DERIVATIVE VARIETIES AND THE PURE BRAID GROUP

By D. Cooper and D. D. Long[1]

Introduction. In the light of the example of [16], which shows there is a braid in the kernel of the Burau representation, there has been renewed interest in the Gassner representation, both from the point of view of faithfulness and in examining the interactions between these two representations. In this paper, we shall show that the approach of [11] highlights the differences between these representations; the upshot being that in the cases $n = 3, 4$ we might expect to see different behavior in the Burau representation.

Briefly, the strategy is that one can regard a braid σ as both an algebraic morphism acting on the representation variety of a free group and as a link in S^3 by taking the closure $\hat{\sigma}$. These ideas interact as the fixed points of the diffeomorphism are precisely the representations of the fundamental group of the complement of $\hat{\sigma}$. One can then ask questions about deformations of representations and the Zariski tangent spaces at certain representations in the algebraic set $\mathrm{Fix}(\sigma)$. This makes clear one of the difficulties in the problem since the Burau and Gassner representations arise in this context from the abelian representations.

We use this information in two ways. Let X be the representation variety of the free group in to $SL(2, \mathbf{C})$. We can use the fact that X is an algebraic group to use left translation to define a generalized derivative $\xi_\sigma(\rho)$ for the morphism σ at the representation ρ. (This is the usual derivative in the case that ρ happens to be a fixed point for σ.) We then define the Derivative variety for σ, denoted $B(\sigma)$, to be the Zariski closure of $\xi_\sigma(X)$. The question of faithfulness can then be rephrased by asking about which representations map to the identity in the derivative variety. This has the advantage that we are considering questions which include irreducible representations of the link complement and that we can deal with both Burau and Gassner representations simultaneously.

Standard analysis, see [17], of the fibers of the morphism $\xi_\sigma : X \to B(\sigma)$ shows that generic point preimages have dimension $\dim_{\mathbf{C}}(X) - \dim_{\mathbf{C}}(B(\sigma))$ so that if the dimension of $B(\sigma)$ is large then we would expect point preimages to

Manuscript received October 19, 1990.
[1]Supported in part by grants from the Alfred P. Sloan Foundation and NSF.
American Journal of Mathematics 115 (1993), 137–160.

have dimension less than the dimension of the abelian representations. As a first step in this direction we show:

THEOREM 2.17. *Suppose that σ is a pure braid with hyperbolic closure and let R be the component of the representation variety of $\pi_1(S^3 \setminus \hat{\sigma})$ which contains the discrete faithful representation. Then the Zariski closure of $\xi_\sigma(R)$ has dimension $n+3$.*

The proof of this fact will occupy §2 and involves Thurston's deformation theory together with some special facts about braids. We have the corollaries:

COROLLARY 2.18. *If σ is a pure braid with hyperbolic closure, then:*

(a) $\dim_{\mathbf{C}}(B(\sigma)) \geq n+3$.

(b) For $\rho \in U$, a Zariski open subset of $B(\sigma)$, we have $\xi_\sigma^{-1}(\rho)$ has dimension $2n-3$.

This is then combined with the results of §3, where we do some analysis at the soluble representations. Letting β denote the Gassner representation, and X_S the soluble representations of the free group of rank n, we show:

THEOREM 3.12. *Suppose that $\sigma \in \gamma_{3n}(\ker(\beta))$. Then $\xi_\sigma^{-1}(I)$ contains a component in its representation variety which contains X_S, in particular, it has dimension at least* $2n+1$.

This result is to be compared with the well known fact (see Theorem 3.16 of [2] & Theorem 3.4 below) that any braid in $\ker(\beta)$ acts as the identity map on F_n/F_n''.

Since $2n+1$ is greater than the dimension $2n-3$ predicted by 2.18, this suggests that a braid in the kernel of β would be unusual. This is to be contrasted with the analogous analysis for the Burau representation; one finds that the best estimate for the dimension of $\xi_\sigma^{-1}(I)$ is $n+2$, and this is only greater than $2n-3$ for $n < 5$; suggesting that for the braid group the cases $n = 3, 4$ may be special. An analysis of the soluble representations also shows:

THEOREM 3.10. *Suppose that $\sigma \in \ker(\beta)$. Then the longitude on each torus lies in $\pi_1(S^3 \setminus \hat{\sigma})''$.*

A further geometric property comes from a somewhat different direction. Note that the trivial braid has the property that the fundamental group of its complement splits over the trivial group. In §4 we use ideas of Hatcher as well as the Culler-Shalen machine to show:

THEOREM 4.8. *Let σ be an element of P_4 which lies in* $\ker(\beta)$. *Then the complement of the $\hat{\sigma}$ contains a closed embedded nonboundary parallel incompressible surface; in particular, its fundamental group splits over a closed surface group.*

In the course of proving this, we observe the following, which are perhaps of some independent interest:

THEOREM 4.1. *Let L be an n component link. Suppose that the representation variety of L has a component A which contains an irreducible representation and has dimension $> n+3$. Then $S^3 \setminus L$ contains a nonboundary parallel closed embedded incompressible surface.*

This has various corollaries in the special cases when one can guarantee the hypotheses of the theorem. For example:

COROLLARY 4.6. *A homology boundary link of two or more components contains in its complement a closed embedded nonperipheral incompressible surface.*

The authors thank C. Hodgson for useful conversations.

1. Preliminaries. In this section we collect some of the basic notions. We appeal to some elementary algebraic geometry, so we collect here a few facts we shall need, this serving the additional purpose of establishing notation. Full details may be found in [17] or [8].

Let $A^n(\mathbf{C})$ be complex space of dimension n; in this context, usually referred to as *affine n-space*. Given an ideal I in the polynomial ring with n indeterminates $\mathbf{C}[X_1,\dots,X_n]$ we define the *affine algebraic set* $V(I)$ to be $\{x \in A^n(\mathbf{C}) \mid g(x) = 0 \;\; g \in I\}$. Conversely, given a set S in $A^n(\mathbf{C})$, we define an ideal in $\mathbf{C}[X_1,\dots,X_n]$ by setting $I(S) = \{f \mid f(x) = 0 \;\; x \in S\}$.

An ideal I is *radical* if $f^k \in I$ for some k implies $f \in I$ and an ideal is *prime* if $g.f \in I$ implies g or $f \in I$. Then Hilbert's Nullstellensatz states that the maps $I \to V(I)$ and $V \to I(V)$ establishes a bijection between radical ideals in $\mathbf{C}[X_1,\dots,X_n]$ and affine algebraic sets in $A^n(\mathbf{C})$. If I is a prime ideal, we say that $V(I)$ is *irreducible* or that $V(I)$ is an *affine algebraic variety*. These are the building blocks in the sense that a radical ideal can be written $I = P_1 \cap \cdots \cap P_r$ as a finite irredundant intersection of prime ideals in just one way, with associated decomposition $V(I) = V(P_1) \cup \cdots \cup V(P_r)$ into algebraic varieties, the *components* of $V(I)$. The sets $V(I)$ form a basis as closed sets of a topology, the *Zariski topology* on $A^n(\mathbf{C})$.

We shall also have to briefly refer to projective varieties. Recall that one defines *projective space* $\mathbf{P}^n$ to be $(A^{n+1}(\mathbf{C}) \setminus 0)/\sim$, where the equivalence relation $\sim$ is given by requiring that equivalence classes are $\{\lambda(x_0,\dots,x_n) \mid \lambda \in \mathbf{C}\}$. The theory is then set up the same way as before, save that we restrict attention to those ideals of $\mathbf{C}[X_1, \ldots, X_{n+1}]$ which are *homogeneous*, that is to say can be generated by homogeneous polynomials.

A map $f : X \to Y$ is a *morphism* if it is the restriction of a polynomial map between the ambient spaces. Morphisms are continuous in the Zariski topology. A morphism is *dominating* if $f(X)$ is Zariski dense in Y. This is not an unduly restrictive notion, as one may always consider the morphism to have target the Zariski closure of $f(X)$.

There is a notion of the *dimension* of a variety; this can be defined in several equivalent ways. Given a variety V we define $\mathbf{C}(V)$ to be the field of fractions of the integral domain $\mathbf{C}[X_1, \ldots, X_n]/I(V)$. Then $\dim_{\mathbf{C}}(X)$ is the transcendence degree $\mathbf{C}(V)$ over $\mathbf{C}$. For us a crucial fact will be

THEOREM 1.1. (See [17] Theorem 3.13)

(a) Suppose that $f : X \to Y$ is a morphism. Then for each $y \in Y$, every component of $f^{-1}(y)$ has dimension as least $\dim_{\mathbf{C}}(X) - \dim_{\mathbf{C}}(Y)$.

(b) On a Zariski open subset of Y, every component of $f^{-1}(y)$ has dimension exactly $\dim_{\mathbf{C}}(X) - \dim_{\mathbf{C}}(Y)$.

Let F_n be the free group of rank n. This has automorphism group $\mathrm{Aut}(F_n)$. Our main interest here will be in the subgroups B_n and P_n: the *braid group* and the *pure braid group*. Full information about these groups and their properties can be found in [2]; we briefly recall the main facts to which we shall appeal.

The n-string *braid group*, B_n we define to be the subgroup of $\mathrm{Aut}(F_n)$ generated by the automorphisms $\{\sigma_i \mid 1 \leq i \leq n-1\}$ where the action of σ_i is given by:

$$\begin{aligned} &x_i \to x_{i+1} \\ &x_{i+1} \to (x_{i+1})^{-1} x_i x_{i+1} \\ &x_j \to x_j \quad j \neq i, i+1 \end{aligned}$$

If we use the symbol Σ_n to denote the symmetric group on n letters, then there is an obvious map from $B_n \to \Sigma_n$ coming from action as a permutation group on the set $\{x_1, x_2, \ldots, x_n\}$. We now define the *Pure Braid group*, denoted P_n, to be the kernel of this homomorphism.

These ideas interact by the use of the representation variety of the free group of rank n into $SL(2, \mathbf{C})$. We set $X = \{\rho \mid \rho : F_n \to SL(2, \mathbf{C})\}$.

Clearly X is an irreducible affine algebraic set and choice of a basis for the free group gives an identification $X = \prod_{i=1}^n SL(2, \mathbf{C})$ a direct product of n copies of $SL(2, \mathbf{C})$. This shows that X has dimension $3n$. Those points of X which correspond to isomorphisms with F_n we refer to as *generic*; see [11] for a justification of this term.

We also pick out the set X_A of abelian representations; one checks easily that this has dimension $n+2$ and the set X_S of soluble representations which has dimension $2n+1$. These are both subvarieties of X.

We shall concentrate on the action of the group $\mathrm{Aut}(F_n)$, which acts on X as a collection of algebraic morphisms by the rule that if $\sigma \in \mathrm{Aut}(F_n)$ and $\rho \in X$ then $\sigma\rho$ is the representation defined by $\sigma\rho(w) = \rho(\sigma^{-1}w)$ (cf. [11]). This is used to give linear representations via:

THEOREM 1.2. [11] *Suppose that the representation $\alpha \in X$ is fixed by a subgroup H of $Aut(F_n)$. Then there is a linear representation of H defined by $h \to dh_\alpha$.*

There are two standard cases of this in the context of the braid groups; we shall be interested in the case that H is the pure braid group and the fixed representations are allowed to run over X_A. This gives a representation $\beta : P_n \to GL(\Lambda)$ usually known as the *Gassner representation.* Here Λ is the ring of Laurant polynomials over $\mathbf{Z}$ in n commuting variables. One also obtains the Burau representation by considering the subvariety of X_A given by considering representations with all generators mapping to the same element and linearizing.

We also have cause to refer to the standard representation of $SL(2,\mathbf{C})$ on its Lie algebra which is usually denoted $Ad : SL(2,\mathbf{C}) \to \mathrm{Aut}(SL(2,\mathbf{C}))$. If we let $I_B : SL(2,\mathbf{C}) \to SL(2,\mathbf{C})$ be the smooth map defined by $A \to BAB^{-1}$ we can define Ad_B as the derivative of I_B at the $\{e\}$. Our conventions imply that the following diagram commutes:

$$\begin{array}{ccc} T_e(SL(2,\mathbf{C})) & \longrightarrow & T_A(SL(2,\mathbf{C})) \\ \downarrow Ad_B & & \downarrow I_B \\ T_e(SL(2,\mathbf{C})) & \longrightarrow & T_{BAB^{-1}}(SL(2,\mathbf{C})) \end{array}$$

2. Derivative varieties. In this section we introduce the notion of derivative varieties in 2.3; this is the image of X under the generalized derivative map. The main result is 2.17 where it is shown that $n+3$ is a lower bound on the dimension of the derivative variety of any hyperbolic braid. In §3 we will use this result to compare the dimension of point preimages with the ambient dimension.

We begin with some preliminaries. Note that in our context X is an algebraic group defined over the complex numbers. For $\rho \in X$, we define the left translation map $L_\rho : X \to X$ by $L_\rho(\alpha) = \alpha.\rho$. Then given a diffeomorphism $f : X \to X$ we may use the derivative of this map to translate the map

$$T_\rho(X) \xrightarrow{df_\rho} T_{f(\rho)}(X)$$

and obtain a canonical map at the identity which we denote by $\xi(df_\rho)$ defined by the following diagram, where e denotes the identity in the Lie group X:

$$\begin{array}{ccc} T_\rho(X) & \xrightarrow{df_\rho} & T_{f(\rho)}(X) \\ \downarrow dL_\rho^{-1} & & \downarrow dL_{f(\rho)}^{-1} \\ T_e(X) & \xrightarrow{\xi(df_\rho)} & T_e(X) \end{array}$$

We shall suppress the Lie algebra notation for $T_e(X)$ as the maps in which we will usually be interested will not preserve the Lie algebra structure. The map $\xi(df_\rho)$ will be called the *generalized derivative* of f at ρ. In the case that f actually fixes ρ this gives the usual derivative in the sense of smooth topology.

We shall actually do better than the lemma which follows (see Theorem 2.8); but it is worth observing that there is much information in the maps $\xi(df_\rho)$ even for a diffeomorphism:

LEMMA 2.1. *Suppose that $f : X \to X$ is a diffeomorphism with the property that for all $\rho \in X$ we have that $\xi(df_\rho) = Id$ and $f(e) = e$. Then there is a neighborhood U of $\{e\}$ on which f is the identity map.*

Proof. Given a vector $v \in T_e(X)$ we define the left invariant field by the rule that $v_\rho = dL_\rho(v)$. For a small neighborhood V of $\mathbf{0} \in T_e(X)$ it is a standard result that the flow lines of these vector fields define a diffeomorphism $\exp : V \to U$, where U is a small neighborhood of the identity in the group X. For f as in the statement of the lemma we have that $df_\rho(v_\rho) = df_\rho dL_\rho(v) = dL_{f(\rho)}(v) = v_{f(\rho)}$ so that f preserves left invariant fields. If now x is any point very close to the identity we may find a vector v so that the flow line through $\{e\}$ of the field defined by v runs through x. Denote this flow line by $\sigma : [0,1] \to X$, where $\sigma(0) = e$ and $\sigma(1) = x$. It satisfies the differential equation $d\sigma/dt = v \cdot \sigma(t)$ with initial condition $\sigma(0) = e$. Applying the chain rule, we see that $f\sigma(t)$ also satisfies this equation and initial condition, so by uniqueness we have that $f\sigma(t) = \sigma(t)$, so that $f(x) = x$. □

COROLLARY 2.2. *Suppose that $f \in Aut(F_n)$ satisfies the hypotheses of 2.1. Then f is the identity automorphism.*

Proof. It is shown in [11] that the only f which is the identity on a neighborhood of $\{e\}$ is the identity automorphism. □

Definition 2.3. Suppose that $\sigma \in \mathrm{Aut}(F_n)$. Then we define a map $\xi_\sigma : X \to \mathrm{Aut}(T_e(X))$ by $\xi_\sigma(\rho) = \xi(d\sigma_\rho)$.

The Zariski closure of the image of ξ_σ we call the *Derivative variety* of σ, denoting this by $B(\sigma)$.

One checks easily that ξ_σ is a morphism and that $B(\sigma)$ is irreducible. Notice that an elementary application of the chain rule shows that the image does indeed lie inside $\mathrm{Aut}(T_e(X))$. One interpretation of Lemma 2.1 is that $B(\sigma)$ is a single point only in the case that $\sigma = \mathrm{Id}$.

Example 2.4. Consider the automorphism $\sigma : F_2 \to F_2$ given by $\sigma(x) = xyx$ and $\sigma(y) = xy$. Using this basis to identify X with $SL(2,\mathbf{C}) \times SL(2,\mathbf{C})$ we see that if $\rho = (M_1, M_2)$ we have the block matrix:

$$\xi_\sigma(\rho) = \begin{bmatrix} I + Ad_{M_1M_2} & Ad_{M_1} \\ I & Ad_{M_1} \end{bmatrix}$$

Remark 2.5. In fact, knowledge of $\xi_\sigma(\rho)$ for all $\rho \in X$ determines σ when $\sigma \in \mathrm{Aut}(F_n)$. The proof is similar to that of 2.1, so we only sketch it. Fix some point ρ; we wish to determine $\sigma(\rho)$. Let $\beta : [0,1] \to X$ be a smooth path with $\beta(0) = e$ and $\beta(1) = \rho$. Define the vectors $\{v_t\}$ by the rule

$$\left.\frac{d\beta}{dt}\right|_{t=t_0} = dL_{\beta(t_0)}(v_{t_0}).$$

Then applying the chain rule to the path $\sigma\beta$ we see that

$$\left.\frac{d\sigma\beta}{dt}\right|_{t=t_0} = d\sigma_{\beta(t_0)} dL_{\beta(t_0)}(v_{t_0}) = dL_{\sigma\beta(t_0)}\xi\left(d\sigma_{\beta(t_0)}\right)(v_{t_0}) = dL_{\sigma\beta(t_0)}\xi_\sigma(\beta(t_0))(v_{t_0}).$$

It follows that if we set $\mathbf{w}_{t_0} = \xi_\sigma(\beta(t_0))(v_{t_0})$; the hypothesis implies that all these vectors are determined once we know $\{v_t\}$. Then the path $\sigma\beta(t)$ satisfies the differential equation $\dot\gamma(t) = \mathbf{w}_t \cdot \gamma(t)$ with initial condition $\gamma(0) = e$. Again, by uniqueness it is the only such path and $\sigma(\rho)$ is recovered as $\sigma\beta(1)$.

A coarse invariant of $B(\sigma)$ is its dimension. Even this is of interest since as explained in §0, we would like to know that very few abelian representations map to the identity under the map ξ_σ. If we could show $\dim_{\mathbf{C}}(X) - \dim_{\mathbf{C}}(B(\sigma)) < \dim_{\mathbf{C}}(X_A)$ this would be true for generic points.

We are interested in analyzing the image of the map ξ_σ in the case that the automorphism σ lies inside the pure braid group. In this case we have extra information coming from certain invariance properties special to this group. The first of these comes from the map $\pi : X \to SL(2,\mathbf{C})$ given by $\pi(M_1, M_2, \dots, M_n) = M_1, \dots, M_n$. Then it is an elementary property [2] that $\sigma \in P_n$ implies that for all $\rho \in X$ we have $\pi\rho = \pi(\sigma\rho)$. Differentiating this condition gives:

$$d\pi_\rho = d\pi_{\sigma(\rho)} d\sigma_\rho$$

and translating this condition to the identity we see:

$$d\pi_\rho dL_\rho = d\pi_{\sigma(\rho)} dL_{\sigma(\rho)} \xi(d\sigma_\rho) \tag{1}$$

LEMMA 2.6. *If $\rho = (M_1, \ldots, M_n)$ then $d\pi_\rho dL_\rho$ has matrix given by*

$$(I, Ad_{M_1}, Ad_{M_1 M_2}, \ldots, Ad_{M_1 M_2 \cdots M_{n-1}})$$

Proof. Consider the tangent vector in the j-th component as coming from a small path v_t left translated to ρ that is, we seek to find the image of $(M_1, \ldots, v_t M_j, \ldots, M_n)$. This is given by $M_1 \ldots v_t M_j \ldots M_n$. Differentiating and translating back to the identity in the target, we see that $(\mathbf{0}, \mathbf{0}, \ldots, v_j, \ldots, \mathbf{0})$ maps to $Ad_{M_1 M_2 \cdots M_j}(v_j)$. This implies the result. □

This has the following consequence for the sequel:

COROLLARY 2.7. *Suppose that $\sigma \in P_n$. Then if $\xi(d\sigma_\rho) = I$, we have $\rho = \sigma(\rho)$ up to signs. In particular,*

(a) There is a K depending only on n for which $\rho = \sigma^K(\rho)$.

(b) If no M_j has trace zero then $\rho = \sigma(\rho)$.

In fact, we show something slightly stronger:

THEOREM 2.8. *Suppose that $\sigma \in P_n$. Then $\xi(d\sigma_\rho)$ and ρ determine $\sigma(\rho)$ up to signs.*

Proof. By equation (1) we have:

$$\begin{aligned}(I, Ad_{M_1}, Ad_{M_1 M_2}, \ldots, Ad_{M_1 M_2 \cdots M_{n-1}}) \xi(d\sigma_\rho)^{-1} \\ = (I, Ad_{\sigma M_1}, Ad_{\sigma M_1 \sigma M_2}, \ldots, Ad_{\sigma M_1 \sigma M_2 \cdots \sigma M_{n-1}})\end{aligned}$$

Expanding, we see that $M_1, \ldots, M_n$ and $\xi(d\sigma_\rho)^{-1}$ are given, we can find an expression for $Ad_{\sigma M_1}$. This expression determines σM_1 up to sign. Similarly, we may determine $\sigma M_1 \sigma M_2$ up to sign, hence σM_2. Continuing in this way we determine every σM_j for $1 \le j \le n-1$; and σM_n is determined by the condition that $M_1 \ldots M_n = \sigma(M_1 \ldots M_n)$. □

Remark 2.9. Part (b) of Corollary 2.7 follows from 2.8 since for P_n we have that each M_j is conjugate to its image $\sigma(M_j)$ so that unless the trace of M_j is zero its sign is also determined.

COROLLARY 2.10. *Suppose that $\sigma \in P_n$. Then for generic ρ, $\xi(d\sigma_\rho)$ determines $\sigma : X \to X$.*

Remark 2.11. Contrast this with what we would like to know, namely, that for abelian ρ, $\xi(d\sigma_\rho)$ determines σ.

Corollary 2.7 shows that the points in which we are interested lie inside the fixed set of the diffeomorphism. We use this as follows. Suppose that σ is a pure braid on n strands, we denote the closure of this braid by $\hat{\sigma}$; this is an n-strand link in S^3. Then one finds easily that a presentation for its fundamental group is given by:

$$G(\hat{\sigma}) = \langle x_1, \dots, x_n \mid x_1 = \sigma x_1, \dots, x_n = \sigma x_n \rangle$$

When $\hat{\sigma}$ is a hyperbolic link, we may use the deformation theory of Thurston to examine the dimension of $B(\sigma)$. From henceforth, we suppose that this is the case. This is justified by the following simple lemma:

LEMMA 2.12. *Let N be any normal subgroup in P_n other than the center. Then N contains braids whose closure is a hyperbolic link in S^3.*

Proof. It is shown in [10] that such an N always contains elements which represent pseudo-Anosov mapping classes of the punctured disc. Let σ be one such and set $M(\sigma)$ to be the mapping torus of σ; this is a hyperbolic manifold with $n+1$ torus boundary components. One of these corresponds to the $\partial D^2 \times S^1$ and we wish to cap off the S^1 factor with a disc in order to get a closed braid in S^3. This corresponds to $(0, 1)$ surgery on the cusp—unfortunately this may fail to be a hyperbolic manifold. However, deformation theory [18] Theorem 5.8.2 implies that for very large k, $(0, k)$ does have a singular hyperbolic structure which may be de-singularized in a branched covering. It follows that the braid σ^k has hyperbolic closure, as was required. □

We see that $\mathrm{Fix}(\sigma)$ is an affine algebraic subset of X; decompose it into its irreducible components $\mathrm{Fix}(\sigma) = F_0 \cup F_1 \cup \cdots \cup F_n$. The connection with the map ξ_σ is given by:

LEMMA 2.13. *Suppose that $\rho \in \mathrm{Fix}(\sigma)$. Then the Zariski tangent space $T_\rho(\mathrm{Fix}(\sigma))$ is a subspace of* $\ker(\xi_\sigma(\rho) - \mathrm{Id})$.

Proof. (See [19].) Consider the image of a smooth path $(v(1)_t x_1, \dots, v(j)_t x_j, \dots, v(n)_t x_n)$ which lies inside the fixed point set. Differentiating the condition that this path satisfies the relations $\sigma x_i = x_i$ gives the result. □

Remarks 2.14. (a) This is also a smooth fact: If $f : X \to X$ is a diffeomorphism then the same proof shows that if x is a smooth point of $\mathrm{Fix}(f)$ then $T_x(\mathrm{Fix}(f))$ is a subspace of $\{v \in T_x(X) \mid df_x(v) = v\}$. Notice that we really only need 2.13 for smooth points.

(b) The calculation of 2.13 shows that the subspace $\ker(\xi_\sigma(\rho) - \mathrm{Id})$ represents the cocycles $Z^1(G(\hat{\sigma}); Ad\rho)$ (see below) and in general this subspace can be larger than the dimension of the Zariski tangent space to the variety $\mathrm{Fix}(\sigma)$—what is really measured by the cocycles is the Zariski tangent space to the scheme defined by the presentation of $G(\hat{\sigma})$. However we do have equality in a special case:

THEOREM 2.15. *Suppose that $\hat{\sigma}$ is a hyperbolic link. Then for a Zariski open set of points on the component containing the complete structure we have* $\ker(\xi_\sigma(\rho) - \mathrm{Id}) = T_\rho(\mathrm{Fix}(\sigma))$.

Proof. There are two points here. Firstly, it is shown in [17] that for any irreducible algebraic set F, there is a Zariski open set U in F on which the dimension of the Zariski tangent space agrees with the dimension of F as defined in §1. These are the *smooth points* of F. Further, two components of $\mathrm{Fix}(\sigma)$ may meet, but they can only do so in a proper subvariety. Thus there is a Zariski open subset U in F_i coming from the intersection of the smooth points of F_i which do not meet the rest of $\mathrm{Fix}(\sigma)$. On such points the Zariski tangent space is the object which one would expect, via, say smooth topology. These considerations have nothing to do with the component containing the complete structure.

The second point is that we need to show that $\ker(\xi_\sigma(\rho) - \mathrm{Id})$ actually gives the Zariski tangent space. Here we need to use the fact that we have a special component and the proof requires some group cohomology, for which we reference [4].

We recall the outline here. Suppose that $\rho : \Gamma \to SL(2, \mathbf{C})$ is a representation and $\cdots \to M_1 \to \mathbf{Z}[\Gamma] \to \mathbf{Z} \to 0$ is a projective resolution. Then composing ρ with the adjoint representation $\mathrm{Ad} : SL(2, \mathbf{C}) \to sl_2$ we can form a chain complex of $\mathbf{Z}[\Gamma]$-modules $\{\mathrm{Hom}(M_i, sl_2)\}$ from this resolution; the action on the first factor being the usual one, the action on the second being twisted by $\mathrm{Ad}\rho$. The cohomology groups $H^*(\Gamma; \mathrm{Ad}\rho)$ are defined to be those of this complex. We need little concerning these cohomology groups other than the fact that they satisfy Poincaré duality and that if M is a $K(\pi, 1)$ then $H^*(\Gamma; \mathrm{Ad}\rho) = H^*(M; \mathrm{Ad}\rho)$.

As above we let M be the exterior of the pure link coming from $\hat{\sigma}$. Then we have an exact sequence coming from the map induced by inclusion $i : \partial M \to M$:

$$\to H^1(M; \mathrm{Ad}\rho) \to H^1(\partial M; \mathrm{Ad}\rho) \to H^2(M, \partial M; \mathrm{Ad}\rho) \to H^2(M; \mathrm{Ad}\rho) \to$$

The fact that we need is that if ρ is the complete representation then Mostow-Weil rigidity gives $H^2(M, \partial M; \mathrm{Ad}\rho) \to H^2(M; \mathrm{Ad}\rho)$ is the zero map. This implies that $H^1(\partial M; \mathrm{Ad}\rho) \to H^2(M, \partial M; \mathrm{Ad}\rho)$ is surjective and by duality that

$H^1(M; \mathrm{Ad}\rho) \to H^1(\partial M; \mathrm{Ad}\rho)$ is injective. If M has n torus boundary components, it is easy to see that $H^1(\partial M; \mathrm{Ad}\rho) = \mathbf{C}^{2n}$ so that by duality again, the image of $H^1(M; \mathrm{Ad}\rho)$ has dimension n; since this group was injected it follows that $\dim_{\mathbf{C}}(H^1(M; \mathrm{Ad}\rho)) = n$. Since $H^1(M; \mathrm{Ad}\rho) = Z^1(M; \mathrm{Ad}\rho)/B^1(M; \mathrm{Ad}\rho)$ we have that $\dim_{\mathbf{C}}(Z^1(M; \mathrm{Ad}\rho)) = n+3$. As we observed above, representations give rise to cocycles, so that we have $T_\rho(\mathrm{Fix}(\sigma)) \leq Z^1(M; \mathrm{Ad}\rho)$. Since the left-hand side of this inequality is known ([18] Theorem 5.6) to satisfy $n+3 \leq T_\rho(\mathrm{Fix}(\sigma))$ we have equality at the complete representation.

The usual considerations show (see [18]) that for all representations γ on the component sufficiently near to the complete structure we have $H^2(M, \partial M; \mathrm{Ad}\gamma) \to H^2(M; \mathrm{Ad}\gamma)$ is the zero map and a Zariski denseness argument completes the proof. □

Now let $R \subset \mathrm{Fix}(\sigma)$ be the component which contains the discrete faithful representation of the link group $G(\hat{\sigma})$. Notice that we do not work with representations up to conjugacy here so that the deformation theory of Thurston ([18] Theorem 5.6 & 5.8.2) implies that $\dim_{\mathbf{C}}(R)$ is at least $n+3$. In fact, one deduces from Weil [19] or Mostow rigidity that since R contains the complete representation, $\dim_{\mathbf{C}}(R) = n+3$; see also the proof of 2.15. It follows that:

COROLLARY 2.16. *For a Zariski open subset U of R we have* $\dim_{\mathbf{C}}(\ker(\xi_\sigma(\rho) - \mathrm{Id})) = n+3$.

Now we consider the map ξ_σ restricted to R. Our aim will be to show that $\xi_\sigma(R)$ has large dimension:

THEOREM 2.17. *The Zariski closure of $\xi_\sigma(R)$ has dimension $n+3$.*

The proof of this fact will occupy the rest of this section. We first draw some corollaries:

COROLLARY 2.18. *If σ is a pure braid with hyperbolic closure, then:*

(a) $\dim_{\mathbf{C}}(B(\sigma)) \geq n+3$.

(b) For $\rho \in U$, a Zariski open subset of $B(\sigma)$, we have $\xi_\sigma^{-1}(\rho)$ has dimension $2n-3$.

Proof. Part (a) is clear and (b) follows directly from Theorem 2.17 and 1.1. □

This already has the following consequence. We have already observed that $\dim_{\mathbf{C}}(X_A) = n+2$. For $n = 3, 4$ we have that $2n-3 < n+2$, so that we would expect there to be an abelian representation which does not map to the identity matrix; that is to say, σ does not lie in the kernel of the Gassner representation.

We shall clarify this situation by doing somewhat better in §3.

We now embark on the proof of 2.17. To do this we must use extra structure coming from the nature of the automorphisms. To this end we define a map $tr : X \to \mathbf{C}^n$ given by:

$$tr((M_1, \ldots, M_n)) = (tr(M_1), \ldots, tr(M_n))$$

where tr denotes the usual trace map in $SL(2, \mathbf{C})$. Again we easily have that for $\sigma \in P_n$ $tr = tr \circ \sigma$ so that

$$dtr_\rho dL_\rho = dtr_{\sigma\rho} dL_{\sigma\rho} \circ \xi(d\sigma_\rho).$$

We shall be interested in the case that $\rho \in R$, so that $\sigma\rho = \rho$ where we may rewrite this as $dtr_\rho dL_\rho(I - \xi(d\sigma_\rho)) = 0$ or in other words

$$\operatorname{Im}(I - \xi(d\sigma_\rho)) \subset \ker(dtr_\rho dL_\rho).$$

It is therefore the analysis of the map $tr : SL(2, \mathbf{C}) \to \mathbf{C}$ which is our next task. Recall that the Lie algebra of $SL(2, \mathbf{C})$ is generated as a $\mathbf{C}$ vector space by

$$e_R = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad e_U = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \text{and} \quad e_L = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

We shall always use this ordered basis in the calculations which follow.

LEMMA 2.19. *Let A be the $SL(2, \mathbf{C})$ matrix*

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Then $dtr_A dL_A : T_e(SL(2, \mathbf{C})) \to \mathbf{C}$ has matrix $[a - d, c, b]$.

Proof. If v is a vector in the Lie algebra, then the given map can be computed by consideration of $tr(vA)$. □

COROLLARY 2.20. *$dtr_A dL_A$ has critical points only when A is central.*

One sees easily that one can recover a finite number of $SL(2, \mathbf{C})$ matrices from any 3-tuple of complex numbers $\{\alpha, \beta, \gamma\}$, and that the only matrices giving rise to $\{0, 0, 0\}$ are central. Our information will come from $\ker(dtr_A dL_A)$. One cannot recover the matrix A from this; only its fixed points:

LEMMA 2.21. *Suppose that* $\ker(dtr_A dL_A) \subset T_e(SL(2,\mathbf{C})) = \mathbf{C}^3$ *is two-dimensional. Then one can recover the fixed points of* A *for the action by fractional linear transformations on* $\mathbf{C}$.

Proof. Consider the inner product on $\mathbf{C}^3$ given by $(\alpha,\beta,\gamma)\cdot(\alpha',\beta',\gamma') = \alpha\alpha'+\beta\beta'+\gamma\gamma'$. This form is not positive definite, but it is nondegenerate, so that for any subspace W we have $\dim(W)+\dim(W^{\perp}) = 3$.

Then the map $dtr_A dL_A$ can be considered as taking the inner product with the matrix $[a-d, c, b]$. This vector cannot be recovered from the subspace $\ker(dtr_A dL_A)$ but it can be recovered up to scaling by a complex number as the orthogonal complement of $\ker(dtr_A dL_A)$. The fixed points for the action of A on $\mathbf{C}$ satisfy

$$z = \frac{az+b}{cz+d};$$

that is to say $cz^2+(d-a)z-b = 0$. Thus simultaneous scaling does not alter the fixed points. □

Provided we adopt the appropriate convention for the case of coincident fixed points, notice that any two-dimensional subspace of $T_e(SL(2,\mathbf{C}))$ gives rise to a nonzero vector defined up to scaling and hence a pair of fixed points for a 1-parameter subgroup of $SL(2,\mathbf{C})$.

COROLLARY 2.22. *Suppose that* $A, B \in SL(2,\mathbf{C})$. *Then* $\mathrm{Fix}(A) = \mathrm{Fix}(B)$ *if and only if* $\ker(dtr_A dL_A) = \ker(dtr_B dL_B)$.

If we return to the context of the map $tr : X \to \mathbf{C}^n$ we see that if $\rho = (\rho(x_1),\dots,\rho(x_n))$ then $\ker(dtr_\rho dL_\rho) = \oplus_{i=1}^n \ker(dtr_{\rho(x_i)} dL_{\rho(x_i)})$ and this subspace has dimension $2n$ provided that none of the $\rho(x_i)$'s is a central matrix. It will also be useful to observe:

LEMMA 2.23. $Ad_B(\ker(dtr_A dL_A)) = \ker(dtr_{BAB^{-1}} dL_{BAB^{-1}})$.

Proof. This is the commutativity of the diagram of §1. □

COROLLARY 2.24. *Suppose that* $Ad_B(\ker(dtr_A dL_A)) = \ker(dtr_B dL_B)$. *Then* A *and* B *commute.*

Proof. By 2.23 & 2.24, BAB^{-1} and B have the same fixed points. But then $\mathrm{Fix}(B) = \mathrm{Fix}(BAB^{-1}) = B(\mathrm{Fix}(A))$ implies that $\mathrm{Fix}(A) = \mathrm{Fix}(B)$. □

Our final observation before completing the proof of 2.17 comes from a slight variation of Theorem 5.8.2 of [18]. There it is shown that on the component containing the complete structure, the traces of meridians control representations up to conjugacy. Here we are dealing with representations and seek to control these by knowledge of the map $\ker(dtr_\rho dL_\rho)$. In the light of our previous remarks, this amounts to asking about the fixed points of the meridians.

THEOREM 2.25. *In the notation established above, let* $G(\hat{\sigma}) = \langle x_1, \ldots, x_n \mid x_i = \sigma x_i \ 1 \le i \le n\rangle$ *be the group of a hyperbolic link. Let* R *be the component of* $\mathrm{Fix}(\sigma)$ *containing the discrete faithful representation and set* R^* *to be the Zariski open subset of those representations for which no* x_j *maps into the center.*

Then the map $\varphi : R^* \to (\mathbf{CP}^2)^n$ *given by*

$$(M_1, \ldots, M_n) \to \{[a_1 - d_1, c_1, b_1], \ldots, [a_n - d_n, c_n, b_n]\}$$

is finite to one on a Zariski open subset of R^* *which contains the complete representations.*

Proof. It is shown in [17] that R^* is can be given the structure of a (abstract) variety. In this structure, the map φ continues to be a morphism, since this is a local notion.

Let ρ be some representative of the discrete faithful representation and suppose for simplicity that the fixed points f_j of the parabolics $\rho(M_j)$ are all finite nonzero points. This means that the equation for the fixed points has the form $(z - f_j)^2 = 0$, so that the associated projective point has coordinates $[2f_j, 1, -f_j^2]$.

Fix some classical neighborhood V of ρ inside R^* and consider where the preimage of the point $p = \{[2f_1, 1, -f_1^2], \ldots, [2f_n, 1, -f_n^2]\}$ meets V.

By Mostow-Weil rigidity the only representations near the complete structure which keep every meridian parabolic are conjugate to the complete representation. If α is any representation inside V mapping to p, it must be parabolic on the boundary and since $n \ge 3$, all conjugacy is spent, so that $\alpha = \rho$. Hence we see that the full preimage $\varphi^{-1}(p)$ is a union of components, one of which is a point, so that by theorem 1.1 $\varphi^{-1}(p)$ is a finite union of points and the Zariski closure of the image has dimension $n+3$. Applying 1.1 again, we deduce that φ is finite to one on a Zariski open set.

The second clause follows from [17]. There it is shown that if we define for $x \in R^*$ a function $e(x) = \max\{\dim_{\mathbf{C}}(A) \mid A \text{ is a component of } \varphi^{-1}(\varphi(x))\}$ then the set $S_n = \{y \mid e(y) \ge n\}$ is closed. We have just seen that $e(\rho) = 0$ so that there are both Zariski and classical neighborhoods of ρ on which φ is finite to one. $\square$

One interpretation of this result is that for most representations on the component containing the complete structure, one cannot vary the representation without changing the fixed set of at least one x_j; or equivalently without changing $\ker(dtr_{\rho(x_j)} dL_{\rho(x_j)})$.

Remark 2.26. Notice that in general, one can vary irreducible representations without varying the fixed points. An example comes from the Borromean rings: If one component is removed, then what remains is a pair of two unlinked circles. It follows that if we map one of the generators to the identity matrix, then the group becomes free of rank two. However there are many curves of representations of this latter group which do not move the fixed points of the images of the generators.

Proof of 2.17. Suppose that the theorem were false and the map $\xi_\sigma : R \to \xi_\sigma(R)$ drops dimension. By Corollary 2.16, there is a Zariski open V in R for which $\dim_{\mathbf{C}}(\ker(\xi_\sigma(\rho) - I) = n + 3$ for all $\rho \in V$. By theorem 2.25, there is a Zariski open V' for which variation in the representation causes at least one $\ker(dtr_{\rho(x_j)} dL_{\rho(x_j)})$ to vary. Finally, there is a Zariski open V'' on which $\rho(x_i)$ does not commute with $\rho(x_j)$ $i \neq j$ for every representation in V''.

Choose some fixed $\rho \in V \cap V' \cap V''$; R is irreducible so that this set is nonempty open. By theorem 1.1, every component of $\xi_\sigma^{-1}(\xi_\sigma(\rho))$ has dimension at least $\dim_{\mathbf{C}}(R) - \dim_{\mathbf{C}}(\xi_\sigma(R)) > 0$ so that we may choose at least a curve C so that $\rho \in C \subset \xi_\sigma^{-1}(\xi_\sigma(\rho))$. Further, we may arrange a small classical neighborhood W satisfying $\rho \in W \subset V \cap V' \cap V''$.

What we have now achieved is that for all $\alpha \in W \cap C$:

(a) The subspace $\mathrm{Im}(\xi_\sigma(\alpha) - I)$ is independent of α and has dimension $2n - 3$.

(b) $\ker(dtr_\alpha dL_\alpha)$ has dimension $2n$.

We have already observed that we always have $\mathrm{Im}(I - \xi(d\sigma_\rho)) \subset \ker(dtr_\rho dL_\rho)$ and the dimensions involved clearly make this situation very restrictive.

By choice of V', we know that as α varies over $W \cap C$ the subspace $\ker(dtr_\alpha dL_\alpha)$ must vary in at least one of its components, without loss of generality the first component has $\ker(dtr_{\alpha x_1} dL_{\alpha x_1})$ varying. We need to observe:

LEMMA 2.27. $\mathrm{Im}(I - \xi(d\sigma_\alpha))$ *cannot contain a nonzero vector of the form* $(\mathbf{0}, \ldots, \mathbf{0}, v, \mathbf{0}, \ldots, \mathbf{0})$.

Proof. Exactly as for tr, we also have $\mathrm{Im}(I - \xi(d\sigma_\alpha)) \subset \ker(d\pi_\alpha dL_\alpha)$. Now Lemma 2.6 shows that a vector of the form of the statement never lies inside $\ker(d\pi_\alpha dL_\alpha)$. □

Let $\{v_1, v_2, \ldots, v_{2n-3}\}$ be a basis for $\mathrm{Im}(I - \xi(d\sigma_\rho))$ this subspace being fixed throughout what follows.

Let $p_i : T_\alpha(R) \to \mathbf{C}^3$ be the projection of the tangent space onto the i-th factor. Considering each v_j as an element of $\oplus_{i=1}^n \ker(dtr_{\alpha(x_i)} dL_{\alpha(x_i)})$ we see that since $\ker(dtr_{\alpha x_1} dL_{\alpha x_1})$ is not constant over W, the intersection is at most one dimensional so that $p_1(v_j)$ are all multiples of some fixed vector in the first factor. By doing row operations, we may obtain a new linearly independent set (which we relabel) $\{v_1, v_2, \ldots, v_{2n-3}\}$ which has $p_1(v_j) = \mathbf{0}$ for $j \geq 2$.

We now work with the set $\{v_2, \ldots, v_{2n-3}\}$ a linearly independent set of $2n-4$ vectors which lie in $\oplus_{i=1}^{n} \ker(dtr_{\alpha(x_i)} dL_{\alpha(x_i)})$ a subspace of dimension $2n-2$. Observe that the phenomenon of the above paragraph can only happen at most once again; for if it happened twice, by doing row operations twice, we could find $2n-6$ independent vectors lying in $n-3$ two-dimensional kernel factors; so we would have equality and hence a vector of the form $(\mathbf{0}, \ldots, \mathbf{0}, v, \mathbf{0}, \ldots, \mathbf{0})$, which is forbidden by 2.27.

There are two cases, basically identical. We do the first. Suppose that $\{p_2(v_j)\}$ contains a basis for the second factor. By reordering, we may suppose that $\{p_2(v_2), p_2(v_3)\}$ is a basis. By row operations we may find a linearly independent set, again renamed $\{v_2, \ldots, v_{2n-3}\}$ so that $p_2(v_j) = 0$ for $j \geq 4$. Continuing in this way, we reduce the number of vectors at each stage by two, and the dimension of the subspace they live in by two. We began with $2n-4$ vectors in $n-1$ factors of dimension two, so that if we perform this operation $n-3$ times we have two vectors v and $\mathbf{w}$ whose only components lie in two kernel factors; that is, they have the shape $v = (\mathbf{0}, \ldots, \mathbf{0}, v_1, v_2)$ and $\mathbf{w} = (\mathbf{0}, \ldots, \mathbf{0}, \mathbf{w}_1, \mathbf{w}_2)$. The set $\{v_1, \mathbf{w}_1\}$ must be linearly independent, else we are done as above. Consideration of the map of 2.6 shows that since v and $\mathbf{w}$ must lie in $\ker(d\pi_\rho dL_\rho)$ it must be that $v_1 = -\mathrm{Ad}_{\rho(x_{n-1})}(v_2)$ and $\mathbf{w}_1 = -\mathrm{Ad}_{\rho(x_{n-1})}(\mathbf{w}_2)$. It follows that $\{v_2, \mathbf{w}_2\}$ is an independent set, and $\mathrm{Ad}_{\rho(x_{n-1})}(\ker(dtr_{\rho x_n} dL_{\rho x_n})) = \ker(dtr_{\rho x_{n-1}} dL_{\rho x_{n-1}})$.

However 2.24 now implies that $\rho(x_n)$ and $\rho(x_{n-1})$ commute which contradicts our choice of ρ. $\square$

3. Soluble representations. Recall that we have denoted X_A to be the abelian representations. Part of the source of the difficulty comes from the fact that these are part of the singular variety of the soluble representations. In this section we attempt to clarify this difficulty. To do this we must first analyze the way that the abelian representations lie inside the soluble ones.

First let us consider an easy case. Consider the representation ρ^* of the free group of rank n given by

$$\rho^*(x_1) = \begin{bmatrix} p & 0 \\ 0 & 1/p \end{bmatrix} \rho^*(x_2) = \begin{bmatrix} q & 0 \\ a & 1/q \end{bmatrix}$$

and $\rho^*(x_j) = \mathrm{Id}$ for $j > 2$.

LEMMA 3.1. *If $p, q \neq \pm 1$ and $a \neq 0$, then ρ^* is a smooth point of X_S.*

Proof. We shall show that we can parameterize all the soluble representations nearby ρ^* as a complex manifold of dimension $2n+1$, from which the result will follow.

Notice that since the elements in the first two coordinates have a unique common fixed point, if ρ is soluble and nearby, there is a unique point f_ρ fixed by the image of ρ which has to be close to 0. Then there is a unique parabolic transformation fixing identity which carries this fixed point to 0, namely

$$\begin{bmatrix} 1 & -f_p \\ 0 & 1 \end{bmatrix}.$$

Conjugating ρ by this parabolic gives a representation which fixes 0 and has entries very close to those of ρ^* so it has the form

$$\rho(x_i) = \begin{bmatrix} p_i & 0 \\ a_i & 1/p_i \end{bmatrix}$$

where p_1 is close to p, p_2 is close to q, etc.

The required parametrization is $\rho \to \{f_\rho, p_1, a_1, \ldots, p_n, a_n\}$. □

The same proof shows:

LEMMA 3.2. *Suppose that ρ is a soluble nonabelian representation. Then ρ is a smooth point of X_S.*

This fails for an abelian representation consisting of hyperbolic elements, since soluble representations nearby may fix either of two points. But this is the worst that can happen generically:

LEMMA 3.3.

(a) Let

$$\Lambda_i = \begin{bmatrix} \lambda_i & 0 \\ 0 & 1/\lambda_i \end{bmatrix}$$

and consider the representation $\rho = \{\Lambda_1, \Lambda_2, \ldots, \Lambda_n\}$ where none of the $\Lambda_i = \pm 1$. Then the set of soluble representations near to ρ with common fixed point near 0 forms a manifold of dimension $2n+1$.

(b) In the notation of §2, the tangent space of the representations of (a) has a basis:

$\{0, 0, \quad, e_R, \ldots, 0\}$ *where the nonzero entry is in the i-th place for all $1 \leq i \leq n$*

$\{0, 0, \quad, e_L, \ldots, 0\}$ *where the nonzero entry is in the i-th place for all $1 \leq i \leq n$*

together with the vector $\{(1-\lambda_1^2)e_U, (1-\lambda_2^2)e_U, \ldots, (1-\lambda_n^2)e_U\}$.

Proof. For (a) we may use the same parametrization as in 3.1. Given (a), we need only show that the given tangent vectors can all arise and then we have shown that they form a basis as they are visibly linearly independent. The first n vectors come from deformations of the form $\{\Lambda_1, \Lambda_2, \Lambda(t)\Lambda_i, \dots, \Lambda_n\}$ where $\Lambda(t)$ is smooth path of diagonal matrices running through Id. The second n vectors come from a path of the form

$$\begin{bmatrix} \lambda_i & 0 \\ a(t) & 1/\lambda_i \end{bmatrix}$$

in the i-th place, where $a(t)$ is a smooth path through 0.

None of the above deformations has moved the 0-fixed point. This gives the last vector: we simultaneously conjugate all the entries by

$$\begin{bmatrix} 1 & f(t) \\ 0 & 1 \end{bmatrix}.$$

An easy calculation shows that this gives the final vector. □

Denoting the small neighborhood of 3.3(a) by $X_S(\rho, 0)$, we also have $X_S(\rho, \infty)$ and clearly the tangent spaces of these two manifolds taken together span $T_\rho(X)$. Further since both contain $T_\rho(X_A)$, a space of dimension $n + 2$, we see from the dimension formula for vector spaces that $X_S(\rho, 0)$ and $X_S(\rho, \infty)$ are transverse and that $T_\rho(X_S(\rho, 0)) \cap T_\rho(X_S(\rho, \infty)) = T_\rho(X_A)$. From this analysis it follows that:

THEOREM 3.4. *The braid* $\sigma \in \ker(\beta) \leq P_n$ *if and only if* $\mathrm{Fix}(\sigma)$ *contains* X_S.

Proof. If σ fixes X_S, then for every generic $\rho \in X_A$ infinitesimal paths in $X_S(\rho, 0)$ are fixed so that $d\sigma_\rho$ acts as the identity map on $T_\rho(X_S(\rho, 0))$. Similarly for $T_\rho(X_S(\rho, \infty)$; taken together these spaces span $T_\rho(X)$ and $\sigma \in \ker(\beta)$.

For the converse let us note that $\sigma \in \ker(\beta)$ means that for any fixed abelian representation, σ acts as the identity to first order nearby. More formally, in the notation of [11], the representations $\{\exp(\mathbf{x}_1)\Lambda_1, \dots, \exp(\mathbf{x}_n)\Lambda_n\}$ are mapped to $\{\exp(\mathbf{x}_1 + o(\mathbf{x}_i))\Lambda_1, \dots, \exp(\mathbf{x}_n + o(\mathbf{x}_i))\Lambda_n\}$.

However, in the special case that all the $\mathbf{x}_i$ are multiples of e_U (or all multiples of e_L) we see that since $\mathrm{Ad}_{\Lambda_i}(e_U) = \lambda_i^2 e_U$ there are no higher order terms when one comes to use the Campbell-Baker-Hausdorff formula (this is just saying that the elements in a one parameter subgroup commute) and therefore if a representation in this subspace is fixed to first order, it is fixed exactly. Further, $\exp(\alpha e_U) = I + \alpha e_U$ so that

$$\exp(\alpha e_U)\Lambda_i = \begin{bmatrix} \lambda_i & \alpha/\lambda_i \\ 0 & 1/\lambda_i \end{bmatrix}$$

so that for generic choices we obtain the generic soluble representation into $SL(2, \mathbf{C})$.

Putting these observations together, we see that if $\sigma \in \ker(\beta)$ then it fixes a generic point of X_S, hence fixes all of X_S, as was required. □

Remarks 3.5. (a) In Theorem 3.16 of [2], it is shown that $\sigma \in \ker(\beta)$ implies that $\sigma(x_i) = A_i \cdot x_i \cdot A_i^{-1}$ $1 \leq i \leq n$, where we may assume that each A_i lies in $[\ker(\alpha), \ker(\alpha)]$ and $\alpha : F_n \to H_1(F_n)$ is the abelianization map. Any soluble representation into $SL(2, \mathbf{C})$ is metabelian, in particular, kills all such A_i so that this gives the forward implication of 3.4.

Further, we also have:

LEMMA 3.6. *The generic soluble representation into* $SL(2, \mathbf{C})$ *is a faithful representation of the free metabelian group.*

Proof. A result of Magnus [15 cf. [13] p. 28] shows that if $\{s_i\}$ and $\{t_i\}$ $1 \leq i \leq n$ are commuting and invertible polynomial elements then

$$x_i \to \begin{bmatrix} s_i & t_i \\ 0 & 1 \end{bmatrix}$$

is a faithful representation if the free metabelian group F_n/F_n''; one sees easily that if we divide all the entries of the i-th image element by the square root of its determinant we obtain a determinant one representation which is easily seen to be faithful. □

Whence theorem 3.4 gives an alternative proof of Theorem 3.16 of [2]. We also see that the proof of 3.4 gives:

THEOREM 3.7. ([1] & [14]) $\mathrm{Aut}(F_n/F_n'')$ *has a faithful linear representation by linearizing at the generic abelian representation.*

(b) It is also worth noting that the analogous idea for the Burau representation gives much less information here. In this case, the same argument only shows that the link group has all soluble representations with all of the generators having the same diagonal entries; a variety of dimension $n + 2$.

An alternative formulation of 3.4 is:

COROLLARY 3.8. *Let* σ *be a braid in* P_n. *Then* $\sigma \in \ker(\beta)$ *if and only if the set of soluble representations of* $G(\hat{\sigma})$ *consists exactly of the soluble representation of the free group of rank n.*

Remark 3.9. This can happen for links other which are not pure braids; see [6] and references therein.

It is shown in [2] that for a braid in the kernel of the Gassner representation the linking number of any two components is zero; hence every component L_i of the link has an associated longitude λ_i on the i-th torus which is the boundary of a surface embedded in the complement of the link. From 3.4 we deduce easily a new piece of information:

THEOREM 3.10. *Each λ_i lies in $G(\hat{\sigma})''$.*

Proof. Because of the above observations we have that λ_i lies in $G(\hat{\sigma})'$. Hence for every soluble representation ρ we have $tr(\rho(\lambda_i)) = 2$. It follows that $\rho(\lambda_i)$ is either the identity or a parabolic. However a parabolic cannot commute with a hyperbolic element. It follows that on the Zariski open set where none of the meridian generators has trace ± 2 we have $\rho(\lambda_i)$ is the identity. Hence $\rho(\lambda_i)$ is the identity on all of X_S.

The result now follows from 3.6. □

Let X_S^* be the soluble representations which fix ∞; that is to say, upper triangular matrices in $SL(2, \mathbf{C})$. This is a subvariety of X_S left invariant by P_n. Unfortunately, the assignment $\sigma \to \xi(d\sigma_\rho)$ is not a homomorphism for $\rho \in X_S^*$ since the general representation in X_S^* is moved. However we can easily analyze this situation. If M is the $SL(2, \mathbf{C})$ matrix

$$\begin{bmatrix} p & a \\ 0 & 1/p \end{bmatrix}$$

then the action of Ad_M on the ordered basis $\{e_U, e_R, e_L\}$ is given by the matrix

$$\begin{bmatrix} p^2 & -2ap & -a^2 \\ 0 & 1 & a/p \\ 0 & 0 & 1/p^2 \end{bmatrix}.$$

From this we easily see that if we choose the ordered basis (the nonzero entry is in the i-th place) $\{\{0, 0, \quad, e_U, \dots, 0\}\ 1 \le i \le n,\ \{0, 0, \quad, e_R, \dots, 0\}\ 1 \le i \le n,\ \{0, 0, \quad, e_L, \dots, 0\}\ 1 \le i \le n\}$ for the tangent space $T_\rho(X)$ at a general representation ρ in X_S^*, then the matrix for $\xi(d\sigma_\rho)$ has the block matrix form:

$$(*) \qquad \begin{bmatrix} [\beta(\sigma)] & ** & ** \\ 0 & [I] & ** \\ 0 & 0 & [\beta_1(\sigma)] \end{bmatrix}$$

where $\beta(\sigma)$ is the standard Gassner representation and $\beta_1(\sigma)$ is the Gassner representation with the usual variables p_i replaced by $1/p_i$.

Definition 3.11. If G is a group, we set $\gamma_k(G)$ to be the k-th term of the lower central series of G. Notice that if G contains a free group of rank two, then $\gamma_k(G)$ is nontrivial.

THEOREM 3.12. *Suppose that* $\sigma \in \gamma_{3n}(\ker(\beta))$. *Then* $\xi(d\sigma_\rho) = \mathrm{Id}$ *on all of* X_S.

Proof. Let σ be any nontrivial element of $\ker(\beta)$. By Theorem 3.4, $\mathrm{Fix}(\sigma)$ contains X_S; in particular, the assignment $\sigma \to \xi(d\sigma_\rho)$ is now a representation of $\ker(\beta)$. Consideration of the matrices of $(*)$ shows that since σ lies in $\ker(\beta)$ (and hence $\ker(\beta_1)$) all the matrices $\xi(d\sigma_\rho)$ are $3n \times 3n$ upper triangular, hence nilpotent of class $3n$. It follows that for every braid σ in $\gamma_{3n}(\ker(\beta))$ we have that $\xi(d\sigma_\rho)$ is the identity matrix. □

We may combine this result with the calculation of 2.17 and 1.1. We see that for such braids there is a component in $\xi_\sigma^{-1}(I)$ of dimension $2n+1$ which is always larger than the expected dimension of the preimage i.e. $2n-3$.

Again it is worth noting that the relevant calculation for the braid group yields less information. The same kind of proof shows that if we pass to an element of $\gamma_{3n}(\ker(\mathrm{Burau}))$ then this fixes all soluble representation with the same diagonal entries (see Remark 3.5(b)) and we obtain a component of dimension $n+2$ in $\xi_\sigma^{-1}(I)$. However if $n \geq 5$ we have $n+2 \leq 2n-3$.

This suggests a uniformity in the Gassner representation for all n, while for the Burau the $n = 3, 4$ appear to be special.

4. Closed incompressible surfaces. In this section we show:

THEOREM 4.1. *Let L be an n component link. Suppose that the representation variety of L has a component A which contains an irreducible representation and has dimension* $> n+3$.

Then $S^3 \setminus L$ *contains a nonboundary parallel closed incompressible surface.*

This is perhaps of independent interest and has some corollaries which we discuss at the end of the section. Our application in this context will be to obtain a closed incompressible surface in the complement of a four strand braid lying in the kernel of the Gassner representation.

In order to prove this theorem, we shall use some ideas of Hatcher [9] for which we need to introduce some notation. Recall that the weighted simple closed curves on a torus are parametrized by (a dense subset of) projective lamination space which is a real projective line $\mathbf{P}^1$. For a manifold with n torus boundary components curve systems in the boundary lie in the n-fold join $\mathbf{P}^1 * \mathbf{P}^1 * \cdots * \mathbf{P}^1 = S^{2n-1}$. An incompressible, ∂-incompressible surface gives rise to a point on this sphere. In [9], it is shown that the set of points so obtained forms a dense set in a polyhedron of dimension $< n$. To prove 4.1 we observe:

PROPOSITION 4.2. *One can find simple closed curves* C_i *on* T_i *so that the projective class of* $\lambda_1 C_1 + \lambda_2 C_2 + \cdots + \lambda_n C_n \lambda_i \geq 0$ *is never a boundary slope.*

Proof. This is essentially general position, but some mild degree of care is necessary. We begin by briefly recalling how Hatcher's proof in [9] goes, since we shall use some of these ideas.

The starting point is the Floyd-Oertel theorem (see [7]), namely there are a finite number of branched surfaces with the properties: (a) Every essential surface is carried by one of these branched surfaces. (b) Every surface carried by one of these branched surfaces with nonnegative weights is essential. It clearly suffices to deal with the case of a single branched surface, $\mathbb{B}$. One then observes that the boundary of $\mathbb{B}$ is a branched 1-submanifold in ∂M (which we refer to as a union of train tracks, although this is not strictly correct, as there may be bigon regions inside some of the boundary tori) which give rise to a collection of open cells in each $\mathbf{P}^1$. One may also orient these train tracks so that if S is carried by $\mathbb{B}$, the orientation induced on ∂S by these train tracks in $H_1(\partial M)$ gives the class of ∂S in S^{2n-1}, namely forget orientation and projectivize.

Then given two surfaces S_1 and S_2 carried by $\mathbb{B}$ we may form a new surface S_1+S_2 which is also carried by $\mathbb{B}$ by double curve sum. This induces a sum on the boundary curves compatible with the train tracks. The theorem of [9] is proven by observing that with (train track induced) orientations, the intersection number $\partial S_1 \partial S_2 = 0$ so that the cell of curves carried by the boundary train tracks is a self-annihilating subspace of $H_1(\partial M)$ hence has dimension $\leq n$. Projectivizing proves the result.

To prove Proposition 4.2 we proceed as follows. Fix curves $C_1^*, \dots, C_n^*$ so that $\mathbf{C}_i^*$ lies in the i-th torus. Our claim will be that by small adjustment of this collection we may find a set with the properties of the proposition.

It clearly suffices to do this for one branched surface at a time. Fix such a $\mathbb{B}$ as above. The above proof shows that it defines a cell in $\mathbf{P}^1 * \mathbf{P}^1 * \cdots * \mathbf{P}^1$ with at most n linearly independent vertices $\{v_1, \dots, v_n\}$ so that every point in the cell is a positive linear combination of the v_i's. Each v_i is defined by some projective class, the t-th one being of the form $v_t = [C_{1,t}, C_{2,t}, \dots, C_{n,t}]$ where it is possible that the extremal vertices are actually represented by foliations, though this does not affect the argument.

It is also possible that some of the vertices do not meet some of the tori; again this only involves a few more words. Further, it could also happen that there is a torus which none of the vertices meet; in this case we are reduced to the case of $n-1$ tori and there are in fact only $n-1$ independent vertices. Henceforth we suppose that this does not happen.

A small initial adjustment arranges that none of the C_i^* is a multiple of $C_{p,q}$ for any p, q. Recall that we want to arrange that the projective class of $\lambda_1 C_1^* + \lambda_2 C_2^* + \cdots + \lambda_n C_n^*$ is never in the cell defined by nonnegative linear combinations of the $\{v_t\}$.

Let us consider the nonnegative solutions to the equation $\lambda_1 C_{1,1} + \lambda_2 C_{1,2} + \cdots + \lambda_n C_{1,n} = \alpha C_1^*$, for some nonnegative α. We see easily that the set of such solutions forms a linear cone V_1 in $\mathbf{R}^n$ and that it is a proper subset of $\mathbf{R}_+^n$ by our initial adjustments. (In particular, this uses that all of the vertices do not miss the first torus.) It could be empty—for example if C_1^* is not carried by the train track. In this case we are already finished. Note that in order for a linear combination of $\{v_t\}$ to hit a point whose first co-ordinate is projectively C_1^* (or zero) we must clearly choose some point on the plane V_1 in $\mathbf{R}^n$.

Now fix some random point on this graph and consider these value substituted into the equation $\lambda_1 C_{2,1} + \lambda_2 C_{2,2} + \cdots + \lambda_n C_{2,n}$. If this yields some multiple of the curve C_2^* this is regarded as bad and we perturb C_2^* slightly so that this does not happen. Now we consider solutions to the equation $\lambda_1 C_{2,1} + \lambda_2 C_{2,2} + \cdots + \lambda_n C_{2,n} = \beta C_2^*$; this defines another linear subspace V_2 as above. In order for 4.2 to fail we must clearly choose some values lying inside $V_{12} = V_1 \cap V_2$. Observe that our initial perturbation guaranteed that V_{12} is a proper subspace of V_1. Proceeding in this way, we see that there is at most one linear combination of the $\{v_t\}$ which yields the projective class of C_i^* for all $i \leq n-1$ and this defines a unique curve on the n-th torus. By perturbation, we may arrange that C_n^* is not this curve, and the proposition is proven. □

Proof of 4.1. It is slightly simpler to work with the characters of representations in A; we denote this set by A_χ; since A contains an irreducible representatoin, A_χ has dimension strictly larger than n.

Consider the function $\phi : A_\chi \to \mathbf{C}^n$ given by $\chi \to \{\chi(C_1), \ldots, \chi(C_n)\}$ where the curves C_i are chosen as in 4.2. This is a morphism. Let χ_ρ be the character of an irreducible representation. Then dimension considerations dictate that $\phi^{-1}\phi(\chi_\rho)$ has dimension at least one, so that there is a curve C through χ_ρ lying inside $\phi^{-1}\phi(\chi_\rho)$. If we now go to infinity in C we see that this curve of characters must be bounded on every torus component of L, for if not, the bounded class would give a boundary slope lying inside the forbidden simplex of 4.2. Thus the results of [5] imply that L contains a closed incompressible surface. □

Definition 4.3. (See [6].) Let $F(k)$ denote the free group of rank k. A k component link in S^3 is a *homology boundary link* if there is a surjective map $\pi_1(S^3 \setminus L) \to F(k)$.

COROLLARY 4.4. *Suppose that L is a link in S^3 of four or more components, with the property that one can delete one of its components and obtain a boundary link. Then the complement of L contains a closed embedded nonboundary parallel incompressible surface.*

Proof. The hypothesis implies that there is a surjection $\rho : \pi_1(S^3 \setminus L) \to F(n-1)$. The characters of irreducible representations of the target group give a variety of dimension $3(n-1) - 3$ and this is strictly larger than n for $n \geq 4$. □

COROLLARY 4.5. *Any Brunnian link of four or more components contains in its complement a closed embedded incompressible surface.*

COROLLARY 4.6. *Any homology boundary link of two or more components contains in its complement a closed embedded incompressible surface.*

Remark 4.7. The hypothesis on the number of components is necessary as it is shown in [12] that the Borromean rings do not contain such a surface.

We can now apply these results to show:

THEOREM 4.8. *Suppose that* $\sigma \in \ker(\beta_4)$. *Then* $\hat{\sigma}$ *contains in its complement a closed embedded incompressible surface.*

Proof. One sees easily that if we delete one component of $\hat{\sigma}$ that we obtain a three strand braid lying in $\ker(\beta_3)$; since this representation is known to be faithful, we have that it is the trivial braid. Therefore we may apply 4.1 to deduce the result. □

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SANTA BARBARA, CA 93106

REFERENCES

[1] S. Bachmuth, Automorphisms of free metabelian groups, *Trans. Amer. Math. Soc.*, **118** (1965), 93–104.
[2] J. S. Birman, Braids, links and mapping class groups, *Ann. of Math. Stud.*, #82.
[3] ———, Braids, *Contemporary Math.*, **78**.
[4] K. Brown, *Cohomology of groups*, Springer-Verlag, 1982.
[5] M. Culler and P. B. Shalen, Varieties of group representations and splittings of 3-manifolds, *Ann. of Math.*, **117** (1983), 109–146.
[6] T. Cochran, Links with trivial Alexander's module but nonvanishing Massey products, *Topology*, **29** (1990), 189–204.
[7] W. Floyd and U. Oertel, Incompressible surfaces via branched surfaces, *Topology*, **23** (1984), 117–125.
[8] R. Hartshorne, Algebraic geometry, *Springer-Verlag LNS*, #52.
[9] A. Hatcher, On the boundary curves of incompressible surfaces, *Pacific J. Math.*, **99** (1982), 373–377.
[10] D. D. Long, A note on the normal subgroups of mapping class groups, *Math. Proc. Cambridge Philos. Soc.*, **99** (1986), 79–87.
[11] ———, On the linear representations of braid groups I & II, *Trans. Amer. Math. Soc.*, **311** (1989), 535–561 & *Duke Math. Journal*, **59** (1989), 443–460.
[12] M. Lozano, Arcbodies, *Math. Proc. Cambridge Philos. Soc.*, **94** (1983), 253–260.
[13] R. Lyndon and W. Schupp, Combinatorial group theory, Ergennisse der Math. #89, 1977.
[14] A. Magid, Tangential linear representations of automorphism groups, Preprint.
[15] W. Magnus, On a theorem of Marshall Hall, *Ann. of Math.*, **40** (1939), 764–768.
[16] J. Moody, The Burau representation is unfaithful for large n, *Bull. Amer. Math. Soc.* (to appear).
[17] D. Mumford, Complex projective varieties, Grundlehren 221, Springer-Verlag, 1976.
[18] W. P. Thurston, The geometry of topology of 3-manifolds, Preprint.
[19] A. Weil, Remarks on the cohomology of groups, *Ann. of Math.*, **80** (1964) 149–157.