Therefore, the set W must be uncountable. Therefore Theorem (4.1) is proved. $\square$

5. Questions

Note that our example of an infinitely periodic knot is non-prime. Are there any prime non-trivial infinitely periodic knots? If so, are all such examples wild at every point? Does the wild point set of an infinitely periodic knot always contain a Cantor set?

Acknowledgment

The authors would like to thank the referees for their valuable suggestions. Both the authors were partially supported by Caterpillar Fellowships.

Received March 28, 2003

Final version received March 4, 2004

Bradley University
Peoria, IL 61625
USA
onanyes@hilltop.bradley.edu
mtimm@bradley.edu

References

[1] A. Delgado, and M. Timm, *Spaces whose finite sheeted covers are homeomorphic to a fixed space*, Topology Appl. **129** (2003), 1–10.
[2] E. Flapan, *Infinitely Periodic Knots*, Can. J. Math. **XXXVII** (1), (1985), 17–28.
[3] R.H. Fox, *Knots and periodic transformations, Topology of* 3*-manifolds and related topics* (Proc. the Univ. of Georgia Inst. 1961), pp. 120–167, Prentice-Hall, Englewood Cliffs, N. J. 1962.
[4] J. McPherson, *A sufficient condition for an arc to be nearly polyhedral*, Proc. Amer. Math. Soc. **28**, (1971), 229–233.
[5] J. McPherson, *Wild arcs in three-space I (Families of Fox-Artin arcs)*, Pacific J. Math. **45**, (1973), 585–598.
[6] J. Munkres, Topology, A First Course, Prentice Hall, Englewood Cliffs, NJ 1975.
[7] H. Seifert, *Topologie dreidimensionalen gefaserter raume*, Acta Math. **60** (1933), 147–238.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

ON THE PURE BRAID GROUP OF A SURFACE

MIGUEL A. XICOTÉNCATL

Abstract. Given a closed surface $M \neq S^2$ or $\mathbb{R}P^2$, the classical pure braid group P_k is known to inject as a subgroup of the pure braid group of the surface, $P_k(M)$. Moreover, its normal closure $\langle\langle P_k \rangle\rangle$ is the kernel of the epimorphism $\varphi_* : P_k(M) \to (\pi_1 M)^k$ induced by the inclusion of the configuration space $F(M,k) \subset M^k$. In this article we exhibit an isomorphism between $\ker \varphi_*$ and $\pi_1 F_G(\mathbb{H}, k)$, where $M \approx \mathbb{H}/G$ and $F_G(\mathbb{H}, k)$ is an appropriate orbit configuration space.

1. Introduction

Given a topological space M, the configuration space of k-tuples of distinct points is defined as

$$F(M,k) = \{(m_1, \ldots, m_k) \in M^k \mid m_i \neq m_j \quad \text{if} \quad i \neq j\}.$$

Since their introduction by Fadell and Neuwirth [5], configuration spaces have proven to be very useful in geometry and algebraic topology. See [4] for a survey.

Furthermore, the case $M = \mathbb{R}^2$ is related to the classical braid groups as follows. Let Σ_k be the symmetric group acting on $F(\mathbb{R}^2, k)$ by permuting coordinates, and set B_k to be the fundamental group of the orbit space $F(\mathbb{R}^2, k)/\Sigma_k$ and $P_k = \pi_1 F(\mathbb{R}^2, k)$. Then B_k is isomorphic to the Artin's braid group on k strands and P_k is the k-stranded pure braid group, which is the kernel of the natural epimorphism $B_k \to \Sigma_k$.

By analogy with the classical case, given a finite dimensional manifold M one defines the pure braid group of M as $P_k(M) = \pi_1 F(M,k)$. Notice that the natural inclusion $\varphi : F(M,k) \subset M^k$ induces a homomorphism

$$\varphi_* : \pi_1 F(M,k) \longrightarrow (\pi_1 M)^k$$

which satisfies the following result of J. Birman [1].

Theorem (1.1). *For a closed smooth manifold M,*
1. *φ_* is an isomorphism if* $\dim M > 2$,
2. *φ_* is an epimorphism if* $\dim M = 2$.

Therefore, the group $P_k(M)$ is isomorphic to $(\pi_1 M)^k$ if $\dim M > 2$, being the most interesting case the one when M is a surface. In the case when $\dim M = 2$,

2000 *Mathematics Subject Classification*: Primary 20F36, 55R80; Secondary 20J06, 55R10, 55R20, 57N05.
Keywords and phrases: surface braid groups; Orbit configuration spaces.
Partially suported by CONACyT Grant No. 37296-E.

one obtains a short exact sequence

$$1 \longrightarrow \ker\varphi_* \longrightarrow P_k(M) \xrightarrow{\varphi_*} (\pi_1 M)^k \longrightarrow 1$$

and a description of the group $\ker\varphi_*$ is given next. Let $V \approx \mathbb{R}^2$ be a euclidean neighborhood in M. Then the inclusion $V \subset M$ induces an inclusion at the level of configuration spaces $i : F(\mathbb{R}^2, k) \to F(M, k)$ and thus a homomorphism of pure braid groups

$$i_* : P_k(\mathbb{R}^2) \longrightarrow P_k(M),$$

which is monomorphism for any compact surface different from S^2 or $\mathbb{R}P^2$, by a theorem of Birman [1] and Goldberg [6]. Thus P_k can be regarded as a subgroup of $P_k(M)$.

In [6] C. Goldberg has given a combinatorial description for $\ker\varphi_*$. Namely, recall that given a subgroup $H \leq G$ (not necessarily normal), its normal closure $\langle\langle H\rangle\rangle$ is defined to be the intersection of all normal subgroups of G containing H. That is to say, $\langle\langle H\rangle\rangle$ consists of all finite products $\prod g_i h_i g_i^{-1}$ of conjugates of elements in H. Then the following result was proven in [6]:

Theorem (1.2). *For any closed surface different from S^2 or $\mathbb{R}P^2$, the subgroup $\ker\varphi_* \leq P_k(M)$ is equal to $\langle\langle P_k\rangle\rangle$, the normal closure of P_k in $P_k(M)$.*

The purpose of this note is to provide a more geometrical description of $\ker\varphi_*$ (in the case when M is an orientable surface), since this group appears naturally as the fundamental group of an appropriate orbit configuration space (see sections 2 and 3). Namely,

Theorem (1.3). *For a closed, orientable surface M of genus $g > 1$, let $\mathbb{H}$ be the universal cover of M such that $M \approx \mathbb{H}/G$, where $G \cong \pi_1(M)$. Then there is a natural isomorphism*

$$\langle\langle P_k\rangle\rangle \cong \pi_1 F_G(\mathbb{H}, k)$$

where $F_G(\mathbb{H}, k)$ is the orbit configuration space of k points in $\mathbb{H}$.

2. Orbit configuration spaces

Let M be a finite dimensional manifold, G a group acting freely on M, and let $G \cdot m$ denote the orbit of the point $m \in M$ under the action of G. Inspired by [5], define the *orbit configuration space* of k points in M by:

$$F_G(M, k) = \{(m_1, \ldots, m_k) \in M^k \mid G \cdot m_i \neq G \cdot m_j \quad \text{if} \quad i \neq j\}$$

The spaces $F_G(M, k)$ were introduced in [8] as generalizations of ordinary configuration spaces, and their basic properties and some applications were studied. In the case when the canonical map $\pi : M \to M/G$ is a bundle projection there is an obvious relation to the ordinary configuration spaces given by the following:

Theorem (2.1). *If the canonical map $M \to M/G$ is a bundle projection, with classifying map $f : M/G \to BG$, then the space $F_G(M, k)$ is the total space of* *the pull-back of the principal fibration $G^k \to EG^k \to BG^k$ along the composition $F(M/G,k) \hookrightarrow (M/G)^k \xrightarrow{f^k} BG^k$. Thus there are maps of G^k-bundles:*

$$\begin{tikzcd}
G^k \arrow[r, equal] \arrow[d] & G^k \arrow[r, equal] \arrow[d] & G^k \arrow[d] \\
F_G(M,k) \arrow[r, hook] \arrow[d] & M^k \arrow[r] \arrow[d] & (EG)^k \arrow[d] \\
F(M/G,k) \arrow[r, hook] & (M/G)^k \arrow[r, "f^k"] & (BG)^k
\end{tikzcd}$$

Therefore, there is a principal G^k-bundle $F_G(M,k) \to F(M/G,k)$. In other words, the group G^k acts "coordinate-wise" on the space $F_G(M,k)$ and the quotient $F_G(M,k)/G^k$ is homeomorphic to $F(M/G,k)$.

In particular, there is a fibration $F_G(M,k) \to F(M/G,k) \to BG^k$ that can be used, for example, to do homological calculations; see for instance [3] and [9]. For simplicity, let us consider here the case when G is a discrete group. Using the Serre spectral sequence of the fibration above, or equivalently, the spectral sequence for a covering, one obtains the following result:

THEOREM (2.2). *Let G be a discrete group such that the canonical map $M \to M/G$ is a covering projection. Then, there is a spectral sequence whose E_2-term is given by*

$$E_2^{*,*} = H^*\left(G^k; \underline{H^* F_G(M,k)}\right)$$

and which converges to $H^*F(M/G,k)$. Here the action of G^k on the local coefficient system is induced by the action of G^k on $F_G(M,k)$.

3. Proof of the theorem (1.3)

Let $\mathbb{H} = \{z \in \mathbb{C} \mid Im(z) > 0\}$ be the complex upper-half plane, acted on by $PSL(2,\mathbb{R})$ through Möbius transformations. By Poincaré's theorem on Fuchsian groups, it is well known that for every compact, orientable surface M of genus $g > 1$, there is a discrete subgroup G of $PSL(2,\mathbb{R})$ acting properly discontinuously on M, and such that $M \approx \mathbb{H}/G$. Thus $\mathbb{H}$ is the universal cover of M. Moreover since $\mathbb{H}$ is contractible, $\pi_1(M) \cong G$.

From the preceding theorem, there is a covering $G^k \to F_G(\mathbb{H},k) \to F(M,k)$ whose long homotopy exact sequence reduces to:

$$1 \longrightarrow \pi_1 F_G(\mathbb{H},k) \longrightarrow P_k(M) \longrightarrow G^k \longrightarrow 1 \tag{3.1}$$

since $\pi_1(M) \cong G$. Now there is a map of extensions

$$\begin{tikzcd}
1 \arrow[r] & \langle\langle P_k \rangle\rangle \arrow[r] \arrow[d, dashed] & P_k(M) \arrow[r, "\varphi_*"] \arrow[d, equal] & (\pi_1 M)^k \arrow[r] \arrow[d, "\cong"', "\text{loop lifting}"] & 1 \\
1 \arrow[r] & \pi_1 F_G(\mathbb{H},k) \arrow[r] & P_k(M) \arrow[r] & G^k \arrow[r] & 1
\end{tikzcd}$$

where the right hand side isomorphism is given by the usual lifting of loops to the total space. This induces an isomorphism $\langle\langle P_k \rangle\rangle \cong \pi_1 F_G(\mathbb{H},k)$. $\square$

Another consequence of Theorem (2.1) is the existence of a spectral sequence abutting to the cohomology of the pure braid group $P_k(M)$, as stated next. Explicit calculations will be done elsewhere.

THEOREM (3.2). *For a closed, orientable surface M of genus $g > 1$, there is a discrete subgroup $G \leq PSL(2, \mathbb{R})$ and a spectral sequence*

$$E_2^{*,*} = H^*\left(G^k; \underline{H^* F_G(\mathbb{H}, k)}\right)$$

that converges to the cohomology of $P_k(M)$.

Proof. Using the fact that M is an Eilenberg–Mac Lane space of type $K(G, 1)$ and the Fadell–Neuwirth fibrations [5], it is easy to prove that $F(M, k)$ is an Eilenberg–Mac Lane space of type $K(P_k(M), 1)$. Now apply Theorem (2.2). Equivalently, this is the Lyndon-Hochschild-Serre spectral sequence associated to the extension (3.1). □

Received June 30, 2003

Final version received April 14, 2004

DEPTO. DE MATEMÁTICAS
CINVESTAV DEL IPN
APDO. POSTAL 14-740
07300 MÉXICO, D.F.
MEXICO
xico@math.cinvestav.mx

REFERENCES

[1] J.S. BIRMAN, *On braid groups*, Comm. Pure and Applied Math. **22** (1969), 41–72.

[2] J.S. BIRMAN, *Braids, links and mapping class groups*, Ann. Math. Studies, **82**, Princeton Univ. Press, Princeton, NJ, 1975.

[3] F.R. COHEN, M. XICOTÉNCATL, *On orbit configuration spaces associated to the Gaussian integers: homology and homotopy groups*, Topology Appl. **118** (2002), 17–29.

[4] E.R. FADELL, S.Y. HUSSEINI, Geometry and Topology of Configuration Spaces, Springer Monographs in Mathematics, Springer-Verlag, 2000.

[5] E. FADELL, L. NEUWIRTH, *Configuration spaces*, Math. Scand. **10** (1962), 119–126.

[6] C.H. GOLDBERG, *An exact sequence of braid groups*, Math. Scand. **33** (1973), 69–82.

[7] S. KATOK, Fuchsian Groups, The University of Chicago Press, 1992.

[8] M.A. XICOTÉNCATL, *Orbit configuration spaces, infinitesimal braid relations in homology and equivariant loop spaces*, Ph.D. Thesis, University of Rochester, 1997.

[9] M.A. XICOTÉNCATL, *On orbit configuration spaces and the rational cohomology of* $F(\mathbb{R}P^n, k)$, Une dègustation topologique: homotopy theory in the Swiss Alps (Arolla, 1999), 233–249, Contemp. Math. **265**, Amer. Math. Soc.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

A SIMPLIFIED INDEX FOR ROOTS

XUEZHI ZHAO

Abstract. Let $f: Y \to X$ be a map between compact connected polyhedra. We shall simplify the existing index for roots at a point $x_* \in X$, where x_* is a local separating point of X. We also show that this simplified index is easy to compute, so it can be used to estimate the number of roots in practice.

1. Root classes and their indices

Let $f: Y \to X$ be a map between compact connected polyhedra. For a given point $x_* \in X$, a point $y \in Y$ is said to be a *root* of f at x_* if $f(y) = x_*$. A natural question is how to describe the set $f^{-1}(x_*)$; this question is discussed in root theory, which is a branch of Nielsen's fixed point theory (see [2] for an introduction to this topic).

Here we give a brief account of root theory (see [5] for more details). Two roots y_1 and y_2 of f at x_* are said to be in the same *root class* if there is a path α in Y joining them, such that $\langle f \circ \alpha \rangle = 1 \in \pi_1(X, x_*)$. Thus, $f^{-1}(x_*)$ can be split into several root classes, each of which is an isolated set (see below) in $f^{-1}(x_*)$.

If we choose a root y_*, then we have a homomorphism $f_\pi: \pi_1(Y, y_*) \to \pi_1(X, x_*)$. For any root y, pick a path γ from y_* to y; we can define an element $\langle f \circ \gamma \rangle \in \pi_1(X, x_*)$. The corresponding element in the right coset of $\mathrm{Im} f_\pi$ in $\pi_1(X, x_*)$ is independent of the choice of the path γ. Such a correspondence is written by

$$\phi\colon f^{-1}(x_*) \to \pi_1(X, x_*)/\mathrm{Im} f_\pi.$$

Moreover, two roots y_1 and y_2 are in the same root class if and only if $\phi(y_1) = \phi(y_2)$ (see [5, p. 133, 6.1 Theorem]). Thus, ϕ gives an alternative way to define root classes.

Given an isolated root set R of f at x_*, i.e. there is an open neighborhood U of R in Y such that $\bar{U} \cap f^{-1}(x_*) = R$, the sequence of maps

$$Y \xrightarrow{j} (Y, Y - R) \xleftarrow{e} (\bar{U}, \bar{U} - R) \xrightarrow{f} (X, X - x_*)$$

induces a homology homomorphism

$$f_* e_*^{-1} j_*\colon H_*(Y) \to H_*(X, X - x_*)$$

that is called the *index homomorphism* or simply *index* of the root set R. Throughout this paper, all homology groups are assumed to have rational coefficients $\mathbb{Q}$.

2000 *Mathematics Subject Classification*: 55M20.

Keywords and phrases: root, local separating point, index.

Supported by NSF of China and Beijing Natural Sciences Foundation.

The most important property of the index of any root class is its homotopy invariance, hence it can be used to estimate the number of roots.

As we have seen, such an index is a homomorphism between homology groups, so it is not numerical and hence difficult to compute. It is a natural idea to try to reduce this index in some useful cases. An interesting case is when X and Y are both oriented closed manifolds of the same dimension; here $H_*(X, X - x_*) = \mathbb{Q}$, and the index is in some sense referred to as the degree of the map f (see [5, 7.3 Corollary]). A great improvement in this case was made in [1] and [3], where manifolds were allowed to be non-orientable.

It is the purpose of this paper to study a simpler "index", which is just the 1-dimensional part of the original index.

2. Local separating points

We shall consider the index homomorphism in dimension one only, namely, the 1-*dimensional index homomorphism.* Let $f\colon Y \to X$ be a map between compact polyhedra, and $x_* \in X$. In order to get a non-trivial 1-dimensional index homomorphism, we have to assume that $H_1(X, X - x_*)$ is non-trivial. Fortunately, this algebraic condition has a simple equivalent geometric interpretation, namely the existence of a local separating point.

Definition (2.1). A point p in a topological space X is said to be a *local separating point* of X if there is a connected neighborhood O_p of p in X such that $O_p - p$ is disconnected.

PROPOSITION (2.2). *Let p be point in a compact polyhedron X. Then p is a local separating point if and only if $H_1(X, X - p)$ is non-trivial.*

Proof. Since X is a compact polyheron, we can find a regular neighborhood W of p in X, which is contractible. In the exact sequence

$$\to H_1(W) \to H_1(W, W - p) \to \tilde{H}_0(W - p) \to \tilde{H}_0(W) \to \tilde{H}_0(W, W - p),$$

we know that $H_1(W) = 0$ and $\tilde{H}_0(W) = H_0(W, p) = 0$. Thus $H_1(W, W - p) \neq 0$ if and only if $\tilde{H}_0(W - p) \neq 0$, which is the case if and only if $W - p$ is disconnected, that is, if and only if p is a local separating point. □

In [6], the definition of local separating point was generalized as follows:

Definition (2.3). A connected subpolyhedron A of X is said to be a *local separating set* of X if there is a connected neighborhood N of A in X such that the set $N - A$ is not connected.

Clearly, p is a local separating point if and only if $\{p\}$ is a local separating set.

3. 1-dimensional index homomorphism

Consider a map $f\colon Y \to X$ and $x_* \in X$. In this section we will assume that x_* is a local separating point of X. The restriction $f_* e_*^{-1} j_*\colon H_1(Y) \to H_1(X, X - x_*)$ of the index homomorphism to dimension one is said to be the 1-*dimensional index homomorphism* and is denoted by $v_1(f, x_*; \cdot)$.

As in the case of the original index, the homotopy invariance of the 1-dimensional index homomorphism implies the following result.

Theorem (3.1). *If f has n root classes at x_* with non-trivial* 1*-dimensional index homomorphism (i.e., $v_1(f, x_*; \cdot)$ is a non-zero homomorphism at each of these n root classes), then any map g homotopic to f has at least n roots at x_*, i.e. $|g^{-1}(x_*)| \geq n$.*

Proof. Observe that for any isolated root set R, a non-trivial 1-dimensional index homomorphism implies certainly a non-trivial index homomorphism. Since a root class corresponding to no root under homotopy must have trivial index homomorphism ([5, p. 137]), we are done. □

Now, we shall illustrate how to compute our 1-dimensional index homomorphism for maps between compact polyhedra.

Take a regular neighborhood W of x_*. As in the proof of Proposition (2.2), we have an exact sequence

$$0 \to H_1(W, W - x_*) \stackrel{\partial_{*,1}}{\to} \tilde{H}_0(W - x_*) \stackrel{i_{*,0}}{\to} \tilde{H}_0(W) \to 0.$$

Thus, $H_1(W, W - x_*)$ is generated by the 1-simplices whose ending points are in different components of $W - x_*$.

Let R be a connected isolated root set of $f \colon Y \to X$ at x_*, and take a regular neighborhood N of R. Then the 1-dimensional index homomorphism $v_1(f, x_*; R)$ of R is given by the composition

$$H_1(Y) \stackrel{j_{*,1}}{\to} H_1(Y, Y - R) \stackrel{e_{*,1}}{\leftarrow} H_1(N, N - R) \stackrel{f_{*,1}}{\to} H_1(X, X - x_*).$$

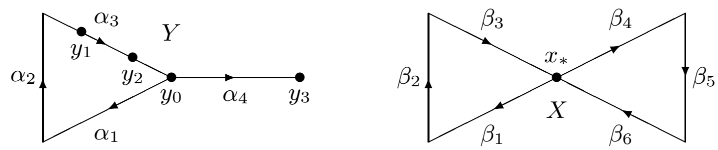
Note that $f_{*,1} e_{*,1}^{-1} j_{*,1}$ is non-trivial if and only if $f_{*,1} e_{*,1}^{-1}|_{\mathrm{Im} j_{*,1}} \colon \mathrm{Im} j_{*,1} \to H_1(X, X - x_*)$ is non-trivial, where $\mathrm{Im} j_{*,1} \subset H_1(Y, Y - R)$. From the exact sequence $\cdots \to H_1(Y) \stackrel{j_{*,1}}{\to} H_1(Y, Y - R) \stackrel{\partial_{*,1}}{\to} H_0(Y - R) \to \cdots$, we get that $\mathrm{Im} j_{*,1} = \mathrm{Ker} \partial_{*,1}$. For an element in $H_1(Y, Y - R)$ represented by a path $\alpha \colon I \to Y$, we know that $\partial_{*,1}(\alpha) = \alpha(1) - \alpha(0)$. Thus, such an element lies in $\mathrm{Im} j_{*,1} = \mathrm{Ker} \partial_{*,1}$ if and only if $\alpha(0)$ and $\alpha(1)$ are in the same component of $Y - R$. Therefore, $\mathrm{Im} j_{*,1}$ is generated by the set of paths $\gamma \colon I \to Y$ whose end points either lie in the same component of $Y - R$ or are the same points in R. Notice that $e_{*,1}$ is an isomorphism, hence the generators of $\mathrm{Im} j_{*,1}$ can be chosen to be the paths in N. Since the neighborhood N of R can be chosen to be very small, we may assume that $f(N) \subset W$. Thus, for any path $\gamma \colon I \to N$ which is considered as element in $H_1(N, N - R)$, $f_{*,1}([\gamma]) = 0 \in H_1(X, X - x_*)$ if and only if either $f(\gamma(0))$ and $f(\gamma(1))$ are in the same component of $W - x_*$ or $f(\gamma(0)) = f(\gamma(1)) = x_*$. So we have proved:

Theorem (3.2). *Let R be a connected isolated root set of $f \colon Y \to X$ at x_*. If $v_1(f, x_*; R) \neq 0$, then there are two components of $N - R$ which are in the same component of $Y - R$ and are mapped by f into different components of $W - x_*$, where N and W are regular neighborhoods of R and x_* respectively.* □

Corollary (3.3). *If a connected isolated root set R of $f \colon Y \to X$ at x_* is not a local separating set in Y, then $v_1(f, x_*; R) = 0$.* □

4. An example

Example (4.1). Let X and Y be the two compact 1-dimensional polyhedra (i.e. graphs) which are shown in the figure below. The map $f\colon Y \to X$ is the piece-wise linear map such that $f\circ\alpha_1 = \beta_1$, $f\circ\alpha_2 = \beta_2$, $f\circ\alpha_3 = \beta_3\beta_4\beta_5\beta_6\beta_1\beta_2\beta_3$ *and* $f \circ \alpha_4 = \beta_1\beta_2\beta_3$.



Clearly, $H_1(Y) = \mathbb{Q}$, and we take as its generator the homology class $[\alpha_1\alpha_2\alpha_3]$. A basis for $H_1(X, X - x_*) = \mathbb{Q}^3$ can be chosen to consist of the classes $\{[\beta_1^{-1}\beta_3^{-1}], [\beta_1^{-1}\beta_4], [\beta_1^{-1}\beta_6^{-1}]\}$. As indicated in the figure, we have the set of roots $f^{-1}(x_*) = \{y_0, y_1, y_2, y_3\}$.

Consider the root y_1. Its regular neighborhood N can be chosen as a small line segment which does not meet any root other than y_1. When N is small enough, it is mapped by f into a sub-path of $\beta_3\beta_4$, centered at x_*. In 1-dimensional homology, for the composition

$$H_1(Y) \overset{j_{*,1}}{\to} H_1(Y, Y - y_1) \overset{e_{*,1}}{\leftarrow} H_1(N, N - y_1) \overset{f_{*,1}}{\to} H_1(X, X - x_*),$$

we have $f_{*,1}e_{*,1}^{-1}j_{*,1}([\alpha_1\alpha_2\alpha_3]) = f_{*,1}([N]) = [\beta_3\beta_4] = -[\beta_1^{-1}\beta_3^{-1}] + [\beta_1^{-1}\beta_4]$. Using the ordered basis of $H_1(Y)$ and $H_1(X, X - x_*)$ mentioned above, the 1-dimensional index homomorphism $v_1(f, x_*; \{y_1\})$ can be simply written as the vector $(-1, 1, 0)$.

Observe that $\pi_1(X, x_*) = \mathbb{Z} * \mathbb{Z}$, with generators the loops $a = \langle\beta_1\beta_2\beta_3\rangle$ and $b = \langle\beta_4\beta_5\beta_6\rangle$. Thus the image of $f_\pi\colon \pi_1(Y, y_0) \to \pi_1(X, x_*)$ in $\pi_1(X, x_*)$ is an infinite cyclic group generated by aba. Let γ be the line segment from y_0 to y_1; then $\langle f\circ\gamma\rangle = a^{-1}b^{-1} \in \pi_1(X, x_*)$, which is the element $\phi(y_1)$ in $\pi_1(X, x_*)/\mathrm{Im} f_\pi$ corresponding to y_1.

In a similar way, we obtain the rest of the data in the following table:

root	$v_1(f, x_*; \cdot)$	$\phi(\cdot)$
y_0	$(-1, 0, 0)$	1
y_1	$(-1, 1, 0)$	$a^{-1}b^{-1}$
y_2	$(0, 0, -1)$	a^{-1}
y_3	$(0, 0, 0)$	a

Since 1, a^{-1}, $a^{-1}b^{-1}$ are different elements in $\pi_1(X, x_*)/\mathrm{Im} f_\pi$, and since $a = (aba)a^{-1}b^{-1}$, f has three root classes, $\{\,\{y_0\}, \{y_1, y_3\}, \{y_2\}\,\}$, each of which has non-trivial 1-dimensional index homomorphism. Thus, by Theorem (3.1), any map homotopic to f has at least 3 roots at x_*.

The author thanks the referee for many helpful comments.

Received June 27, 2003

Final version received May 05, 2004

Department of Mathematics
Capital Normal University
Beijing 100037
P. R. China
zhaoxve@mail.cnu.edu.cn

References

[1] R. Brooks, R. F. Brown and H. Schirmer, *The absolute degree and the Nielsen root number of compositions and Cartesian product of maps*, Topology Appl. **116** (1992), 5–21.

[2] R. F. Brown, The Lefschetz Fixed Point Theorem, Scott Foresman and Co., Glenview, IL, 1971.

[3] R. Dobrenko and J. Jezierski, *The coincidence Nielsen number on non-orientable manifolds*, Rocky Mountain J. Math. **23** (1993), 67–85.

[4] B. Jiang, *Lectures on Nielsen fixed point theory*, Contemp. Math. **14**, Amer. Math. Society, Providence, Rhode Island, 1983.

[5] T. H. Kiang, The Theory of Fixed Point Classes, Springer-Verlag, Berlin/Science Press, Beijing, 1989.

[6] X. Zhao, *Realization of fixed point sets*, Acta Math. Sinica, New Series **12** (1996), 71–76.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

ON GENUS ACTIONS OF FINITE SIMPLE GROUPS ON HANDLEBODIES AND BOUNDED SURFACES

BRUNO P. ZIMMERMANN

ABSTRACT. Let G be a finite nonabelian simple group of isometries of a 3-dimensional handlebody V or of a compact bounded surface of least possible genus g (a "genus action" of G). Generalizing analogous results for genus actions of finite simple groups on closed surfaces, we prove that G is a normal subgroup of small index in the orientation-preserving isometry group H of V, and that H is canonically isomorphic to a subgroup of the automorphism group of G. We present also some related results on finite group actions of maximal possible order, in particular for nilpotent and simple groups.

1. Introduction

By the *genus* of a finite group G we understand the least genus of a closed orientable surface on which G acts by orientation-preserving diffeomorphisms. In analogy, the *handlebody genus* of G is the smallest genus of a closed orientable 3-dimensional handlebody V with an orientation-preserving G-action. In each case, such an action on a surface or handlebody of least genus is called a *genus action* of G.

The algebraic or real genus of a compact surface with nonempty boundary is the rank of its free fundamental group, and the *algebraic* or *real genus* of a finite group G is the least algebraic genus of a compact bounded surface (possibly nonorientable) with a G-action (possibly orientation-reversing). Given an action of a finite group G on a bounded surface of algebraic genus g, by taking a (possibly twisted) product of the surface and the G-action with a closed interval, one obtains an orientation-preserving action of G on an orientable handlebody of genus g. Hence the handlebody genus of a finite group G is smaller or equal to its algebraic genus.

We consider genus actions of finite simple groups on handlebodies (and simple group will always mean nonabelian simple group in the following). The only finite simple group which acts on the 3-ball (the handlebody of genus zero) is the dodecahedral or alternating group $\mathbb{A}_5$, and there are no finite simple groups acting on the solid torus. Therefore, in the following, we can concentrate on the case of handlebodies of genus $g > 1$.

Three-dimensional handlebodies V are uniformized by Schottky groups (free Kleinian groups acting by isometries on hyperbolic 3-space and by conformal maps on its sphere at infinity). Then the interior of V becomes a complete

2000 *Mathematics Subject Classification*: 57M60, 57S25, 30F50.

Keywords and phrases: handlebody, bounded surface, finite group action, finite simple group.

hyperbolic 3-manifold of infinite volume and its boundary a Riemann surface (or a hyperbolic surface if $g > 1$). By the Proposition in the introduction of [12], if the action of a finite group G on a surface F extends to a 3-dimensional handlebody V then V can be uniformized by a Schottky group such that the action of G on F extends to an action of G on the interior of V by isometries (also, G acts by conformal maps resp. isometries on $F = \partial V$). Hence, if a finite group acts by orientation-preserving diffeomorphisms on a handlebody V then it acts also by isometries, for some uniformization of V by a Schottky group. Note that the isometry group of a handlebody of genus $g > 1$ is finite because this is the case for the isometry group of a closed hyperbolic surface.

We say that a group is 2-*generated* if it is generated by two elements, and (m, n)-*generated* if it is generated by two elements of orders m and n. Our main result is the following

Theorem (1.1). *Let G be a finite simple group acting by isometries on a 3-dimensional handlebody V of least genus g, or on a bounded surface V of least algebraic genus g (a genus action of G); assume $g > 1$. Then G is a normal subgroup, of index at most 10, in the orientation-preserving isometry group H of V. If G is $(2, n)$-generated than G has index 1, 2 or 4 in H, and if it is $(2, 3)$-generated than the index is 1 or 2. In any case, the map induced by conjugation from H to the automorphism group $\operatorname{Aut} G$ of G is injective, hence $G \subset H \subset \operatorname{Aut} G$.*

Analogous results for genus actions on closed surfaces are obtained in the papers [17] and [2] which motivated the present paper. For bounded surfaces, a version of the Theorem (under the hypothesis that G is $(2, n)$-generated) is proved in [6]. In all these papers, the proofs use the theory of Fuchsian or noneuclidean crystallographic groups and the formula of Riemann-Hurwitz. In the present paper, we apply instead the corresponding theory for finite group actions on handlebodies which uses the language of finite graphs of finite groups and their Euler characteristics [8].

It is conjectured (and known for many classes of simple groups, see [17]) that every finite simple group is $(2, n)$-generated. The part of the Theorem without such a hypothesis uses the classification of the finite simple groups, or more precisely the fact that every finite simple group is 2-generated [1].

A particular case of genus actions are the actions of maximal possible order. In the case of closed orientable surfaces of genus $g > 1$, by a classical result of Hurwitz and as a consequence of the formula of Riemann-Hurwitz, the maximal order of a finite group of orientation-preserving diffeomorphisms is $84(g-1)$, and the corresponding groups are called *Hurwitz groups*. In the case of a handlebody of genus $g > 1$ or of a bounded surface of algebraic genus $g > 1$, the maximal order is $12(g - 1)$, and we call these groups *maximal handlebody* resp. *maximal bounded surface groups* (see [19], [8], Theorem 7.2, for the case of handlebodies, and [5] for the case of bounded surfaces). Note that, by an above observation, the bound for handlebodies implies that for bounded surfaces, and that every maximal bounded surface group is a maximal handlebody group.

We complement the main Theorem by some results on finite group actions of maximal possible order on closed and bounded surfaces, and on handlebodies.

PROPOSITION (1.2). *Every maximal handlebody group is 2-generated. Every maximal bounded surface group G is $(2,n)$-generated, for some n, and $(2,3)$-generated if G is simple (or perfect). There are infinitely many different maximal handlebody groups which are not $(2,n)$-generated, for any n.*

Comparing the maximal actions on surfaces extending to handlebodies with the Hurwitz actions, the following holds

PROPOSITION (1.3). *Let G be a finite group of orientation-preserving isometries of maximal possible order $12(g-1)$ of a handlebody V of genus $g>1$. Then either G is a normal subgroup of index at most three in the orientation-preserving isometry group H of the hyperbolic surface $F=\partial V$, or H is a Hurwitz group of order $84(g-1)$. In the latter case, if H is simple then H is the smallest Hurwitz group $\mathrm{PSL}(2,7)$ of order 168, acting on Klein's quartic of genus three, and G is the symmetric group $\mathbb{S}_4$ of order 24.*

In fact, if H is a Hurwitz group it might always be true that H is isomorphic to $\mathrm{PSL}(2,7)$.

The finite simple groups which are maximal handlebody or maximal bounded surface groups of even genus can be characterized as follows.

PROPOSITION (1.4). a) *For a prime power q, the linear fractional group $\mathrm{PSL}(2,q)$ is a maximal handlebody group of order $12(g-1)$ (respectively a maximal bounded surface group) if and only if q is different from 2, 7, 9, 11 and 3^{2m+1}.*

b) *A finite simple group G is a maximal handlebody group (respectively a maximal bounded surface group) of even genus if and only if G is isomorphic to a linear fractional group $\mathrm{PSL}(2,q)$, for a prime power $q\equiv\pm 3 \bmod 8$ different from 11 and 3^{2m+1}.*

Finally, for finite nilpotent groups our methods imply the following result (proved in [23] for the case of closed surfaces, and in [7] for the case of bounded surfaces; see also the Remark in section 6 for the case of closed surfaces). Various upper and lower bounds for finite group actions on handlebodies and bounded surfaces are obtained in [10], and in [8], Theorem 7.6, for the case of finite cyclic and abelian groups.

PROPOSITION (1.5). *A finite nilpotent group G of orientation-preserving diffeomorphisms of a handlebody or a bounded surface of (algebraic) genus $g>1$ has order at most $8(g-1)$, and this upper bound is attained for infinitely many different genera g. Moreover, if G has maximal order $8(g-1)$ then it is a 2-group.*

2. Proof of Theorem (1.1)

For a finite graph of finite groups $(\Gamma,\mathcal{G})$, we denote by $\pi_1(\Gamma,\mathcal{G})$ its fundamental group (that is the iterated free product with amalgamation and HNN-extension over the vertex groups, amalgamated over the edge groups of a maximal tree, with the HNN-generators corresponding to the edges in the complement of the chosen maximal tree), and by

$$\chi(\Gamma,\mathcal{G}) := \sum 1/|G_v| - \sum 1/|G_e|$$

its Euler characteristic; the sum is extended over all vertex groups G_v resp. edge groups G_e of $(\Gamma, \mathcal{G})$. For example, the graph of groups

$$\Gamma(B_1, A, B_2)$$

with one edge with edge group A and two vertices with vertex groups B_1 and B_2 has Euler charateristic $\chi(\Gamma(B_1, A, B_2)) = 1/|B_1| - 1/|B_2| - 1/|A|$ and fundamental group $\pi_1(\Gamma(B_1, A, B_2)) \cong B_1 *_A B_2$.

A finite graph of finite groups $(\Gamma, \mathcal{G})$ is *admissible* if it satisfies a certain set of normalized or unnormalized conditions (see [8]); most importantly, the vertex groups of $(\Gamma, \mathcal{G})$ are spherical groups (finite subgroups of the orthogonal group $SO(3)$), and the edge groups are cyclic groups which are either trivial or maximal cyclic in the adjacent vertex goups. Equivalently, $(\Gamma, \mathcal{G})$ is associated to a handlebody orbifold whose orbifold fundamental group is isomorphic to $\pi_1(\Gamma, \mathcal{G})$ (see [21]).

For a finite group G, a finite graph of finite groups $(\Gamma, \mathcal{G})$ is *G-admissible* if it is admissible and there exists an epimorphism from $\pi_1(\Gamma, \mathcal{G})$ onto G with torsionfree kernel (or equivalently, the epimorphism is injective on the vertex groups). Now, in a similar way as finite group actions on closed surfaces are connected to Fuchsian groups and the formula of Riemann-Hurwitz, finite group actions on handlebodies are connected to admissible graphs of groups, their fundamental groups and the multiplicativity of their Euler characteristics under finite coverings. In particular, the following holds ([8], [21]).

PROPOSITION (2.1). *A finite group G acts on a handlebody V_g of genus g if and only if there exists a G-admissible finite graph of finite groups $(\Gamma, \mathcal{G})$ such that*

$$g - 1 = -\chi(\Gamma, \mathcal{G})|G| \, .$$

An analogous result for finite group actions on bounded surfaces is proved in [9]; in this case the vertex groups are finite subgroups of the orthogonal group $O(2)$, that is cyclic or dihedral groups (cyclic groups in the orientation-preserving case), and the edge groups are trivial or of order two (corresponding to reflections in dihedral vertex groups).

The maximal negative Euler characteristic of an admissible graph of groups is $-1/12$, and there are exactly the following four admissible graphs of groups with this Euler characteristic ([20], [8], p.401, Chart B)

$$\Gamma(\mathbb{D}_2, \mathbb{Z}_2, \mathbb{D}_3), \;\; \Gamma(\mathbb{D}_3, \mathbb{Z}_3, \mathbb{A}_4), \;\; \Gamma(\mathbb{D}_4, \mathbb{Z}_4, \mathbb{S}_4), \;\; \Gamma(\mathbb{D}_5, \mathbb{Z}_5, \mathbb{A}_5)$$

(where $\mathbb{D}_n$ denotes the dihedral group of order $2n$ and $\mathbb{A}_4$, $\mathbb{S}_4$ and $\mathbb{A}_5$ the tetrahedral, octahedral and dodecahedral group, respectively). It follows then from Proposition (2.1) that the maximal possible order of a finite group G acting on a handlebody of genus $g > 1$ is $12(g - 1)$, and that the maximal handlebody groups are exactly the finite admissible quotients (i.e., by torsionfree subgroups) of one of the following four free products with amalgamation

$$\mathbb{D}_2 *_{\mathbb{Z}_2} \mathbb{D}_3, \;\; \mathbb{D}_3 *_{\mathbb{Z}_3} \mathbb{A}_4, \;\; \mathbb{D}_4 *_{\mathbb{Z}_4} \mathbb{S}_4, \;\; \mathbb{D}_5 *_{\mathbb{Z}_5} \mathbb{A}_5.$$

The maximal bounded surface groups are exactly the finite quotients of the first group $\mathbb{D}_2 *_{\mathbb{Z}_2} \mathbb{D}_3$, isomorphic to the extended modular group $\mathrm{PGL}(2,\mathbb{Z})$ (which is the only one of the four groups with cyclic and dihedral vertex groups).

The second largest Euler characteristic is $-1/8$, realized by the two admissible graphs of groups

$$\Gamma(\mathbb{D}_2, \mathbb{Z}_2, \mathbb{D}_4), \quad \Gamma(\mathbb{D}_3, \mathbb{Z}_3, \mathbb{S}_4)\,,$$

with fundamental groups

$$\mathbb{D}_2 *_{\mathbb{Z}_2} \mathbb{D}_4, \quad \mathbb{D}_3 *_{\mathbb{Z}_3} \mathbb{S}_4\,,$$

and consequently the second largest order is $8(g-1)$. After this, the next largest orders are $20(g-1)/3$ and $6(g-1)$ (see [8], p. 401, chart B). We note that, for *orientation-preserving* actions on bounded surfaces of algebraic genus $g > 1$, the largest possible order is $6(g-1)$ and realized by the graph of groups $\Gamma(\mathbb{Z}_2, 1, \mathbb{Z}_3)$ whose fundamental group is isomorphic to the modular group $\mathrm{PSL}(2,\mathbb{Z}) \cong \mathbb{Z}_2 * \mathbb{Z}_3$. Hence the corresponding groups are exactly the finite admissible quotients of the modular group or, equivalently, the finite groups which are $(2,3)$-generated.

Suppose now that G is a finite simple group as in Theorem (1.1). We have to prove that G is normal in the orientation-preserving isometry group H of V. The crucial step of the proof is the following. Using the classification of the finite simple groups, it is shown in [1] that every finite simple group is generated by two elements. If G is generated by two elements of orders m and n then the graph of groups $\Gamma(\mathbb{Z}_m, 1, \mathbb{Z}_n)$ is G-admissible, and, by Proposition (2.1), G acts on a handlebody of genus $|G|(1 - 1/m - 1/n) + 1$. Since, by hypothesis, g is the least genus for a G-action, this implies

$$g - 1 < |G| \leq |H| \leq 12(g-1).$$

It follows that the index $|H|/|G|$ of G in H is at most 11. By left multiplication of left cosets of G in H with elements of H, we obtain a homomorphism $\phi : H \to \mathbb{S}_{11}$ from H to the symmetric group $\mathbb{S}_{11}$ of degree 11. Since $\phi(G)$ fixes the coset G of the unit element, ϕ restricts to a homomorphism $\phi : G \to \mathbb{A}_{10}$ from G to the alternating group $\mathbb{A}_{10}$. Note that the image $\phi(G)$ is trivial if and only if G is normal in H.

Assume, by contradiction, that G is not normal in H. Then $\phi(G)$ is a non-trivial subgroup of $\mathbb{A}_{10}$, and the kernel of $\phi : G \to \mathbb{A}_{10}$ is trivial (because G is simple). Thus G is isomorphic to a subgroup of the alternating group $\mathbb{A}_{10}$. By [4] or [13], the simple subgroups of $\mathbb{A}_{10}$ are exactly the alternating groups $\mathbb{A}_5$ - $\mathbb{A}_{10}$ and the linear fractional groups $\mathrm{PSL}(2,7)$ and $\mathrm{PSL}(2,8)$. It is well known and easy to prove that all these groups are generated by an element of order two and another element of some order n. As above, this implies that G acts on a handlebody of genus $|G|(1 - 1/2 - 1/n) + 1$, and that

$$2(g-1) < |G| \leq |H| \leq 12(g-1).$$

It follows that G has index at most five in H. Repeating the above argument we find that G is isomorphic to a subgroup of the alternating group $\mathbb{A}_4$ which is a contradiction because $\mathbb{A}_4$ is solvable.

Hence we have proved that G is a normal subgroup of H. Moreover, G has index at most 11 in H and $|G| > g - 1$. If this index is at least 7 then

$|H| > 7(g-1)$ and H has order $8(g-1)$ or $12(g-1)$. The fundamental groups of the corresponding 6 graphs of groups listed above are all generated by involutions. It follows that H has no quotients of odd orders 7, 9 or 11, and in particular the index of G in H is at most 10. Similarly, if G is $(2,n)$-generated then $|G| > 2(g-1)$ and has index 1, 2 or 4 in H; if G is $(2,3)$-generated then $|G| \geq 6(g-1)$ and has index 1 or 2 in H.

Finally, we prove that the action of H on G by conjugation is faithful or, equivalently, that the centralizer $C_H G$ of G in H is trivial. Suppose that $C_H G$ is nontrivial. As the center of G is trivial this implies that H has a subgroup $G \times \mathbb{Z}_n$, for some $n > 1$. It is a consequence of the equivariant Dehn Lemma/Loop Theorem that the quotient $\bar{V} = V/\mathbb{Z}_n$ is again a handlebody (see [21], Proposition 1). The action of G on V projects to an action of G on $\bar{V}$. However, the genus of $\bar{V}$ is strictly smaller than the genus g of V (because $g > 1$; one may apply, for example, the formula of Riemann-Hurwitz to the induced branched covering between the boundaries of V and $\bar{V}$). This is a contradiction because, by hypothesis, g was the least genus for a G-action on a handlebody.

3. Proof of Proposition (1.2)

The maximal handlebody groups are exactly the finite admissible quotients of the four free products with amalgamation associated to the four admissible graphs of groups of maximal negative Euler characteristic $-1/12$, and the maximal bounded surface groups are the quotients of the first group $\mathbb{D}_2 *_{\mathbb{Z}_2} \mathbb{D}_3$ (isomorphic to the extended modular group). Each of these four free products with amalgamation is a quotient, with torsionfree kernel, of the (quadrilateral) Fuchsian group with signature $(0; 2, 2, 2, 3)$ and presentation

$$\langle x_1, x_2, x_3, x_4 \mid x_1^2 = x_2^2 = x_3^2 = x_4^3 = x_1x_2x_3x_4 = 1 \rangle$$

(the image of x_1x_2 generates the amalgamated subgroups). This Fuchsian group is generated by the two elements x_1x_2 and x_2x_3 (noting that $(x_1x_2)(x_3x_2)(x_2x_1)(x_2x_3) = (x_1x_2x_3)^2 = x_4^{-2} = x_4$ and $(x_1x_2)(x_3x_2) = x_4^{-1}x_2$). This quadrangle group is in fact one of the exceptional Fuchsian groups for which the geometric and the algebraic rank do not coincide, see [18], chapter 4.16. Noting that the image of x_1x_2 in $\mathbb{D}_2 *_{\mathbb{Z}_2} \mathbb{D}_3$ has order two, this proves that every maximal handlebody group is 2-generated, and that every maximal bounded surface group is $(2,n)$-generated. Suppose that G is a maximal bounded surface group which is simple (or perfect). The modular group $\mathbb{Z}_2 * \mathbb{Z}_3$ is a subgroup of index two in the extended modular group $\mathbb{D}_2 *_{\mathbb{Z}_2} \mathbb{D}_3$. Then G which is a surjective image of the extended modular group is also a surjective image of the modular group, and hence $(2,3)$-generated.

An example of a maximal handlebody group which is not $(2,n)$-generated is the central product (the direct product with identified centers) $\mathbb{A}_5^* \times_{\mathbb{Z}_2} \mathbb{A}_5^*$ of two binary dodecahedral groups $\mathbb{A}_5^*$, isomorphic to the orientation-preserving symmetry group of the 4-dimensional 120-cell or to the orientation-preserving subgroup of the Coxeter group [3,3,5]. Its quotient by the diagonal subgroup (isomorphic to $\mathbb{A}_5$) is the binary dodecahedral group $\mathbb{A}_5^*$ which has a unique (central) involution and hence is not $(2,n)$-generated. Then also $\mathbb{A}_5^* \times_{\mathbb{Z}_2} \mathbb{A}_5^*$ is not $(2,n)$-generated; in particular, it is not a maximal bounded suface group. On the other hand, $\mathbb{A}_5^* \times_{\mathbb{Z}_2} \mathbb{A}_5^*$ is a maximal handlebody group because there is an admissible surjection $\phi : \mathbb{D}_5 *_{\mathbb{Z}_5} \mathbb{A}_5 \to \mathbb{A}_5^* \times_{\mathbb{Z}_2} \mathbb{A}_5^*$, by mapping $\mathbb{A}_5$ to the diagonal subgroup of $\mathbb{A}_5^* \times_{\mathbb{Z}_2} \mathbb{A}_5^*$, and an involution in $\mathbb{D}_5$ to a suitable element of the form (a, b) such that $a^2 = b^2$ is the central element and ab^{-1} has order ten in $\mathbb{A}_5^*$ (so a and b generate a binary dihedral subgroup $\mathbb{D}_5^*$ of $\mathbb{A}_5^*$). The quotients of $\mathbb{D}_5 *_{\mathbb{Z}_5} \mathbb{A}_5$ by characteristic subgroups of the kernel of ϕ (a free group) give infinitely many maximal handlebody groups which are not maximal bounded surface groups (see also section 2 of [22]).

4. Proof of Proposition (1.3)

Let G be a finite group of isometries of maximal possible order $12(g-1)$ of a connected orientable hyperbolic surface F of genus $g > 1$ which extends to a handlebody of genus g. By [19], the G-action on F is of type $(2, 2, 2, 3)$, that is the quotient orbifold F/G has signature $(0; 2, 2, 2, 3)$, i.e. is the 2-sphere with four branch points of orders 2,2,2 and 3. The group of all lifts of elements of G to the universal covering of F, the hyperbolic plane, is a Fuchsian group $\tilde{G} = (2, 2, 2, 3)$ of signature $(0; 2, 2, 2, 3)$ and characteristic $-1/6$ (see [18]); this is a subgroup of finite index of the Fuchsian group $\tilde{H}$ consisting of all lifts of elements of the orientation preserving isometry group H of F.

We can assume that $\tilde{H}$ is different from $\tilde{G}$. By the formula of Riemann-Hurwitz, $\tilde{H}$ is a triangle group (p, q, r) (i.e., of signature $(0; p, q, r)$), of characteristic $-1/12$, $-1/24$, $-1/18$ or $-1/42$ (there are no triangle groups of characteristics $-1/30$ and $-1/36$). The triangle groups of characteristic $-1/12$ are $(2, 4, 6)$, $(3, 3, 4)$ and $(2, 3, 12)$, but only the first one has the quadrangle group $(2, 2, 2, 3)$ as a subgroup of index two. The only triangle group of charateristic $-1/18$ is $(2, 3, 9)$, containing $(2, 2, 2, 3)$ as a subgroup, necessarily normal, of index three (see [14] for the determination of the signature of a subgroup of a Fuchsian group). The only triangle group of characteristic $-1/24$ is $(2, 3, 8)$ which does not have $(2, 2, 2, 3)$ as a subgroup. The only remaining possibility is the triangle group $(2, 3, 7)$ which contains $(2, 2, 2, 3)$ as a non-normal subgroup of index seven; in this case H is a Hurwitz group of order $84(g-1)$ containing G as a subgroup of index seven.

Finally, suppose that H is a non-abelian simple Hurwitz group of order $84(g-1)$. The subgroup G of index seven in H gives a homomorphism from H to the symmetric group $\mathbb{S}_7$ (by permutation of the left cosets of G in H induced by left multipication). As H is simple the kernel of this homomorphism is trivial and H is isomorphic to a subgroup of S_7. The simple Hurwitz groups of order at most 7! are listed in [3] and are of linear fractional type $\mathrm{PSL}_2(7)$, $\mathrm{PSL}_2(8)$ or $\mathrm{PSL}_2(13)$; the only one among these groups with a subgroup of index seven is $\mathrm{PSL}_2(7)$, acting on Klein's quartic of genus 3, with a subgroup S_4 (see [15], p. 415).

5. Proof of Proposition (1.4)

Part a) of the Proposition follows from [11], Theorem 3.1 and Corollary 3.3 (we note that in the case $k = 5$ of Theorem 3.1 and in Corollary 3.3 of [11] the condition $q \neq 11$ has to be added).

The only simple group which is a maximal handlebody group of genus 0 is the alternating group $\mathbb{A}_5 \cong \mathrm{PSL}(2,4) \cong \mathrm{PSL}(2,5)$. A maximal handlebody group G of genus $g > 1$ has order $12(g-1)$. It follows that the handlebody genus g of G is even if and only if the Sylow 2-subgroups of G have order 4. By [16], p. 582, Theorem 11.1, the finite simple groups with abelian Sylow 2-subgroups of order 4 are the linear fractional group $\mathrm{PSL}(2,q)$, for a prime power $q \equiv \pm 3 \mod 8$. Part b) of Proposition (1.4) follows now from part a).

6. Proof of Proposition (1.5)

The first part of the Proposition follows from the fact that a finite nilpotent group G is not an admissible quotient of one of the four extremal products with amalgamation $\mathbb{D}_2 *_{\mathbb{Z}_2} \mathbb{D}_3$, $\mathbb{D}_3 *_{\mathbb{Z}_3} \mathbb{A}_4$, $\mathbb{D}_4 *_{\mathbb{Z}_4} \mathbb{S}_4$ and $\mathbb{D}_5 *_{\mathbb{Z}_5} \mathbb{A}_5$ in section 2 (because the factors $\mathbb{D}_3$, $\mathbb{A}_4$, $\mathbb{S}_4$ and $\mathbb{A}_5$ are not nilpotent). This implies that the maximal order $12(g-1)$ is not achieved for nilpotent groups of diffeomorphisms of handlebodies and bounded surfaces of (algebraic) genus $g > 1$.

The groups G of the second largest order $8(g-1)$ are exactly the admissible quotients of one of the two groups $\mathbb{D}_2 *_{\mathbb{Z}_2} \mathbb{D}_4$ and $\mathbb{D}_3 *_{\mathbb{Z}_3} \mathbb{S}_4$, of Euler characteristic $-1/8$ (see section 2). Note that the second group does not occur for nilpotent groups G. The nilpotent 2-group $\mathbb{D}_4$ is an admissible quotient, with kernel the free group F_2 of rank two, of the first group $\mathbb{D}_2 *_{\mathbb{Z}_2} \mathbb{D}_4$. It follows that $8(g-1)$ is an upper bound for nilpotent groups of diffeomorphisms; moreover this upper bound is attained for infinitely many different genera, by considering the quotients of $\mathbb{D}_2 *_{\mathbb{Z}_2} \mathbb{D}_4$ by characteristic subgroups of F_2 whose index is a power of two (so the quotient is still a 2-group and hence nilpotent).

Finally, suppose that the nilpotent group G is an admissible quotient of $\mathbb{D}_2 *_{\mathbb{Z}_2} \mathbb{D}_4$. Then G is generated by the 2-groups $\mathbb{D}_2$ and $\mathbb{D}_4$ which are contained in the unique normal Sylow 2-subgroup of G (a nilpotent group is the direct sum of its Sylow subgroups). Hence G coincides with its Sylow 2-subgroup and is a 2-group.

Remark. By [23], the maximal possible order of a finite nilpotent group G of orientation-preserving diffeomorphisms of a closed orientable surface of genus $g > 1$ is $16(g-1)$. This can be seen also as follows. A nilpotent group is the direct sum of its Sylow subgroups. It follows that G is not an admissible quotient of a triangle group of type $(2,3,m)$, for $m > 6$ (because the product of two elements of orders two and three in G has order six). For a similar reason, G is not an admissible quotient of the triangle groups $(2,4,5)$, $(2,4,6)$, $(2,4,7)$ and $(2,5,5)$. Excluding all these triangle groups, the largest order remaining is $16(g-1)$ and realized by the triangle group $(2,4,8)$. In this maximal case, G has to be again a 2-group because it is generated by elements of orders 2 and 4 and hence coincides with its Sylow 2-subgroup.

Received August 07, 2003

Final version received January 26, 2004

Università degli Studi di Trieste
Dipartimento di Matematica e Informatica
34100 Trieste

Italy
zimmer@units.it

References

[1] M. Aschbacher, R. Guralnick, *Some applications of the first cohomology group*, J. Algebra **90** (1984), 446–460.

[2] S.A. Broughton, *Simple group actions on hyperbolic Riemann surfaces of least area*, Pacific J. Math. **158** (1993), 23–48.

[3] M. Conder, *The genus of compact Riemann surfaces with maximal automorphism group*, J. Algebra **108** (1987), 204–247.

[4] J.H. Conway, R.T. Curtis, S.P. Norton, R.A. Parker, R.A. Wilson, Atlas of Finite Groups. Oxford University Press 1985.

[5] C.L. May, *Automorphisms of compact Klein surfaces with boundary*, Pacific J. Math. **59** (1975), 199–210.

[6] C.L. May, *Real genus actions of finite simple groups*, Rocky Mountain J. Math. **31** (2001), 539–551.

[7] C.L. May, *Nilpotent automorphism groups of bordered Klein surfaces*, Proc. Amer. Math. Soc. **101** (1987), 287–292.

[8] D. McCullough, A. Miller, B. Zimmermann, *Group actions on handlebodies*, Proc. London Math. Soc. **59** (1989), 373-415.

[9] D. McCullough, A. Miller, B. Zimmermann, *Group actions on nonclosed 2-manifolds*, J. Pure Appl. Algebra **64** (1990), 269–292.

[10] A. Miller, B. Zimmermann, *Large groups of symmetries of handlebodies*, Proc. Amer. Math. Soc. **106** (1989), 829–838.

[11] L. Paoluzzi, B. Zimmermann, *On finite quotients of the Picard group and related hyperbolic tetrahedral groups*, Rend. Istit. Mat. Univ. Trieste Suppl. 1 **XXXII** (2001), 257–288.

[12] M. Reni, B. Zimmermann, *Handlebody orbifolds and Schottky uniformizations of hyperbolic 2-orbifolds*, Proc. Amer. Math. Soc. **123** (1995), 3907–3914.

[13] C. Sims, *Computational methods in the study of permutation groups*, Computational Problems in Abstract Algebra (Proc. Conf. Oxford 1967, ed. J. Leech), Pergamon, Oxford 1970, 169–173.

[14] D. Singerman, *Subgroups of Fuchsian groups and finite permutation groups*, Bull. London Math. Soc. **2** (1970), 319–323.

[15] M. Suzuki, Group Theory I, Grundlehren der mathematischen Wissenschaften 247, Springer-Verlag 1986.

[16] M. Suzuki, Group Theory II, Grundlehren der mathematischen Wissenschaften 248, Springer-Verlag 1986.

[17] A. J. Woldar, *Genus actions of finite simple groups*, Illinois J. Math. **33** (1989), 438–450.

[18] H. Zieschang, E.Vogt, H.-D.Coldewey, Surfaces and planar discontinuous groups. Lecture Notes in Mathematics 835, Springer, Berlin 1980.

[19] B. Zimmermann, *Über Abbildungsklassen von Henkelkörpern*. Arch. Math. **33** (1979), 379–382.

[20] B. Zimmermann, *Über Homöomorphismen n-dimensionaler Henkelkörper und endliche Erweiterungen von Schottky-Gruppen*, Comm. Math. Helv. **56** (1981), 474–486.

[21] B. Zimmermann, *Finite orientation reversing group actions on handlebodies and 3-manifolds*, Rend. Circ. Mat. Palermo **48** (1999), 549–562.

[22] B. Zimmermann, *Finite group actions on handlebodies and equivariant Heegaard genus for 3-manifolds*, Topology Appl. **43** (1992) 263–274.

[23] R. Zomorrodian, *Nilpotent automorphism groups of Riemann surfaces*, Trans. Amer. Math. Soc. **288** (1985), 241–255.

Bol. Soc. Mat. Mexicana (3) Vol. 10, Special issue, 2004

SOME RESULTS ON ONE-RELATOR SURFACE GROUPS: ERRATUM

JAMES HOWIE

In [1], a number of results were proved about groups of the form $\pi_1(S)/N(R)$, where S is a closed orientable surface, and $N(R)$ is the normal closure of a single element $R \in \pi_1(S)$.

Unfortunately, one of the results of that paper, Theorem 4.1 (the Freiheitssatz), is not true in the full generality stated there. The purpose of the present note is to explain the gap in the proof, give some examples where the statement of the theorem is false, and to present a conjecture regarding the additional hypotheses necessary for the Freiheitssatz to hold.

The error in the proof of [1], Theorem 4.1, does not affect any of the other results stated in [1].

Elements of $\pi_1(S)$ can be represented as closed curves α in the surface S (usually with self-intersections). Any closed curve in S defines a homology class in $H_1(S)$. Given two closed curves α, β in S, we let $\langle\alpha, \beta\rangle \in \mathbb{Z}$ denote the value of the intersection form on $H_1(S) \times H_1(S)$, applied to the pair of homology classes represented by α and β.

If $R \in \pi_1(S)$ is represented by the closed curve α, then we also use the notation $\pi_1(S)/\alpha$ for $\pi_1(S)/N(R)$.

Given an *embedded* (that is, *simple*) closed curve β in S, I falsely stated in [1], Theorem 4.1, the following assertion, intended as the analogue of Magnus' Freiheitssatz for one-relator groups [3]:

ASSERTION (1). *Let S be a closed oriented surface, α a closed curve in S, and β a simple closed curve in S such that α is not homotopic to a curve disjoint from β. Then $\pi_1(S \setminus \beta) \to \pi_1(S)/\alpha$ is injective.*

The alleged proof of Assertion 1 in [1] is wrong. It seeks to apply Magnus' Freiheitssatz [3] to show that a map $\pi_1(F) \to \pi_1(F')/\gamma$ is injective, where F is a certain non-compact surface, F' is obtained from F by adjoining an annulus to a pair of circle components of ∂F, and γ is a certain closed curve in F'. This argument is not valid, for the following reason. Although $\pi_1(F)$ and $\pi_1(F')$ are both free, and the inclusion-induced homomorphism $\pi_1(F) \to \pi_1(F')$ is injective, the image of this homomorphism is not a free factor of $\pi_1(F')$, so Magnus' result does not apply.

Indeed there are some quite simple counterexamples to Assertion 1, as follows:

Example (2). Let β be a non-separating, embedded closed curve in the closed orientable surface S, and let α be any closed curve that intersects β transversely in a single point. Then

$$1 \neq [\alpha, \beta] \in \pi_1(S \setminus \beta) \cap N(\alpha).$$

Example (3). Let α, β be essential simple closed curves contained in a punctured torus $T \subset S$, such that $\langle \alpha, \beta \rangle \neq 0$. Then

$$1 \neq \partial T \in \pi_1(S \smallsetminus \beta) \cap N(\alpha).$$

These examples share the following common feature: in each case, the *geometric intersection number* $|\alpha \cap \beta|$ (that is, the minimum number of points of intersection of β with any curve isotopic to α) is equal to the absolute value of $\langle \alpha, \beta \rangle$. In other words, α and β can be isotoped so as to intersect transversely in finitely many points, all of the same sign. This suggests the following

Conjecture (4). Let α, β, S be as in Assertion 1. Suppose that $|\alpha \cap \beta|$ is strictly greater than the absolute value of $\langle \alpha, \beta \rangle$. Then $\pi_1(S \smallsetminus \beta) \to \pi_1(S)/\alpha$ is injective.

There is some evidence in favour of Conjecture 4. Firstly, in [1], a special case of Assertion 1 is proved, using a much simpler (and, more importantly, correct) argument:

THEOREM (5) ([1], Proposition 3.10). *Let S be a closed oriented surface, α a closed curve in S, and β a simple closed curve in S such that α is not homotopic to a curve disjoint from β, and that $\langle \alpha, \beta \rangle = 0$. Then $\pi_1(S \setminus \beta) \to \pi_1(S)/\alpha$ is injective.*

The proof of this result uses the fact that the infinite cyclic cover of S corresponding to the kernel of the homomorphism $\langle -, \beta \rangle : \pi_1(S) \to \mathbb{Z}$ can be constructed from a collection F_n, $n \in \mathbb{Z}$ of copies of $F := S \smallsetminus \beta$, with $F_n \cap F_{n+1}$ an annulus. See [1], Proposition 2.1.

Combining this fact with a result of Klyachko [2], we can also prove the following special case of Conjecture 4:

THEOREM (6). *Let S be a closed oriented surface, α a closed curve in S, and β a simple closed curve in S such that α is not homotopic to a curve that meets β at most once, and that $\langle \alpha, \beta \rangle = \pm 1$. Then $\pi_1(S \smallsetminus \beta) \to \pi_1(S)/\alpha$ is injective.*

The methods of [1] can also be used to prove the following:

THEOREM (7). *Let S be a closed oriented surface, α a closed curve in S, and β_1, β_2 two disjoint simple closed curves in S such that α is not homotopic to a curve disjoint from β_1 or from β_2. Then $\pi_1(S \smallsetminus (\beta_1 \cup \beta_2)) \to \pi_1(S)/\alpha$ is injective.*

The proofs of the last two results will appear in a separate article.

DEPARTMENT OF MATHEMATICS
HERIOT-WATT UNIVERSITY
EDINBURGH EH14 4AS
SCOTLAND
Email: J.Howie@hw.ac.uk

REFERENCES

[1] J. Howie, Some results on one-relator surface groups, *Bol. Soc. Mat. Mexicana* (3) **10**, Special issue, (2004).

[2] A. A. Klyachko, A funny property of sphere and equations over groups, *Comm. Algebra* **21** (1993), 2555–2575.

[3] W. Magnus, Über diskontinuierliche Gruppen mit einer definierenden Relation (Der Freiheitssatz), *J. reine angew. Math.* **163** (1930), 141–165.