TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 330, Number 1, March 1992

EXAMPLES OF PSEUDO-ANOSOV HOMEOMORPHISMS

MAX BAUER

ABSTRACT. We generalize a construction in knot theory to construct a large family $\mathscr{GR} = \bigcup GR(\mathscr{P})$ of mapping classes of a surface of genus g and one boundary component, where $\mathscr{P}$ runs over some finite index set. We exhibit explicitly the set $\mathscr{GR}^* \subset \mathscr{GR}$ that consists of pseudo-Anosov maps, find the map that realizes the smallest dilatation in $\mathscr{GR}^*$, and for every $\mathscr{P}$, we give a set of defining relations for $GR(\mathscr{P})$.

1. NOTATION AND BACKGROUND

Introduction. Denote by $F(g, b, s)$ an orientable surface of genus g, b boundary components and s punctures. If the topological type of $F(g, b, s)$ is not important we simply refer to it as F. Throughout the paper $F = F(g, b, s)$ will be assumed to be *hyperbolic*, i.e., F has negative Euler characteristic. Suppose that $(\mathscr{F}, \nu)$ is a measured foliation (see [FLP]) and ϕ is an orientation preserving homeomorphism of F, then we define $\phi(\mathscr{F})$ to be the folitation whose leaves are the images of the leaves of $\mathscr{F}$. Furthermore, $\phi_*(\nu)$ is the measure on $\phi(\mathscr{F})$ that is defined as the push forward of the measure ν under ϕ. To be more explicit, if α is an arc transverse to the folitation $\phi(\mathscr{F})$, then $\phi_*(\nu)(\alpha) = \nu(\phi^{-1}(\alpha))$. We define $\overline{\phi}(\mathscr{F}, \nu) = (\phi(\mathscr{F}), \phi_*(\nu))$. An orientation preserving homeomorphism ϕ of F is *pseudo-Anosov* (or *p.A.*) if there is a pair of transverse arational (i.e. no closed leaves) measured foliations $(\mathscr{F}, \nu)$ and $(\mathscr{F}^\perp, \nu^\perp)$ in F, such that $\overline{\phi}(\mathscr{F}, \nu) = (\mathscr{F}, \lambda\nu)$ and $\overline{\phi}(\mathscr{F}^\perp, \nu^\perp) = (\mathscr{F}^\perp, (1/\lambda)\nu^\perp)$, for some $\lambda > 1$. The foliations $\mathscr{F}$ and $\mathscr{F}^\perp$ are called the *stable foliation* and *unstable foliation*, respectively, of ϕ, and we refer to λ as the *dilatation* of ϕ. Note that the property of being p.A. is an invariant of the conjugacy class of a map, as is the dilatation. The first examples of p.A. maps were given by Nielsen (see [Ni and Gi]), however without the machinery of measured foliations. Measured foliations (and the related machinery of train tracks discussed below) were introduced by Thurston (see [Th]) to pursue a systematic investigation of surface homeomorphisms. Thurston's original construction of p.A. maps was generalized in [P1]. Recently, [GK] recognized certain maps arising in knot theory as being p.A., and we will consider a natural generalization of these maps.

This paper is organized as follows. After introducing the necessary background material we proceed in §2 to construct large classes $GR(\mathscr{P})$, where

Received by the editors January 12, 1990.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 57N05.

$\mathscr{P}$ is in some finite index set, of homeomorphisms of $F(g, 1, 0)$ and find in §4 explicitly the subset $GR^*(\mathscr{P})$ of $GR(\mathscr{P})$ that consists of p.A. maps. The p.A. recognition theorem we use is a result by Casson [CB]. We then proceed in §5 to find the map that realizes the smallest dilatation among the maps of $\bigcup_{\mathscr{P}} GR^*(\mathscr{P})$, where $\mathscr{P}$ runs over the index set, and find an upper bound for the smallest dilatation. Although this bound is larger than the upper bound given in [P2], ideas from the present investigation can be used to improve the result in [P2] as is done in [B2]. In the last section we give a faithful linear representation for $GR(\mathscr{P})$, for every $\mathscr{P}$, that allows us to find a set of defining relations for $GR(\mathscr{P})$. Section 3 is devoted to a simplification of the analysis. We remark that most proofs in §§5 and 6 are rather rough outlines; the details are given in my thesis [B1].

I would like to take this opportunity to thank Robert Penner for his support and encouragement as my thesis advisor and for suggesting the problem. I would also like to thank the unknown referee for many helpful suggestions concerning the exposition and for having pointed out a mistake in an earlier version.

Surface homeomorphisms. For more background information on the material presented in this section consult [FLP, CB and Th] or for a more analytic point of view [Ab].

Let $\mathrm{Homeo}(F)$ denote the group of orientation preserving homeomorphisms of F, and $\mathrm{Homeo}_0(F)$ the subgroup of $\mathrm{Homeo}(F)$ consisting of those maps which are isotopic to the identity map. The mapping class group $MC(F)$ is then defined as $\mathrm{Homeo}(F)/\mathrm{Homeo}_0(F)$, and the equivalence class of $\phi \in \mathrm{Homeo}(F)$ is denoted $[\phi] \in MC(F)$. If ϕ, $\psi \in \mathrm{Homeo}(F)$, then we write $\phi\psi$ or $\phi \circ \psi$ for the composition of ψ followed by ϕ.

We say that $\phi \in \mathrm{Homeo}(F)$ is *periodic*, if ϕ^n is the identity map, for some n; ϕ is *reducible* if there is a union C of essential nonboundary, nonpuncture parallel simple closed curves such that $\phi(C) = C$. We also say that $[\phi] \in MC(F)$ itself is periodic, reducible, or p.A., if $[\phi]$ contains a periodic, reducible, or p.A. representative.

We have

Theorem A (Thurston). *A mapping class of an orientable hyperbolic surface is either periodic, reducible, or p.A. The only overlap is between reducible and periodic maps.*

Proof. See [Th]. □

Remark. Some remarks concerning the definition of a reducible map ϕ are in order. Note first that we can assume that the collection C of reducing curves for ϕ does not contain connected components that are isotopic. If we take a maximal (in the sense of point inclusion) set C of reducing curves with this property, then C decomposes the surface into subsurfaces G_j, for j in some (finite) index set J. It follows that ϕ^n, for some $n > 0$ induces a homeomorphism ϕ_j of G_j, for $j \in J$. Thus, ϕ^n can be "reduced" to a collection of "simpler" homeomorphisms, and as C is maximal and does not contain parallel components we conclude from Theorem A that G_j is either periodic or p.A., for $j \in J$.

Nonnegative matrices. Before we state a well-known property of a certain class of nonnegative matrices, we will need to establish some notation. For a general discussion of nonnegative matrices consult [Ga].

$\mathrm{Mat}_l(\mathbb{Z})$ is defined to be the group of $l \times l$ matrices over $\mathbb{Z}$ and A^t, for $A \in \mathrm{Mat}_l(\mathbb{Z})$ is the transpose of A. $\mathrm{Spec}(A)$, for $A \in \mathrm{Mat}_l(\mathbb{Z})$ denotes the *spectrum* of A, i.e. the set of eigenvalues of A listed with multiplicity. If the modulus of $\delta \in \mathrm{Spec}(A)$ is strictly larger than the modulus of any element in $\mathrm{Spec}(A) \setminus \{\delta\}$, then we refer to δ as the eigenvalue of *maximum modulus* of A. Further, suppose that B and C are $n \times m$ matrices with $n, m \geq 1$, then by an expression of the form $B > C$, $B \geq C$, etc. we mean that the relevant relation holds for each component. Similarly, $B > x$, $B \geq x$, etc., for $x \in \mathbb{R}$ means that each component of B satisfies the relevant inequality. Finally, $A \in \mathrm{Mat}_l(\mathbb{Z})$ is said to be *Perron-Frobenius*, (or *P.F.*) if $A \geq 0$ and for some $n \geq 1$ we have $A^n > 0$.

Theorem B (Perron-Frobenius). (a) *The spectrum of a P.F. matrix* $A \in \mathrm{Mat}_l(\mathbb{Z})$ *contains an element* λ *of maximum modulus that is positive real with corresponding eigenvector* x^* *strictly positive.* x^* *is the unique positive eigenvector and* λ *is a simple root of the characteristic polynomial of* A. *Moreover,*

(b) λ *satisfies*

$$\lambda = \min\left\{\max_{1\leq i\leq l} \frac{(Ax)_i}{x_i} : x \in \mathbb{R}^l,\ x > 0\right\}.$$

For the proof we refer the reader to [Ga] where also the following consequence of this theorem is derived.

Corollary C. *Any nonnegative* $A \in \mathrm{Mat}_l(\mathbb{Z})$ *has an eigenvalue* $\lambda > 0$ *such that no element of* $\mathrm{Spec}(A)$ *has modulus exceeding* λ. *In addition, there is a nonnegative eigenvector corresponding to* λ.

We will need later the following easy facts concerning nonnegative matrices.

Lemma 1. *Suppose that* $A_j \in \mathrm{Mat}_l(\mathbb{Z})$ *satisfies* $A_j \geq 0$, *for* $j = 1, 2$, *and denote by* λ_j *the spectral radius of* A_j.

(a) *If* $\mu > \lambda_1$, *then for* $z \in \mathbb{R}^l$ *we have that* $\lim_{n\to\infty}(1/\mu^n)A_1^n z = 0$.

(b) *If* $A_1 \leq A_2$, *then* $\lambda_1 \leq \lambda_2$.

(c) *If* A_1 *is a principal minor of* A_2, *then again* $\lambda_1 \leq \lambda_2$.

(d) *If* $\lambda_1 \in \mathrm{Spec}(A_1)$ *is an eigenvalue of maximum modulus that is a simple root of the characteristic equation then for any nonnegative and nonzero* $z \in \mathbb{R}^l$, *there exists a* $c \in \mathbb{R}$ *such that* $\lim_{n\to\infty}(1/\lambda_1^n)A_1^n z = cx^*$, *where* x^* *is the unique (up to multiples) nonnegative eigenvector corresponding to* λ_1. *In the special case where* A_1 *is P.F. we know that* $c > 0$.

Proof. For the proof of part (a) we will need the following fact. Suppose that $J \in \mathrm{Mat}_t(\mathbb{Z})$ is an elementary Jordan block corresponding to the eigenvalue ρ, then by writing $J = I + (J - I)$, where $I \in \mathrm{Mat}_t(\mathbb{Z})$ is the identity matrix, we conclude, using the binomial formula that $J^n = O(n^{t-1}|\rho|^n)$. It follows that

$$\text{(i)} \qquad \lim_{n\to\infty} \frac{J^n}{\sigma^n} = \overline{0}, \qquad \text{if } \sigma > \rho,$$

where $\overline{0} \in \mathrm{Mat}_l(\mathbb{Z})$ is the zero matrix.

We now write $A_1 = T\Lambda T^{-1}$, where Λ is in Jordan normal form. We readily see that part (a) follows from this and equation (i).

To prove part (d), we note that the columns of T are generalized eigenvectors of A_1 and the rows of T^{-1} are generalized eigenvectors of A_1^t. We denote the eigenvalue of A_1^t that corresponds to λ_1 by y. Suppose that ρ is an eigenvalue of A_1 whose modulus is second largest in the spectrum of A_1 and whose corresponding elementary Jordan block $J' \in \mathrm{Mat}_r(\mathbb{Z})$ has largest size among those corresponding to eigenvalues of the same modulus as ρ. Using equation (i) we see that

$$A_1^n = \lambda_1^n x^* y^t + O(n^{r-1}|\rho|^n),$$

and the first statement of part (d) follows, with $c = y^t z$.

To prove the second statement we assume that A_1 is P.F. and note that as A^t is also P.F. the Perron-Frobenius theorem implies that $y > 0$. It follows that $c = y^t z > 0$ as desired.

For part (b), assume to derive a contradiction that $\lambda_1 > \lambda_2$. We choose an eigenvector w of A_1 that corresponds to the eigenvector λ_1. As $A_j \geq 0$, for $j = 1, 2$ and $w \geq 0$ by Corollary C, we have $0 \leq w = (1/\lambda_1^n)A_1^n w \leq (1/\lambda_1^n)A_2^n y$. But part (a) applies to show that the right-hand side converges to 0, hence $w = 0$ which is absurd.

For part (c), suppose that A_1 arises from A_2 by deleting the vth row and column of A_2, and denote by A_3 the matrix we get if we replace the vth row and column of A_2 by a zero row and zero column, respectively. If λ_3 denotes the spectral radius of A_3, then we conclude from part (b) that $\lambda_3 \leq \lambda_2$.

We proceed to show that $\lambda_1 = \lambda_3$. For a vector $x \in \mathbb{R}^{l-1}$, denote by $\overline{x} \in \mathbb{R}^l$ the vector we get by inserting 0 between the $(v-1)$th and the vth position of x, and if $x \in \mathbb{R}^l$ is a vector with $x_v = 0$, then we denote by $x' \in \mathbb{R}^{l-1}$ the vector we get from x by deleting the vth coordinate. It is immediate that for $x \in \mathbb{R}^{l-1}$, $A_1 x = \delta x$ implies that $A_3\overline{x} = \delta\overline{x}$, and conversely, if $x \in \mathbb{R}^l$ and $\delta \neq 0$, then $A_3 x = \delta x$ implies that $x_v = 0$ and hence $A_1 x' = \delta x'$. We conclude that the spectrum of A_3 differs from the spectrum of A_1 only by containing an extra eigenvalue 0, in particular we showed that $\lambda_1 = \lambda_3$. □

Measured train tracks. We next introduce the notion of train track that enables us to translate the problem of identifying a map as p.A. into a problem in linear algebra. For more background information consult [Pa, P1, PP, Th] and the monograph [HP].

A *train track* τ in a hyperbolic surface F is defined as follows

(a) τ is a smooth branched one-submanifold embedded in F; edges are called *branches*, vertices are called *switches* and a *half-branch* of τ is an edge of the first barycentric subdivision of τ;

(b) for every switch v of τ and half-branch b incident at v, there is a half-branch b' incident at v and a smooth arc α contained in τ, such that α contains v in its interior and intersects b and b' (for a local model of the half-branches incident at a switch see Figure 1);

(c) the number of half-branches incident at a switch is at least three;

(d) no connected component of $F \setminus \tau$ is an embedded null-gon, mono-gon, bi-gon, once punctured null-gon, or smooth annulus.

For an example of a train track see Figure 9.

A branched submanifold of F that fails to be a train track only by possibly having complementary bi-gons is called a *bi-gon track* or simply *track*. A track τ

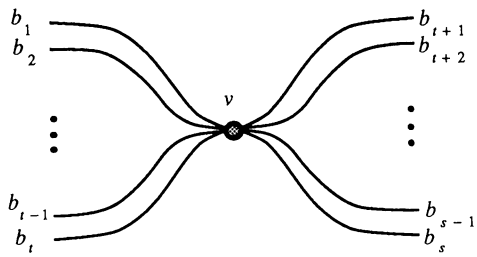

FIGURE 1. Incoming and outgoing branches at a switch

is said to *fill* F if each component of $F \setminus \tau$ is topologically a disc, a punctured disc or a boundary-parallel annulus. Of course by a *subtrack* of a (train) track we mean a branched submanifold that is itself a (train) track.

For each switch v of a track τ we arbitrarily choose a half-branch $b = b(v)$ incident at v. Any half-branch b' incident at v for which there is a smooth arc α contained in τ that intersects b and b' and contains v is called *outgoing* at v, and the half-branches incident at v that are not outgoing are called *incoming* at v. As an example we consider the half-branches incident at v as shown in Figure 1. Assume that, in the notation of Figure 1, we made the choice $b(v) = b_1$, then the incoming half-branches are $b_1, \dots, b_t$ and the outgoing ones are $b_{t+1}, \dots, b_s$. We remark in passing that the two half-branches that make up a branch might be both incoming or both outgoing as in general there is no global orientation on a track.

A function that assigns a nonnegative number to each branch of a track τ is called a *weight* on τ and we next define a special kind of weight. As any weight μ on τ assigns a number to each branch of τ, there is an induced assignment of numbers (also denoted μ) to half-branches. Suppose that at a switch v we have t incoming half-branches, denoted $b_1, \dots, b_t$ and $s - t$ outgoing half-branches, denoted $b_{t+1}, \dots, b_s$ (as in Figure 1), then we say that the *switch condition* on μ holds at v if

$$\sum_{i=1}^{t} \mu(b_i) = \sum_{i=t+1}^{s} \mu(b_i).$$

μ is called *measure* on τ if the switch condition holds at each vertex of τ, and the set of all measures on τ is denoted $V(\tau)$.

We usually assume that the branches of τ are ordered, say $(b_1, \dots, b_l)$. Corresponding to a weight μ on τ, we then define $x = (\mu(b_1), \dots, \mu(b_l)) \in \mathbb{R}^l$. $V(\tau)$ can in this way be identified with a closed cone in the positive quadrant of $\mathbb{R}^l$. By abuse of notation we also refer to x as weight on τ, and if $\mu \in V(\tau)$, then we refer to x as measure on τ.

We next choose a regular neighborhood of τ that is foliated by arcs that are transverse to τ away from the switches of τ and such that each arc intersects τ exactly once. Moreover, a local model of the foliated neighborhood in a neighborhood of a switch v is as in Figure 2(b), where we took the case of three incoming and two outgoing half-branches (or vice versa) at v, (see Figure 2(a)). The arcs of the foliationa re called *ties* and we refer to the foliated neighborhood as *standard tie neighborhood* of τ. Finally, for each branch b of τ we choose a tie, called *central tie* that intersects b exactly once.

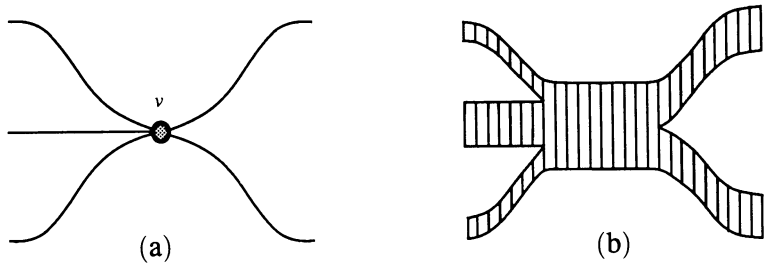

FIGURE 2. Standard tie neighborhood of τ

For a track τ embedded in a surface F we say that a track τ' is *carried* by τ if τ' is contained in a standard tie neighborhood of τ and transverse to the ties. In the special case where $\tau' = \phi(\tau)$, for an orientation preserving homeomorphism ϕ, we also say that τ is *invariant* under ϕ. We remark that although the notion of carrying is transitive it is not symmetric.

Assume now that the track τ is invariant under ϕ, ϕ then induces a map

$$\hat{\phi}\colon V(\tau) \to V(\tau)$$

to be described next. We choose a standard tie neighborhood of τ together with a collection of central ties. Suppose that $\{b_1, \dots, b_l\}$ denotes the branches of τ, then for $\mu \in V(\tau)$ we define

$$\hat{\phi}(\mu)(b_i) = \sum_{j=1}^{l} n(i, j)\mu(b_j),$$

where $n(i, j)$ counts the number of times $\phi(b_j)$ intersects the central tie corresponding to b_i. As each $\mu \in V(\tau)$ satisfies the switch conditions, we see that the definition of $\hat{\phi}$ is independent of the choice of central tie, although $n(i, j)$ is not.

We remark that if τ is a train track, then there is an injection $\mathcal{J}$ of $V(\tau)$ into the space of (equivalence classes of) measured foliations $\mathcal{MF}(F)$ (see [CB] or [FLP]). Moreover, the following diagram commutes:

$$\text{(ii)} \qquad \begin{array}{ccc} \mathcal{MF}(F) & \xrightarrow{\overline{\varphi}} & \mathcal{MF}(F) \\ \mathcal{J}\downarrow & & \mathcal{J}\uparrow \\ V(\tau) & \xrightarrow{\hat{\phi}} & V(\tau) \end{array} .$$

In case τ has complementary bi-gons then $\mathcal{J}$ is still defined, but is no longer injective.

The following distributive property of the 'hat' operator is most useful. Suppose that τ is invariant under ϕ_1 and ϕ_2, then τ is invariant under $\phi_2\phi_1$ (by transitivity of the notion 'is carried by') and $\widehat{\phi_2\phi_1} = \hat{\phi}_2\hat{\phi}_1$.

We now choose a standard tie neighborhood of τ together with a central tie for each branch. We order the branches of τ, say $(b_1, \dots, b_l)$, and identify $V(\tau)$ (as before) with a subset of $\mathbb{R}^l$. Clearly, we can represent $\hat{\phi}$ by an integral matrix $M \in \mathrm{Mat}_l(\mathbb{Z})$, such that the $i \times j$ entry of M is $n(i, j)$, for

$1 \leq i, j \leq l$. M is referred to as *incidence matrix* of ϕ with respect to τ. (Note that M also depends on our choice of ordering of the branches of τ and our choice of central ties.)

A p.A. recognition theorem.

Theorem D (Casson). *Suppose that an orientation preserving homeomorphism ϕ of a hyperbolic surface F leaves a train track τ invariant but leaves no proper subtrack invariant. Then ϕ is isotopic to a p.A. map if and only if*

(a) *τ fills F, and*

(b) *if τ itself is a subtrack of a ϕ invariant train track τ'' (not necessarily proper) then the induced map $\hat{\phi}'': V(\tau'') \to V(\tau'')$ has no nonzero fixed point.*

Proof. See [CB]. □

2. Construction of the examples

We must recall the notion of *Dehn twist* along a simple closed curve c embedded in F: Suppose that A denotes the annulus

$$A = \bigcup_{1 \leq t \leq 2} C_t \subset \mathbb{R}^2,$$

where $C_t = \{t \in \mathbb{R}^2 : |x| = t\}$. We define a homeomorphism $\tilde{r}$ of A as rotation of C_t by $2\pi(t-1)$ in the clockwise direction, for $1 \leq t \leq 2$. Let $\tilde{h}$ be an orientation preserving homeomorphism of A onto a neighborhood of c such that $h(C_{1.5}) = c$. The Dehn twist $\tilde{\tau}_c$ is then defined as $\tilde{h} \circ \tilde{r} \circ \tilde{h}^{-1}$ on $\tilde{h}(A)$ and is the identity map outside. Clearly, $\tilde{\tau}_c$ is a homeomorphism of F whose isotopy class does not change if we isotope c or $\tilde{h}$. See Figure 3 for the effect of $\tilde{\tau}_c$ on an arc α transverse to c.

We next define the class of examples we will analyze. By a *partition* $\mathscr{P}$ of $2g$, we mean a k-tuple, where $k \geq 2$, of positive integers

$$\mathscr{P} = (m_1, \dots, m_k) \in \mathbb{N}^k,$$

such that $2g = m_1 + \cdots + m_k$. Note that if $\mathscr{P}' = (m_1', \dots, m_s')$ is another partition of $2g$, then $\mathscr{P} = \mathscr{P}'$, if and only if $k = s$ and $m_i = m_i'$, for $i = 1, \dots, k$.

Denote by $c_1, \dots, c_{2g}$ the simple closed curves in $F(g, 1, 0)$ depicted in the top portion of Figure 4. For a given partition $\mathscr{P} = (m_1, \dots, m_k)$ of $2g$

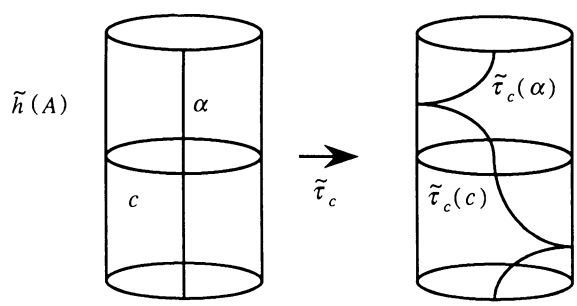

FIGURE 3. Dehn twist

we define

$$v(i) = \begin{cases} 0, & \text{for } i = 1; \\ \sum_{j=1}^{i-1} m_j, & \text{for } 1 < i \leq k, \end{cases}$$

and the double sequence

$$c(i, j) = c_{v(i)+j}, \qquad \text{for } 1 \leq j \leq m_i \text{ and } 1 \leq i \leq k.$$

So $c(1, 1), \ldots, c(1, m_1)$ denotes the first m_1 curves of $(c_1, \ldots, c_{2g})$, $c(2, 1), \ldots, c(2, m_2)$ denotes the next m_2 curves and so forth. Finally, $c(k, 1), \ldots, c(k, m_k)$ denotes the last m_k curves.

Define

$$\varepsilon(i) = \begin{cases} +1, & \text{if } i \text{ odd}, \\ -1, & \text{if } i \text{ even} \end{cases}$$

and

$$\tilde{\tau}(i, j) = \tilde{\tau}_{c(i,j)}^{\varepsilon(i)}.$$

We finally set

$$\widetilde{S}_i = \tilde{\tau}(i, m_i) \cdots \tilde{\tau}(i, 1), \qquad \text{for } 1 \leq i \leq k.$$

To illustrate the maps $\widetilde{S}_i$, we take $g = 3$ and $\mathscr{P} = (2, 3, 1)$. The curves $c_1, \ldots, c_6$ are partitioned into three sets, the first being $c(1, 1) = c_1$, $c(1, 2) = c_2$, the second $c(2, 1) = c_3$, $c(2, 2) = c_4$, $c(2, 3) = c_5$ and the third $c(3, 1) = c_6$. We also have $\widetilde{S}_1 = \tilde{\tau}(1, 2)\tilde{\tau}(1, 1)$, $\widetilde{S}_2 = \tilde{\tau}(2, 3)\tilde{\tau}(2, 2)\tilde{\tau}(2, 1)$, and $\widetilde{S}_3 = \tilde{\tau}(3, 1)$.

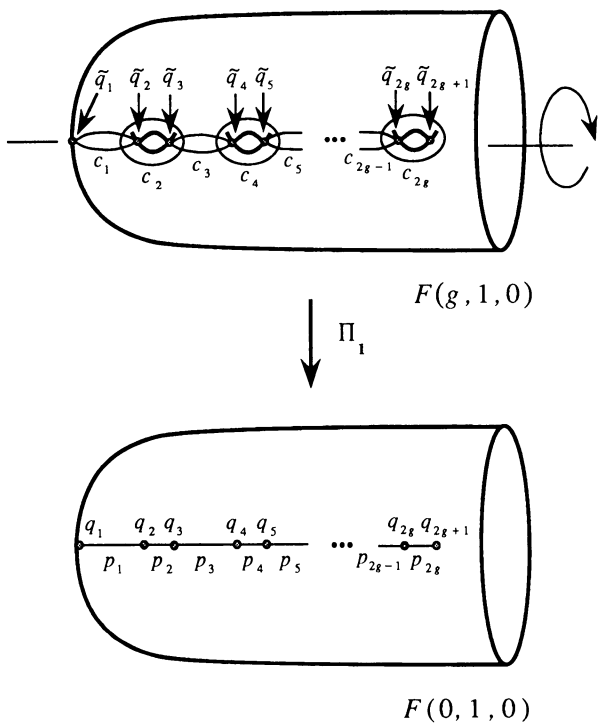

FIGURE 4. Two-fold branched covering $\Pi_i\colon F(g, 1, 0) \to F(0, 1, 0)$

We will consider formal *words* in the letters $\{\widetilde{S}_1, \dots, \widetilde{S}_k\}$, where a word is defined as usual, except we allow only *positive* exponents. We call such a word *complete* if each letter $\widetilde{S}_i$ appears at least once.

For a partition $\mathscr{P} = (m_1, \dots, m_k)$ of $2g$, we define the following semigroups

$$\widetilde{S}(\mathscr{P}) = \{\text{formal words in } \{\widetilde{S}_1, \dots, \widetilde{S}_k\}\}$$

and

$$GR(\mathscr{P}) = \{\widetilde{W} \in \widetilde{S}(\mathscr{P}) \colon \widetilde{W} \text{ a complete word in } \{\widetilde{S}_1, \dots, \widetilde{S}_k\}\}.$$

Each word $\widetilde{W} \in \widetilde{S}(\mathscr{P})$ corresponds to an element $\widetilde{W} \in MC(F(g, 1, 0))$ in the obvious way. The context will make it clear in which sense $\widetilde{W}$ is to be understood. Notice the obvious relations among the $\widetilde{S}_i$:

(iii) $$\widetilde{S}_i\widetilde{S}_j = \widetilde{S}_j\widetilde{S}_i, \quad \text{for } |i-j| \neq 1.$$

To conform with our convention of reading a composition of functions from right to left, we read the letters of a word $W \in \widetilde{S}(\mathscr{P})$ in the same order.

The next lemma asserts that to find the p.A. maps in $\widetilde{S}(\mathscr{P})$ it is enough to consider $\widetilde{GR}(\mathscr{P})$.

Lemma 2. *Suppose that $\mathscr{P}$ is a partition of $2g$, then any $\widetilde{W} \in \widetilde{S}(\mathscr{P}) \setminus \widetilde{GR}(\mathscr{P})$ corresponds to a reducible mapping class, hence is not p.A. by Thurston's theorem.*

Proof. Assume that $\widetilde{W}$ does not contain the letter $\widetilde{S}_i$, for some $i \in \{1, \dots, k\}$. It is easy to find an essential nonboundary, nonpuncture parallel simple closed curve γ embedded in $F(g, 1, 0)$ that intersects c_j only for $j = v(i) + 1$. Thus, γ is a reducing curve for $\widetilde{W}$. □

Associated with a partition $\mathscr{P} = (m_1, \dots, m_k)$ is a "simplest" complete word in $\widetilde{GR}(\mathscr{P})$, namely $\widetilde{W}(\mathscr{P}) = \widetilde{S}_k \cdots \widetilde{S}_1$.

Three problems. To pose the three main problems that we will solve in the sequel we denote by $\widetilde{GR}^*(\mathscr{P})$ the words in $\widetilde{GR}(\mathscr{P})$ that corresponds to p.A. maps and set

$$\widetilde{\mathscr{GR}}^* = \bigcup_{\mathscr{P}} \widetilde{GR}^*(\mathscr{P}),$$

where $\mathscr{P}$ varies over all partitions of $2g$.

Problem 1. Find $\widetilde{\mathscr{GR}}^*$.

Problem 2. Find the map in $\widetilde{\mathscr{GR}}^*$ that realizes the smallest dilatation arising from a word in $\widetilde{\mathscr{GR}}^*$.

Problem 3. Find a set of defining relations for $\widetilde{S}(\mathscr{P})$ and hence for $\widetilde{\mathscr{GR}}(\mathscr{P})$.

Remark. If we define $\widetilde{T}_i$ in the same way as $\widetilde{S}_i$, for $1 \le i \le k$, except we set $\varepsilon(i) = -1, +1$, if i is even, odd respectively, then a word in $\{\widetilde{S}_i \colon 1 \le i \le k\}$ is p.A., if and only if the corresponding word in $\{\widetilde{T}_i \colon 1 \le i \le k\}$ is. This can be seen by changing the orientation of $F(g, 1, 0)$.

3. Simplification of the problems

We discuss in this section two simplifications of the solution to the three problems posed above. Throughout this section, we fix the genus g of $F(g, 1, 0)$, and choose a partition $\mathscr{P} = (m_1, \dots, m_k)$ of $2g$.

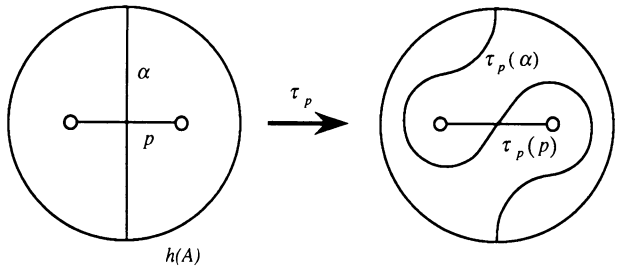

FIGURE 5. Path twist

The hyperelliptic involution. If F is a noncompact surface and p is a path embedded in F with the ends of p contained in 2 different punctures, then we define a *path twist* along p as follows: Take A to be the annulus as in the definition of Dehn twist and let h be a map of A onto a neighborhood of p such that h restricted to $\{x: 1 < |x| \leq 2\}$ is an orientation preserving homeomorphism onto its image and $h(C_1) = p$. Furthermore, if $x = (x_1, x_2) \in C_1$ and $y = (x_1, -x_2)$ then $h(x) = h(y)$. We define a self-homeomorphism r of A as rotation of C_r by $\pi(r-1)$ in the clockwise direction, for $1 \leq t \leq 2$. The path twist τ_p is then defined as $h \circ r \circ h^{-1}$ on $h(A)$ and is the identity map on the complement of $h(A)$. Clearly, τ_p is a homeomorphism of F. For the effect of τ_p on an arc α transverse to p, see Figure 5.

Note that $F(g, 1, 0)$ is invariant under the hyperelliptic involution ι, i.e. rotation by π in the axis shown in Figure 4. The orbit space of ι is a disc $F(0, 1, 0)$, and we have a regular two-fold branched covering

$$\Pi_1 : F(g, 1, 0) \to F(0, 1, 0),$$

which is one-to-one over each point in the branch set

$$Q = \{q_1, \dots, q_{2g+1}\} \subset F(0, 1, 0),$$

where we adopt the notation for points and paths as in Figure 4. We further write

$$\widetilde{Q} = \Pi_1^{-1}(Q) = (\tilde{q}_1, \dots, \tilde{q}_{2g+1}).$$

Denote by $F(g, 1, 2g+1)$ and $F(0, 1, 2g+1)$ the surfaces we get from $F(g, 1, 0)$ and $F(0, 1, 0)$, respectively, by removing the set Q and $\widetilde{Q}$ from the relevant surface. We then have an induced regular unbranched two-fold cover

$$\Pi_2 : F(g, 1, 2g+1) \to F(0, 1, 2g+1).$$

We observe next that we can take the simple closed curves $c(i, j)$, for $1 \leq j \leq m_i$ and $1 \leq i \leq k$ in $F(g, 1, 0)$ to be invariant under ι. Hence, $\tilde{\tau}(i, j)$ can be isotoped to be fiber preserving. We conclude that $\tilde{\tau}(i, j)$ restricts to a homeomorphism of $F(g, 1, 2g+1)$ and hence projects to a map $\tau(i, j)$ of $F(0, 1, 2g+1)$. If we define $c'(i, j)$ as the restriction of $c(i, j)$ to $F(g, 1, 2g+1)$ and $p(i, j) = \Pi_2(c'(i, j))$, then we readily convince ourselves that $\tau(i, j) = \tau_{p(i,j)}^{\varepsilon(i)}$, for $1 \leq j \leq m_i$ and $1 \leq i \leq k$. (Regarding the definition of $\varepsilon(i)$, see the definition of $\widetilde{S}_i$.) In the notation of Figure 4, we have of course

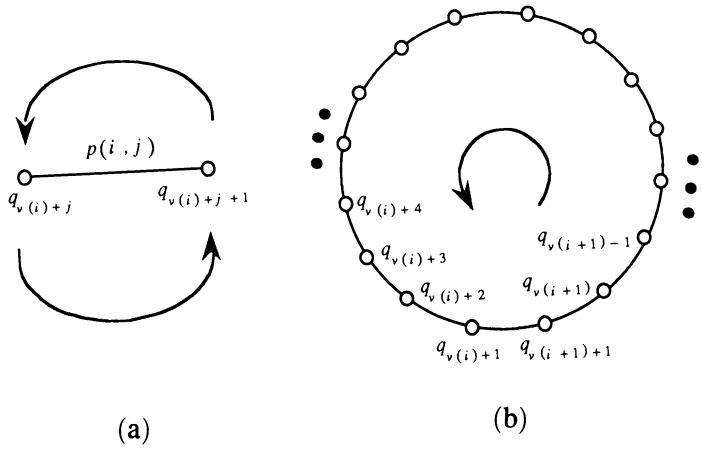

FIGURE 6. Effect of S_i on punctures, in case i odd

$p(i, j) = p_{v(i)+j}$, for $1 \le j \le m_i$ and $1 \le i \le k$. We also define

$$S_i = \tau(i, m_i) \cdots \tau(i, 1), \qquad \text{for } 1 \le i \le k.$$

(Compare to the definition of $\widetilde{S}_i$.)

The expressions $S(\mathscr{P})$, $GR(\mathscr{P})$, $GR^*(\mathscr{P})$, $\mathscr{GR}$, and $\mathscr{GR}^*$ are defined as $\widetilde{S}(\mathscr{P})$, $\widetilde{GR}(\mathscr{P})$, $\widetilde{GR}^*(\mathscr{P})$, $\widetilde{\mathscr{GR}}$, and $\widetilde{\mathscr{GR}}^*$, respectively, except that we use S_i instead of $\widetilde{S}_i$. Similarly, we define

(iv) $$W(\mathscr{P}) = S_k \cdots S_1,$$

analogous to $\widetilde{W}(\mathscr{P})$.

To visualize the effect S_i, for $1 \le i \le k$ has on the surface $F(0, 1, 2g+1)$, note first that for $1 \le j \le m_i$ and $1 \le i \le k$, the path twist $\tau(i, j)$ permutes the punctures $q_{v(i)+j}$ and $q_{v(i)+j+1}$ (see Figure 6(a)). We conclude that for $1 \le i \le k$, S_i cyclically permutes the punctures $\{q_j : v(i)+1 \le j \le v(i+1)+1\}$. We can therefore arrange these punctures on the embedded image of a euclidean circle so that S_i rotates this circle counterclockwise (clockwise) by an angle $2\pi/(m_i + 1)$ in case i is odd (respectively even). Note that the ith circle intersects the $(i+1)$th in exactly one puncture, namely in $q_{v(i+1)+1}$, for $i = 1, \ldots, k-1$. The action of S_i is illustrated in Figure 6(b), in case i odd.

Note that the relations (iii) translate into

(v) $$S_i S_j = S_j S_i, \qquad \text{for } |i - j| \ne 1.$$

The next lemma asserts that to solve our three problems posed at the end of the last section it is enough to restrict attention to $S(\mathscr{P})$ and $\mathscr{GR}$.

Lemma 3. *Suppose that* $\widetilde{W} = \widetilde{S}_{w(n)}^{\alpha(n)} \cdots \widetilde{S}_{w(1)}^{\alpha(1)} \in \mathrm{Homeo}(F(g, 1, 0))$, *and* $W = S_{w(n)}^{\alpha(n)} \cdots S_{w(1)}^{\alpha(1)} \in \mathrm{Homeo}(F(0, 1, 2g+1))$, *where* $1 \le w(i) \le k$, *for* $1 \le i \le n$. *Suppose further that* $W_1 \in GR(\mathscr{P})$ *is related to* $\widetilde{W}_1 \in \widetilde{GR}(\mathscr{P})$ *in the same way as* W *is related to* $\widetilde{W}$, *then*

(a) *For* $\widetilde{W}_1$ *to be isotopic to* $\widetilde{W}$ *it is necessary and sufficient that* W_1 *be isotopic to* W.

(b) W *represents a reducible, periodic or p.A. map if and only if* $\widetilde{W}$ *represents a reducible, periodic or p.A. map, respectively.*

(c) *If* $W \in GR^*(\mathscr{P})$, *then the dilatation* λ *of* W *equals the dilatation* $\tilde{\lambda}$ *of* $\widetilde{W}$.

Proof. The sufficiency statement of part (a) is immediate and the necessity statement follows from [BH] or the more general result in [Zi].

For part (b) we readily convince ourselves that W is periodic or reducible if and only if $\widetilde{W}$ is periodic or reducible, respectively. The claim then follows from Theorem A.

We are left to prove the statement in part (c). Let $(\mathscr{F}, \nu)$ denote the unstable measured foliation of $F(0, 1, 2g+1)$ projectively invariant by W, and let $(\widetilde{\mathscr{F}}, \tilde{\nu})$ denote the lift (and extension) of $(\mathscr{F}, \nu)$ to $F(g, 1, 0)$. Since $\widetilde{W}$ is fiber preserving, $(\widetilde{\mathscr{F}}, \tilde{\nu})$ is projectively invariant by $\widetilde{W}$, and since $\widetilde{W}$ is p.A. (by part (b)), $(\widetilde{\mathscr{F}}, \tilde{\nu})$ is the unstable measured foliation projectively invariant by $\widetilde{W}$. Thus, we may choose a small arc $\alpha \subset F(0, 1, 2g+1)$ transverse to $\mathscr{F}$, take a lift and extension $\tilde{\alpha}$ of α to $F(g, 1, 0)$ and compute

$$\lambda = \frac{\nu(W(\alpha))}{\nu(\alpha)} = \frac{\tilde{\nu}(\widetilde{W}(\tilde{\alpha}))}{\tilde{\nu}(\tilde{\alpha})} = \tilde{\lambda},$$

as was asserted. □

Normal forms. We close this section by defining a normal form for the words in $\mathscr{GR}(\mathscr{P})$.

If $W \in GR(\mathscr{P})$ is given by $W = S_{w(t)} \cdots S_{w(1)}$, where $w(i) \in \{1, \dots, k\}$, for $1 \le i \le t$, then a *subword* W' of W is of the form

$$W' = S_{w(s)} \cdots S_{w(r+1)} S_{w(r)},$$

for some $1 \le r \le s \le t$. The number $s - r + 1$ is called the *length* of W'. If $r = 1$, then we say that W' is an *initial* subword, and if $s = t$, then we say that W' is a *terminal* subword of W. (Recall that we read letters from right to left.) We say for instance that $S_{w(r)}$ appears before $S_{w(s)}$ in W'.

We say that $W \in GR(\mathscr{P})$ is in *normal form*, if:

(a) the first S_i in W appears before the first S_{i+1}, for $1 \le i \le k-1$;
(b) if a letter S_i is immediately followed by S_j, then $j \ge i - 1$;
(c) the last S_1 in W is followed by at least one S_2.

We have

Lemma 4. (a) *Any* W *in* GR *can be altered to be in normal form via cyclic permutation and the relations* (v).

(b) *The unique normal form of a word in* $S_1, \dots, S_k$ *so that each* S_i *occurs exactly once in* $W(\mathscr{P})$.

Sketch of proof. For the proof of part (a) we arrange successively, using induction that each of the conditions in the definition of normal form hold.

Part (b) is a direct consequence of property (a) in the definition of normal form. □

Note that any cyclic permutation of a word in $GR(\mathscr{P})$ can be achieved by conjugation. As the property of a map being p.A. is an invariant of the conjugacy class as is the dilatation and as the relations (v) hold among the homeomorphisms corresponding to words in $GR(\mathscr{P})$, we can, to solve the first

two problems posed in §2 assume that the word $W \in GR(\mathscr{P})$ in question is in normal form.

4. A NECESSARY AND SUFFICIENT CONDITION FOR W TO BE IN $\mathscr{GR}^*$

Throughout this chapter we consider the surface $F = F(g, 1, 2g+1)$, for a fixed $g \geq 1$. Moreover, we fix a partition $\mathscr{P} = (m_1, \ldots, m_k)$ of $2g$ unless we specifically say otherwise.

The strategy to solve the first of the problems posed at the end of the second section will be to find a train track $\tau(\mathscr{P})$ that is invariant under any word $W \in GR(\mathscr{P})$. For a given $W \in GR(\mathscr{P})$, an analysis of the incidence matrix of W with respect to the branches of $\tau(\mathscr{P})$ will show that we can delete certain of the branches of $\tau(\mathscr{P})$ and still get an W invariant train track $\tau^*(W, \mathscr{P})$. This train track will satisfy all the conditions of Casson's theorem, except it might not fill the surface. There is a simple number theoretic condition under which $\tau^*(W, \mathscr{P})$ fills F, hence under which W corresponds to a p.A. map.

Consider the track $\overline{\tau} = \overline{\tau}(\mathscr{P})$ as illustrated in Figure 7 in case $k = 3$. For $1 \leq i \leq k$, we set $l_i = m_i + 1$ and adopt the notation for some of the branches of $\overline{\tau}$ as shown in the same figure in case $k = 3$. It is clear how $\tau(\mathscr{P})$ is defined and what our choice of notation is in case $k = 2, 4, 5, \ldots$. To be more specific, if $2 \leq i \leq k-1$, then the branches $y(i-1, l_{i-1})$, and $y(i, 1)$ bound a mono-gon whose interior contains the puncture $q_{v(i)+1}$. For $1 \leq i \leq k$ and $2 \leq j \leq l_i - 1$, the branch $y(i, j)$ bounds a mono-gon whose interior contains the puncture $q_{v(i)+j}$. Finally, the branches a_1 and $y(1, 1)$ bound a mono-gon whose interior contains the puncture q_1 and the branches $y(k, l_k)$ and a_2 bound a mono-gon whose interior contains the puncture q_{2g+1}. Note that $\overline{\tau}(\mathscr{P})$ has as complementary regions a l_i-gon, for $1 \leq i \leq k$, further $2g+1$ punctured mono-gons and a $(k-1)$-gon with a disk removed. In particular, for each i such that $m_i = 1$, $\overline{\tau}$ has a complementary bi-gon and we readily see that $\overline{\tau}(\mathscr{P})$ is a bi-gon track.

In view of Casson's theorem we like to modify $\overline{\tau}$ to get a train track. For each complementary bi-gon of $\overline{\tau}$ we identify the two smooth arcs that make up the boundary of the bi-gon. See Figure 8 for an illustration. If we remove in this way all complementary bi-gons we get a train track $\tau = \tau(\mathscr{P})$ as illustrated in Figure 9 (the notation for branches of τ as shown in the figure will be explained momentarily). Note that each branch $y(i, j)$ of $\overline{\tau}$, for which $l_i > 2$ corresponds to a unique branch of τ in the obvious way, and we call this branch $x(i, j)$. If $l_i = 2$, then the two branches $y(i, 1)$ and $y(i, 2)$ combine to a branch, called $x(i, 1)$, of τ. For $1 \leq i \leq k$, we set $n_i = m_i$,

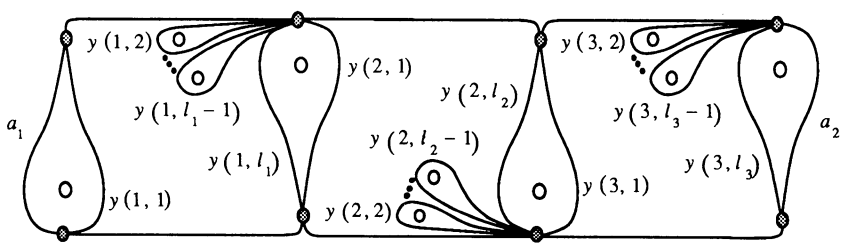

FIGURE 7. $\overline{\tau}(\mathscr{P})$, for $k = 3$

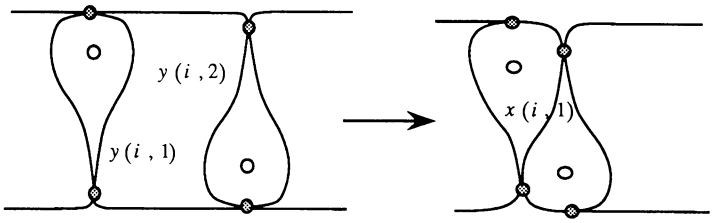

FIGURE 8. Removing a complementary bi-gon

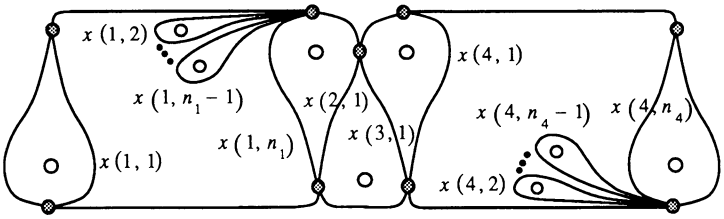

FIGURE 9. $\tau(\mathscr{P})$, for $\mathscr{P} = (m_1, 1, 1, m_4)$

if $m_i = 1$ and $n_i = m_i + 1$ $(= l_i)$, otherwise. For an example, suppose that $\mathscr{P}(m_1, 1, 1, m_4)$, where $m_1 = m_4 > 1$, then $\tau(\mathscr{P})$, together with the notation for certain branches of $\tau(\mathscr{P})$ is illustrated in Figure 9.

We will need the following subsets of the set of branches of $\tau(\mathscr{P})$. Set

$$X(i) = \{x(i, j) \colon 1 \le j \le n_i\}, \qquad \text{for } 1 \le i \le k,$$

$$X = \bigcup_{i=1}^{k} X(i) \quad \text{and} \quad X^c = \{\text{branches of } \tau\} \setminus X.$$

We next choose $i \in \{1, \dots, k\}$ and specify how $S_i \in S(\mathscr{P})$ acts on $\tau(\mathscr{P})$. We perform the path twists in the definition of S_i so that only the branches of X are moved, in particular the vertices of $\tau(\mathscr{P})$ remain fixed. We also choose the central ties corresponding to a standard tie neighborhood of $\tau(\mathscr{P})$ such that the following holds. Suppose that $i \in \{2, \dots, k-1\}$ is such that $m_i > 1$, then a local picture of the subset of $\tau(\mathscr{P})$ that is affected by S_i is as shown in Figure 10(a). Using the notation for two of the branches of X^c as in the same figure, we may assume that $S_i(x(i, 1))$ intersects the central tie corresponding to $x(i-1, n_{i-1})$, a_i' (twice), $x(i+1, 1)$, and $x(i, n_i)$; also $S_i(x(i, 2))$ intersects the central ties corresponding to a_i (twice), $x(i-1, n_{i-1})$, and $x(i, 1)$; for $2 < j < n_i$, $S_i(x(i, j))$ intersects the central tie corresponding to $x(i, j-1)$; finally, $S_i(x(i, n_i))$ intersects the central ties corresponding to $x(i+1, 1)$ and $x(i, n_i - 1)$. S_i fixes all the other branches. If $i \in \{2, \dots, k-1\}$ is such that $m_i = 1$, then the local picture of branches affected by S_i is as in Figure 10(b). Note that the figure is general enough so that we do not have to distinguish between the cases $m_t > 1$ and $m_t = 1$, where $t = i-1, i+1$. In the notation of the figure we may assume that $S_i(x(i, 1))$ intersects the central ties corresponding to $x(i-1, n_{i-1})$, a_i, $x(i, 1)$, a_i', and $x(i+1, 1)$. The cases $i = 1, k$ are similar.

We conclude:

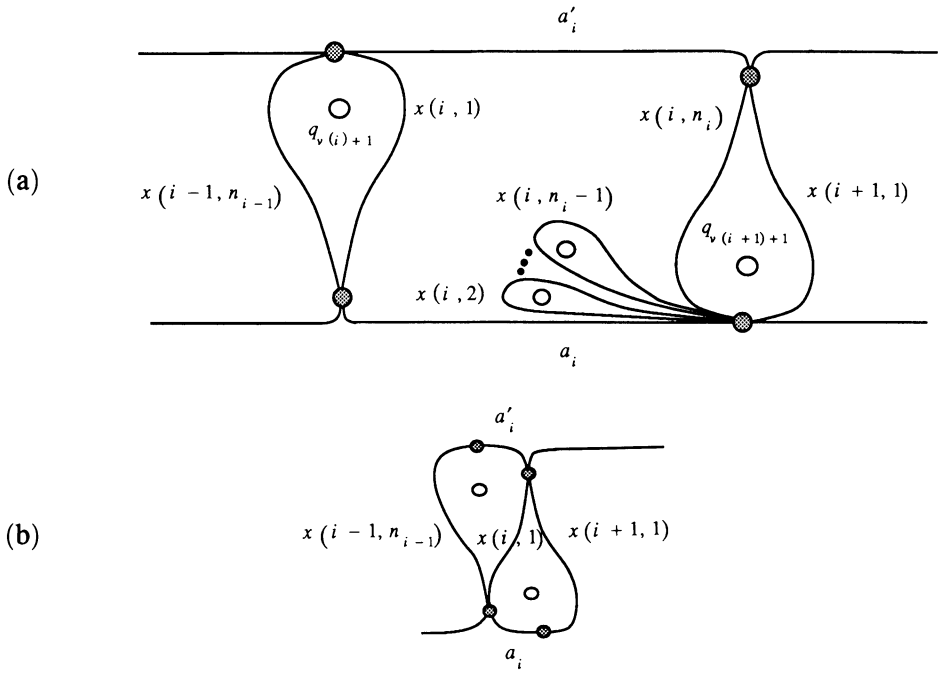

FIGURE 10. Local figure of the branches affected by S_i

Lemma 5. *Suppose that $W \in GR(\mathscr{P})$, then $\tau = \tau(\mathscr{P})$ is a train track that is invariant under S_i, for $1 \le i \le k$, and as the relation "is carried by" is transitive, $\tau(\mathscr{P})$ is invariant under W.*

In view of the last lemma, we can define the incidence matrix M_i that corresponds to S_i and some ordering of the branches of $\tau(\mathscr{P})$, for $1 \le i \le k$.

If $W \in GR(\mathscr{P})$, then the incidence matrix M corresponding to W and $\tau(\mathscr{P})$ is a product of M_i, where $1 \le i \le k$. This is a consequence of the distributive property of the 'hat' operator. We apply to M the notion of initial subword and length of a subword as introduced in §3 for subwords of W. For L some large number, and $1 \le n \le L \times$ (length of M), denote by $\mathrm{Init}(n)$, the initial subword of M^L of length n, and by $\mathrm{Let}(n)$ the nth letter of M^L (as usual reading letters of M^L from right to left).

As a first step in the analysis of the incidence matrix we show

Lemma 6. *Denote by ρ a weight on $\tau(\mathscr{P})$ that is positive on a branch $b \in X$. Then, for $2 \le i \le k$ the following assertion, denoted* Supp(i) *holds: there exists an N_i such that for all $n > N_i$,* $\mathrm{Init}(n)(\rho)$ *is either positive on $x(i-1, n_{i-1})$ or $x(i, 1)$, where the first case occurs if the last occurrence of a letter M_i in* $\mathrm{Init}(n)$ *is after the last occurrence of a letter M_{i-1}, and the last case occurs otherwise.*

Proof. We begin by stating two basic facts that follow immediately from our convention of how $\tau(\mathscr{P})$ is carried by S_i. Namely, if δ is a weight on $\tau(\mathscr{P})$ that is positive on $x(i, 1)$, then so is $M_t\delta$, for $t \ne i$, whereas $M_i\delta$ is positive on $x(i-1, n_{i-1})$. Similarly, if δ is a weight on $\tau(\mathscr{P})$ that is positive on $x(i-1, n_{i-1})$, then so is $M_t\delta$, for $t \ne i-1$, whereas $M_{i-1}\delta$ is positive on $x(i, 1)$.

Assume now that $r \in \{1, \dots, k\}$ is such that ρ (as in the lemma) is positive

on some branch of $X(r)$. We may assume that $r > 1$ as in case $r = 1$, we find some n such that $\rho' = \mathrm{Init}(n)\rho$ is positive on $x(2, 1) \in X(2)$. We then prove the statement of the lemma for ρ' which clearly implies the statement of the lemma for ρ.

We prove the assertion of the lemma by induction on the number of $i \in \{2, \dots, k\}$ such that $\mathrm{Supp}(i)$ holds. For the basis step, we find the smallest integer N_r such that $\mathrm{Init}(N_r)\rho$ is positive on $x(r-1, n_{r-1})$. Clearly, $\mathrm{Let}(N_r) = M_r$ and it follows from the two basic facts stated in the beginning of the proof that $\mathrm{Supp}(r)$ holds.

For the induction step we first assume that $\mathrm{Supp}(j)$ holds for some $j \in \{2, \dots, k-1\}$, but not $\mathrm{Supp}(j+1)$. Suppose now that $m > N_j$ is such that $\mathrm{Let}(m) = M_{j-1}$, then $\mathrm{Supp}(j)$ implies that $\mathrm{Init}(m)\rho$ is positive on $x(j, 1)$. If m' is the smallest integer such that $m' > m$ and $\mathrm{Let}(m') = S_j$, then $\mathrm{Init}(m')\rho$ is positive on $x(j+1, 1) \in X(j+1)$. It follows as in the basis step from the two basic facts that $\mathrm{Supp}(j+1)$ holds for $N_{j+1} = m'$. We next assume that $\mathrm{Supp}(j+1)$ holds but not $\mathrm{Supp}(j)$. As $\mathrm{Supp}(j+1)$ holds we can find $n > N_j$, such that $\mathrm{Init}(n)\rho$ is positive on $x(j, n_j)$, hence there exists $n' > n$, such that $\mathrm{Init}(n')\rho$ is positive on $x(j-1, n_{j-1})$. Clearly, $\mathrm{Let}(n') = S_j$, and it follows again from the two basic facts that $\mathrm{Supp}(j)$ holds for $N_j = n'$. This constitutes the induction step and proves the claim. □

We now assume that $W \in GR(\mathscr{P})$ is in normal form. For $1 \le i \le k$, we define $W(i)$ as the word which we get from W by deleting all letters S_v, for $v \ne i, i+1$. After amalgamating adjacent letters, we have for $1 \le i < k$,

$$\text{(vi)} \qquad W(i) = S_{i+1}^{\beta(i,p_i)} S_i^{\alpha(i,p_i)} \cdots S_{i+1}^{\beta(i,1)} S_i^{\alpha(i,1)},$$

where $\alpha(i, j) > 0$, for $2 \le j \le p_i$ and $\beta(i, j) > 0$, for $1 \le j < p_i$. Moreover, $\beta(i, p_i) \ge 0$ and condition (a) of the definition of normal form implies that $\alpha(i, 1) > 0$. For the convenience of the reader we spell out $W(i-1)$, where $1 \le i-1 < k$:

$$\text{(vii)} \qquad W(i-1) = S_i^{\beta(i-1,p_{i-1})} S_{i-1}^{\alpha(i-1,p_{i-1})} \cdots S_i^{\beta(i-1,1)} S_{i-1}^{\alpha(i-1,1)}.$$

For $1 \le i \le k$, we define the *exponent sum* $\sigma_i = \sigma_i(W)$ of S_i in W as

$$\sigma_i(W) = \alpha(i, 1) + \cdots + \alpha(i, p_i) = \beta(i-1, 1) + \cdots + \beta(i-1, p_{i-1}),$$

(if the terms are defined; of course expressions of the form $\alpha(k, j)$ and $\alpha(0, j)$ are not defined).

We denote by $M(i)$ the incidence matrix corresponding to $W(i)$, for $1 \le i \le k$ and by M the incidence matrix corresponding to W (with respect to some order of the branches of τ). Of course we get $M(i)$ from M by deleting the letters M_v, for $v \ne i, i+1$.

In view of the preceding lemma, we assume that ρ is a weight that is either positive on $x(i-1, n_{i-1})$ or $x(i, 1)$, for $2 \le i \le k$, but zero on all other branches. We like to describe next the branches of X on which $M^L\rho$ is positive, for L large, and to the end, we choose $i \in \{1, \dots, k\}$. It is easy to see (or will become clear from what follows) that if $i \in \{2, \dots, k-1\}$, then to find the branches of $X(i)$ on which $M^L\rho$ is positive, we only need to consider $W(i)^L$ and $W(i-1)^L$. Similarly, to find the branches of $X(i)$ on which $M^L\rho$ is positive in case $i = 1$ or $i = k$, we only need to consider $W(1)^L$ or $W(k)^L$, respectively.

If $i \in \{1, \dots, k-1\}$, we apply $M(i)$ and we first assume that $\beta(i, p_i) = 0$. By the two basic facts stated in the proof of the preceding lemma we know that for $v \in \{1, \dots, p_i - 1\}$, $M_{i+1}^{\beta(i,v)}\rho$ is positive on $x(i, n_i)$, hence $M(i)\rho$ is positive on $x(i, s)$, where $s = n_i - (\alpha(i, v+1) + \cdots + \alpha(i, p_i)) \pmod{n_i}$. If $\beta(i, p_i) > 0$, then we have to consider the case $v = p_i$. We see that $M(i)\rho$ is also positive on $x(i, s)$, where $s = n_i - 0$.

If $i \in \{2, \dots, k\}$, we apply $M(i-1)$. An argument similar to the one in the last paragraph shows that if $1 \le v \le p_{i-1}$, then $M_{i-1}^{\alpha(i-1,v)}\rho$ is positive on $x(i, 1)$, hence $M(i-1)\rho$ is positive on $x(i, s)$, where $s = 1 - (\beta(i-1, v) + \cdots + \beta(i-1, p_{i-1})) \pmod{n_i}$.

We therefore define for $1 \le i \le k-1$, using the powers of S_i in $W(i)$

$$R'(i) = \begin{cases} \{\alpha(i, p_i), (\alpha(i, p_i - 1) + \alpha(i, p_i)), \dots, (\alpha(i, 2) + \cdots + \alpha(i, p_i))\} \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } \beta(i, p_i) = 0; \\ \{0, \alpha(i, p_i), \alpha(i, p_i - 1) + \alpha(i, p_i)), \dots, (\alpha(i, 2) + \cdots + \alpha(i, p_i))\}, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } \beta(i, p_i) > 0, \end{cases}$$

and for $2 \le i \le k$, we define using the powers of S_i in $W(i-1)$,

$$L'(i) = \{\beta(i-1, p_{i-1}), (\beta(i-1, p_{i-1} - 1) + \beta(i-1, p_{i-1})), \dots, (\beta(i-1, 2) + \cdots + \beta(i-1, p_{i-1}))\}.$$

Set

$$R(i) = \{n_i - m\sigma_i - \alpha \colon m \ge 0;\ \alpha \in R'(i)\} \pmod{n_i}, \quad \text{for } 1 \le i \le k-1;$$

$$L(i) = \{1 - m\sigma_i - \beta \colon m \ge 0;\ \beta \in L'(i)\} \pmod{n_i}, \quad \text{for } 2 \le i \le k,$$

and

$$R(k) = L(1) = \varnothing.$$

Finally,

$$T(i) = R(i) \cup L(i) \subset \{1, \dots, n_i\},$$

for $1 \le i \le k$.

Remark. Of course, we can restrict m in the definition of $R(i)$ and $L(i)$ to $n_i - 1 \ge m \ge 0$.

Returning to our analysis, we can state, using the notation just introduced, that $M(i)\rho$ is positive on branches $x(i, n_i - w)$, where $w \in R'(i)$, and hence $M(i)^L\rho$, for L large enough, is positive on the branches $x(i, v)$, for $v \in R(i)$. Similarly, $M^L\rho$ is positive on $x(i, v)$, for $v \in L(i)$. Together, we conclude that $M^L\rho$, for L sufficiently large, is positive on the branches $x(i, v)$, for $v \in T(i) = R(i) \cup L(i)$.

We define

$$X^* = X^*(W, \mathscr{P}) = \bigcup_{i=1}^{k} \{x(i, j) \colon j \in T(i)\},$$

and prove the crucial

Lemma 7. *We denote the incidence matrix of $W \in GR(\mathscr{P})$ (with respect to some ordering of the branches of τ) by M and claim*

(a) *If ρ is a weight on $\tau(\mathscr{P})$ that is positive on some branch of X, then there exists $L > 0$, such that for $l > L$, $M^l\rho$ is positive on all branches of $X^c \cup X^*$.*

(b) *If* $\rho \in V(\tau)$ *is positive only on branches of* $X^c \cup X^*$, *then so is* $M\rho$.

(c) W *indices a permutation of the branches of* $X \setminus X^*$. *To be more explicit, if* ρ *is a weight that is positive on a single branch of* $X \setminus X^*$, *then there is a unique branch of* $X \setminus X^*$ *on which* $M\rho$ *is positive.*

Proof. To show part (a), we first apply Lemma 6 to assure that $M^n\rho$, for n large enough, is positive on $x(i-1, n_{i-1})$ or $x(i, 1)$, for $2 \le i \le k$. But then the argument preceding Lemma 7 shows that $M^n\rho$ is positive on branches of X^*, for n large enough.

We next prove part (b). If ρ is a measure such that $M\rho$ is positive on a branch $x(i, j)$ of X, then there are two cases of how this can happen. The first case is as described before Lemma 7, in which case $j \in T(i)$, and hence $x(i, j) \in X^*$. The second case is that ρ is positive on $x(i, j+\sigma_i)$. By hypothesis $j + \sigma_i \in T(i)$, hence $j \in T(i)$ as desired.

This also proves part (c). Indeed, if ρ is positive on a branch $x(i, j) \in X \setminus X^*$, then the only branch of $X \setminus X^*$ on which $M\rho$ is positive is $x(i, j-\sigma_i)$, which defines the desired correspondence. □

The previous lemma suggests that corresponding to $W \in GR(\mathscr{P})$, we define the branched submanifold $\tau^* = \tau^*(W, \mathscr{P})$ of $\tau(\mathscr{P})$ that consists of the branches $X^c \cup X^*$ together with the appropriate switches. The action of S_i, for $1 \le i \le k$ on τ^* is defined to be the restriction of the action of S_i on $\tau(\mathscr{P})$ to τ^*.

The next lemma asserts that τ^* is in fact a train track.

Lemma 8. *Suppose that* $W \in GR(\mathscr{P})$ *and* $\tau^* = \tau^*(W, \mathscr{P})$, *then*

(a) τ^* *is an* W *invariant train track.*

(b) *For* τ^* *to fill the surface it is necessary and sufficient that* $\# T(i) \ge m_i$, *for* $1 \le i \le k$, *where* $\# D$ *denotes the cardinality of the set* D.

(c) *If* $\mathscr{P} = (g, g)$ *and* $W = W(\mathscr{P}) = S_2S_1$, *then* $\tau^*(W, \mathscr{P}) = \tau(\mathscr{P})$.

Proof. To show that τ^* is a train track, we note that we get τ^* from τ by deleting certain of the branches of $X \setminus X^*$. It therefore suffices to show that τ^* had no dead ends, which is clearly guaranteed if we can show that for $i = 2, \dots, k$, either $x(i, 1)$ or $x(i-1, n_{i-1})$ is a branch of τ^*. To that end we take $W(i)$ as in equation (vi) and consider two cases. If $\beta(i-1, p_{i-1}) = 0$, we conclude that $0 \in L'(i)$, and taking $m = 0$ and $\beta = 0$ in the definition of $L(i)$, we see that $1 \in L(i) \subset T(i)$, hence $x(i, 1)$ is a branch of τ^*. If however $\beta(i-1, p_{i-1}) > 0$, then $0 \in R'(i-1)$, and taking $m = 0$ and $\alpha = 0$ in the definition of $R(i-1)$, we see that $n_{i-1} \in R(i-1) \subset T(i-1)$, hence $x(i-1, n_{i-1})$ is a branch of τ^*. We showed that τ^* is a train track.

The fact that τ^* is W invariant follows from the fact that τ is W invariant and Lemma 7(b).

For the proof of part (b), we note that the condition on complementary regions in the definition of train track implies that no complementary region can contain more than one puncture. We choose $i \in \{1, \dots, k\}$ and assume first that $m_i > 1$, hence we are in the situation of Figure 10(a). We readily see that τ^* must contain at least m_i $(= n_i - 1)$ of the branches $x(i, 1), \dots, x(i, n_i)$ as desired. If however $m_i = 1$, (see Figure 10(b)), then τ^* must contain the branch $x(i, 1)$, but then $1 = \# T(i) = m_i$. This proves part (b).

For part (c) we note that the exponent sum σ_i of S_i in W equals 1, for

$i = 1, 2$, and if follows that $\# T(i) = n_i$. The proof of the lemma is now complete. □

For the rest of this section we assume that, corresponding to $W \in GR(\mathscr{P})$, the branches $\{b_i : 1 \leq i \leq t(3)\}$ of $\tau(\mathscr{P})$ are ordered in such a way that $X^* = \{b_i : 1 \leq i \leq t(1)\}$, and $X = \{b_i : 1 \leq i \leq t(2)\}$, for some integers $1 \leq t(1) \leq t(2) < t(3)$.

We will need the following facts concerning various incidence matrices.

Lemma 9. *For $w \in GR(\mathscr{P})$, denote by* $\mathrm{Inc}(\tau)$ ($\mathrm{Inc}(X)$, $\mathrm{Inc}(X^*)$ *and* $\mathrm{Inc}(\tau^*)$) *the incidence matrix corresponding to W and the branches of τ (respectively the branches X, X^* and the branches of τ^*). Then,*

(a)

$$\mathrm{Inc}(\tau) = \begin{pmatrix} \mathrm{Inc}(X) & 0_1 \\ D_1 & I_1 \end{pmatrix} \in \mathrm{Mat}_{t(3)}(\mathbb{Z}),$$

where 0_1 is the $t(2) \times (t(3) - t(2))$ zero matrix, and I_1 is the $(t(3) - t(2)) \times (t(3) - t(2))$ identity matrix. (We do not need to know about the structure of D_1.) Moreover,

(b)

$$\mathrm{Inc}(X) = \begin{pmatrix} \mathrm{Inc}(X^*) & D_2 \\ 0_2 & P \end{pmatrix} \in \mathrm{Mat}_{t(2)}(\mathbb{Z}),$$

where 0_2 is the $(t(2)-t(1)) \times t(1)$ zero matrix, and P is a $(t(2)-t(1)) \times (t(2)-t(1))$ permutation matrix.

(c)

$$\mathrm{Inc}(\tau^*) = \begin{pmatrix} \mathrm{Inc}(X^*) & 0_3 \\ D_3 & I_3 \end{pmatrix} \in \mathrm{Mat}_{t(1)+t(3)-t(2)}(\mathbb{Z}),$$

where 0_3 is the $t(1) \times (t(3) - t(2))$ zero matrix, and I_3 is the $(t(3) - t(2)) \times (t(3) - t(2))$ identity matrix.

(d) $\mathrm{Inc}(X^*) \in \mathrm{Mat}_{t(1)}(\mathbb{Z})$ *is Perron-Frobenius.*

Proof. The subblocks 0_1 and I_1 of $\mathrm{Inc}(\tau)$ and the subblocks O_3 and I_3 of $\mathrm{Inc}(\tau^*)$ are explained by the fact that W fixes the branches of $\tau \backslash X$ and $\tau^* \backslash X^*$, respectively.

The 0_2 subblock of $\mathrm{Inc}(X)$ is justified by Lemma 7(b) and the P subblock is justified by Lemma 7(c).

We are left to show that $\mathrm{Inc}(X^*)$ is P.F. For $1 \leq i \leq t(1)$, denote by $e_i \in \mathbb{R}^{t(1)}$ the ith unit vector. Of course it suffices to show that there exist an $n > 0$ such that for $1 \leq i \leq t(1)$ we have $\mathrm{Inc}(X^*)^n e_i > 0$. To show that, we first extend e_i to a weight $\overline{e}_i$ on $\tau^*(W, \mathscr{P})$, where $\overline{e}_i \in \mathbb{R}^{t(1)+t(3)-t(2)}$ assigns the same values to branches of X^* as e_i, and assigns the value zero to branches of $\tau^* \setminus X^*$. Lemma 7(a) then implies that for n large enough, $\mathrm{Inc}(\tau^*)^n \overline{e}_i$ is positive on branches of X^*, which implies, using part (c), that $\mathrm{Inc}(X^*)^n e_i > 0$ as desired. □

We remark in passing that it is not necessarily true that $S_i(\tau^*)$ is carried by τ^*, for $1 \leq i \leq k$. □

We can now solve the first problem posed at the end of §3.

Theorem 10. *Suppose that* $\mathscr{P} = (m_1, \dots, m_k)$ *is a partition of* $2g$ *and* $W \in GR(\mathscr{P})$ *is in normal form, then*

$$W \in GR^*(\mathscr{P}) \Leftrightarrow \# T(i) \geq m_i, \quad \text{for } 1 \leq i \leq k.$$

Before proceeding with the proof of Theorem 10, we list three special cases.

Corollary 11. *For* $\mathscr{P}$ *and* W *as in the theorem, we have*:

(a) *If* W *is such that* $\gcd(n_i, \sigma_i) = 1$, *for* $1 \leq i \leq k$, *then* $W \in GR^*(\mathscr{P})$.

(b) $W(\mathscr{P}) \in GR^*(\mathscr{P})$.

(c) *If* $\mathscr{P}_1 = (1, \dots, 1) \in \mathbb{Z}^{2g}$, *then* $GR(\mathscr{P}_1) = GR^*(\mathscr{P}_1)$.

Remarks. (a) $W(\mathscr{P})$ is the map that was shown to be p.A. in [GK]. It arises as the monodromy of a Murasugi sum of Hopf-bands.

(b) [P1] gives a recipe for constructing p.A. maps that are defined as compositions of Dehn twists along certain curves. The recipe applied to the curves c_i, for $1 \leq i \leq 2g$ yields the maps $GR^*(\mathscr{P}_1)$ which is a proper subset of $\mathscr{GR}^*$. [P1] requires that if a curve c intersects d, we twist in different directions along c and d. On the other hand the recipe in [P1] applies to Dehn twists along curves other than the collection $\{c_i : 1 \leq i \leq 2g\}$ that we consider in this investigation.

(c) $\mathscr{P}_1 = (1, 1)$ is the only permutation of $2g$ in case $g = 1$ and it is a folk theorem that, in our notation, each p.A. map ϕ of $F(1, 1, 0)$ is conjugate (up to the hyperelliptic involution) to an element ϕ' of $\widetilde{GR}(\mathscr{P}_1) = \widetilde{GR}^*(\mathscr{P}_1)$. Moreover, the invariant projective measure of ϕ' on an appropriate train track is determined by a positive real number x and both x and the dilatation of ϕ' have a periodic continued fraction expansion with period related to the powers of ϕ', written as a word in $\widetilde{S}_1$ and $\widetilde{S}_2$. This can be generalized as follows. Each p.A. map ϕ of a hyperbolic surface has an invariant train track τ with P.F. incidence matrix. If $\phi(\tau)$ is carried by τ in a special way, then the dilatation of ϕ and the invariant measure on τ can be expressed in continued fraction form, using generalized continued fractions in the sense of Perron and Jacobi (see [Be, Hu]). This is done in [B3], and it is likely that this technique can be extended to include *all* p.A. maps.

Proof of Corollary 11. For part (a) we note that for $1 \leq i \leq k$, $\#\{m\sigma_i : m \geq 0\}$ (mod n_i) $= n_i$, if n_i and σ_i are relatively prime, hence $\#\{T(i)\} = n_i$. Part (b) follows from part (a) as the exponent sum σ_i of S_i in $W(\mathscr{P})$ satisfies $\sigma_i = 1$, for $1 \leq i \leq k$. For part (c) we remark that $n_i = m_i = 1$, hence $\# T(i) = 1 = m_i$, for $1 \leq i \leq 2g$. □

Proof of Theorem 10. Corresponding to $W \in GR(\mathscr{P})$, we construct $\tau^* = \tau^*(W, \mathscr{P})$. We check the conditions of Casson's theorem and begin by noting that Lemma 8(a) asserts that τ^* is an W invariant train track.

We next assume that $\tau' = \tau'(W, \mathscr{P})$ is a subtrack of τ^* that is invariant under W. We will show that τ' cannot be a proper subtrack of τ. We choose a nonzero measure μ' on τ' and extend μ' to a nonzero measure $\mu \in V(\tau^*)$ by defining $\mu(b) = \mu'(b)$, if b is a branch of τ' and $\mu(b) = 0$, otherwise. It is easy to see that any nonzero measure on τ^* is positive on some branch of X^*. By Lemma 7(a), we can find $n > 0$ such that $\widehat{W}^n(\mu)$ is positive on all branches of τ^* ($\widehat{W}$ being the self-map of $V(\tau^*)$ that is induced by W). As

we assumed that τ' is invariant under W, and as μ is zero on branches of $\tau^* \setminus \tau'$, the measure $\widehat{W}^n(\mu)$ must also be zero on these branches. If follows that $\tau' = \tau^*$ as desired.

We next show that if τ^* is itself a subtrack of some train track $\tau'' = \tau''(W, \mathscr{P})$ then $\widehat{W}'': V(\tau'') \to V(\tau'')$ has no nonzero fixed point. We first assume that $\tau'' = \tau$, hence we need to show that $\widehat{W}$ has no nonzero fixed point. Assume to derive a contradiction that $\mu \in V(\tau^*)$ is a nonzero measure such that $\widehat{W}\mu = \mu$. As before, we note that μ must be positive on some branch of X^*, hence Lemma 7(b) implies that μ is positive on all branches of τ^*. If we represent μ by $x \in \mathbb{R}^{t(1)+t(3)-t(2)}$ then we can write $\mathrm{Inc}(\tau^*)x = x$. Using Lemma 9(c), we see that x restricts to a positive fixed point $\overline{x}$ of $\mathrm{Inc}(X^*)$. But as $\mathrm{Inc}(X^*)$ is integral and, by Lemma 7(d), P.F., it follows that for some n, $\mathrm{Inc}(X^*)^n \geq 1$. This however contradicts the existence of a fixed point of $\mathrm{Inc}(X^*)$ and we conclude that our assumption of $\widehat{W}$ having a fixed point was absurd.

We next assume that τ^* is a proper subtrack of τ'' and, to derive a contradiction, that μ'' is a nonzero fixed point of $\widehat{W}''$. If μ'' is zero on branches of $\tau'' \setminus \tau^*$, then μ'' restricts to a nonzero fixed point of $\widehat{W}$, which is absurd as we just showed. We may therefore assume that there is a branch b of $\tau'' \setminus \tau^*$ such that μ'' is positive on b. We will show that for some $n > 0$ and $i \in \{1, \dots, k\}$, $(\widehat{W}'')^n \mu''$ is positive on $x(s-1, n_{s-1})$ or $x(s, 1)$, for $s = i$ and $s = i+1$. To that end, we recall first that complementary regions of τ^* are once punctured mono-gons, various n-gons with k punctures removed, where $n \geq 2$ and $k \geq 0$, and finally one $(k-1)$-gon with one open disk removed. As b cannot be contained in a punctured mono-gon we assume first that b is contained in a k times punctured n-gon, so for some $i \in \{1, \dots, k\}$, b is contained in a complementary region as shown in Figure 10(a) or (b), where $n_i - \# T(i)$ of the branches of $X(i)$ are deleted. We only deal with the first case as the second case is similar but simpler. Recall that τ^* contains either the branch $x(i-1, n_{i-1})$ or $x(i, 1)$, so we readily convince ourselves that b must intersect at least one of the paths $p(i, j)$, where $1 \leq j \leq m_i$. After possibly isotoping τ'' we may assume that each endpoint of b is a vertex of $\tau^* \subset \tau''$, hence remains fixed under S_i and that b after the isotopy intersects the same paths $p(i, j)$ as before the isotopy. It follows now rather easily from the way S_i permutes certain punctures as described earlier (and as illustrated in Figure 6(b)) that $W^n(b)$, for some n, intersects the central tie corresponding to $x(s-1, n_{s-1})$ or $x(s, 1)$, for $s = i$ and $s = i+1$. Indeed, in case both of the endpoints of b agree with the endpoint, call it v, of $x(i, 2)$, we define branches $x'(i, j)$, for $1 \leq j \leq n_i$, as follows. The endpoints of $x'(i, j)$, for $1 \leq j \leq n_i$, are v and $x'(i, j) = x(i, j)$, for $2 \leq j \leq n_i - 1$. Furthermore, $x(i, 1)$ is a mono-gon that enclosures the puncture $q_{v(i)+1}$ in its interior and whose boundary can be isotoped into the union of the branches $\{x(i-1, n_{i-1}), x(i, 1), a_i\}$, and similarly, $x'(i, n_i)$ is a mono-gon that encloses the puncture $q_{v(i+1)+1}$ in its interior and whose boundary can be isotoped into the union of the branches $\{x(i, n_i), x(i+1, 1)\}$. We may then think of b (up to homotopy) as a union of branches that are isotopic to $x'(i, j)$, for some $1 \leq j \leq n_i$. The claim follows by an argument along the lines of the basis step in the induction proof of Lemma 6. The other cases are similar.

We next consider the case where b is contained in the $(k-1)$-gon with an open disk removed. We easily convince ourselves that μ'' must be positive on a branch of X or on a branch of $\tau'' \setminus \tau^*$. In the first case the claim follows again from the basis step of the proof of Lemma 6, and the second case was treated in the last paragraph.

We showed that for some n and some $1 \leq i \leq k$, $(\widehat{W''})^n \mu''$ is positive on $x(s-1, n_{s-1})$ or $x(s, 1)$, for $s = i$ and $s = i+1$. It follows from Lemma 7(a) that after possibly increasing n, $\mu'' = (\widehat{W''})^n \mu''$ is positive on the branches of τ^*. μ'' therefore restricts to a positive fixed point of $\widehat{W}$ which is absurd as we saw before. This contradiction shows that $\widehat{W''}$ cannot have a nonzero fixed point.

We conclude from Casson's theorem that W is p.A. if and only if τ^* fills the surface. This however is equivalent to the condition stated in the theorem as follows from Lemma 8(b). □

5. The smallest dilatation arising from $W \in \mathscr{GR}^*$

In this section we address the problem of finding the smallest dilatation arising from a word in $\mathscr{GR}^*$. All proofs in this section are rather rough outlines.

Incidence matrices. We choose for this subsection a fixed partition $\mathscr{P}$ of $2g$. Corresponding to $W \in GR(\mathscr{P})$, we use the notation $\mathrm{Inc}(C)$ for the incidence matrix of W with respect to C, where C denotes (branches of) a train track as in Lemma 9.

We will show using a standard technique that if $W \in GR^*(\mathscr{P})$ then the dilatation of W is given by the spectral radius of $\mathrm{Inc}(\tau^*)$, where $\tau^* = \tau^*(W, \mathscr{P})$. We then proceed to show that it is enough to spectrally analyze the smaller incidence matrix $\mathrm{Inc}(X^*)$. The fact that the size of this matrix depends on W turns out to be inconvenient, so we finally show that the dilatation of W is also given by the spectral radius of $\mathrm{Inc}(X)$. After this subsection all incidence matrices considered in the remainder of this paper are with respect to the branches X.

Lemma 12. *Suppose that $W \in GR(\mathscr{P})$, then*

(a) *The spectral radii of* $\mathrm{Inc}(\tau^*)$, $\mathrm{Inc}(X)$ *and* $\mathrm{Inc}(X^*)$ *agree.*

(b) *The dilatation of W equals the spectral radius λ of* $\mathrm{Inc}(X)$. *Moreover,*

(c) *If $z \in \mathbb{R}^{t(2)}$ is nonnegative and nonzero then* $\lim_{n\to\infty}(1/\lambda^n)\mathrm{Inc}(X)^n z = cx^{(0)}$, *where $c > 0$ and $x^{(0)} = (x_1, \ldots, x_{t(2)})$ is the unique (up to multiples) nonnegative eigenvector of* $\mathrm{Inc}(X)$ *that corresponds to λ. In addition, $x_i > 0$, for $1 \leq i \leq t(1)$.*

Proof. It is a straightforward exercise, using Lemma 9, to show that the spectra of $\mathrm{Inc}(\tau^*)$, $\mathrm{Inc}(X)$ and $\mathrm{Inc}(X^*)$ differ only by eigenvalues of modulus 1. As $\mathrm{Inc}(X^*)$ is P.F. (by Lemma 9(d)) and integral, we see that the spectral radius λ of $\mathrm{Inc}(X^*)$ satisfies $\lambda > 1$, hence λ is also the spectral radius of $\mathrm{Inc}(X)$ and $\mathrm{Inc}(\tau^*)$. This proves part (a).

To prove part (b), we identify $V(\tau^*)$ with a closed cone in $\mathbb{R}^{t(1)+t(3)-t(2)}$, hence $(V(\tau^*) \setminus \{0\})/\mathbb{R}_+$, where $\mathbb{R}_+$ denotes the positive real numbers, can be represented by a closed cell. $\mathrm{Inc}(\tau^*)$ induces a continuous self-map of this cell and we conclude from the Brouwer fixed point theorem that there exist $x \in$

$\mathbb{R}^{t(1)+t(3)-t(2)}$ and $\delta > 0$ such that $\mathrm{Inc}(\tau^*)x = \delta x$. Moreover, x corresponds to a nonzero measure on τ^*.

We show that δ is the spectral radius of $\mathrm{Inc}(X)$. As before, we conclude that x is positive on some branch of X^*, hence Lemma 7(a) implies that x is positive on branches of X^*. We define $x^* \in \mathbb{R}^{t(1)}$ to be the restriction of x onto its first $t(1)$ components, and conclude from Lemma 9(c) that $\mathrm{Inc}(X^*)x^* = \delta x^*$. The uniqueness statement of the Perron-Frobenius theorem implies that δ is equal to the spectral radius of $\mathrm{Inc}(X^*)$ and hence (by part (a)) to the spectral radius λ of $\mathrm{Inc}(X)$.

We next show that $\delta = \lambda$ is the dilatation of W. To that end, we define $\mu \in V(\tau^*)$ to be the measure that corresponds to x and $\widehat{W}$ to be the self-map of $V(\tau^*)$ that is induced by W. If we define $(\mathscr{F}, \nu) = \mathscr{I}(\mu)$, then the commutative diagram (ii) implies that $\overline{W}(\mathscr{F}, \nu) = (\mathscr{F}, \lambda\nu)$. But there is only one such foliation (class) for a pseudo-Anosov map with $\lambda > 1$, and we therefore can conclude that λ is the dilatation of W.

We are left to prove part (c). We claim first that the spectral radius λ of $\mathrm{Inc}(X)$ satisfies the condition of Lemma 1(d), i.e. λ is an eigenvalue of maximum modulus and a simple root of the characteristic equation. Indeed, the Perron-Frobenius theorem implies that the spectral radius $\lambda > 1$ of $\mathrm{Inc}(X^*)$ satisfies the conditions of Lemma 1(d). But the spectra of $\mathrm{Inc}(X^*)$ and $\mathrm{Inc}(X)$ differ only by eigenvalues of modulus 1, hence $\mathrm{Inc}(X)$ satisfies the condition of Lemma 1(d). We conclude that $z_n = (1/\lambda^n)\mathrm{Inc}(X)^n z$ converges to $cx^{(0)}$, for some $c \geq 0$, and we assert that in fact $c > 0$. To that end we define $\pi w \in \mathbb{R}^{t(1)}$, where $w \in \mathbb{R}^{t(2)}$, to be the projection of w onto its first $t(1)$ coordinates. We readily conclude from Lemma 9(b) that

$$\pi z_n = \pi \frac{1}{\lambda^n}(\mathrm{Inc}(X))^n z \geq \frac{1}{\lambda^n}\mathrm{Inc}(X^*)^n \pi z .$$

As the power method works for P.F. matrices (Lemma 1(d)), we see that the right-hand side converges to the positive eigenvector of $\mathrm{Inc}(X^*)$ that corresponds to λ, hence z_n does not converge to 0. Note that we also showed that $\pi x^{(0)} > 0$, which completes the proof of the lemma. □

As already mentioned, for the remainder of the paper we will take all incidence matrices with respect to the branches X of τ, ordered as follows:

$$x(1, 1), \dots, x(1, n_1), \dots, x(k, 1), \dots, x(k, n_k).$$

The smallest dilatation arising from a word in $GR^*(\mathscr{P})$. The key step in the argument is a monotonicity result stated in Proposition 14 and we start by stating a technical lemma.

Lemma 13. *Suppose that for some partition $\mathscr{P} = (m_1, \dots, m_k)$ of $2g$, $W' \in GR(\mathscr{P})$ and that we get $W \in GR(\mathscr{P})$ from W' by deleting a letter S_t, for some $1 \leq t \leq k$. Denote by M and M' the incidence matrix of W and W', respectively (with respect to the branches X of τ). Suppose further that ρ is a weight on τ that is positive on a branch of $X(t)$ and zero otherwise. We then claim that*

$$(M^n\rho)_v \leq (M'^n\rho)_v, \quad \forall n > 0, \ \forall b_v \in X \setminus X(t).$$

Sketch of proof. For L large to be determined in the course of the proof, we associate with each initial subword N' of M'^L an initial subword N of

M^L. Namely, we get N by deleting the "extra" letters S_t of N'. A (rather lengthy) combinatorial argument shows that for any such pair N and N' we have $(N\rho)_v \leq (N'\rho)_v$, where v is such that $b_v \in \{x(i+1, 1), x(i-1, n_{i-1})\}$. Here we took the case $1 < t < k$; the other cases are similar. We can easily deduce from this the claim of the lemma. □

Remark. Note that in the previous lemma W and W' are not required to be elements of $GR^*(\mathscr{P})$.

We derive from the last lemma a partial ordering of the set of dilatations arising from $GR^*(\mathscr{P})$, for a fixed partition $\mathscr{P}$ of $2g$.

Proposition 14. *In the notation of Lemma* 13, *we claim that the spectral radius of M does not exceed the spectral radius of M'.*

Proof. Denote by x the eigenvector of M corresponding to the spectral radius λ of M and by x' the eigenvector of M' corresponding to the spectral radius λ' of M'. Assume to derive a contradiction that $\lambda > \lambda'$. We take ρ as in Lemma 13 and conclude from the same lemma that

$$0 \leq \frac{(M^n\rho)_v}{\lambda^n} \leq \frac{(M'^n\rho)_v}{\lambda^n}, \quad \forall n > 0, \ \forall b_v \in X \setminus X(t).$$

Lemma 12(c) asserts that the left-hand side converges to cx, with $c > 0$, and as the right-hand side converges to 0 by Lemma 1(a) we conclude that x is zero on branches of $X \setminus X(t)$. This however contradicts the fact stated in Lemma 12(c) that x is positive on branches of $X^* \subset X$. We see that our assumption $\lambda > \lambda'$ was absurd. □

This next lemma solves the problem of finding the smallest dilatation arising from $W \in GR^*(\mathscr{P})$, for a fixed partition $\mathscr{P}$ of $2g$. Recall the definition of $W(\mathscr{P})$ as in equation (iv).

Lemma 15. *Suppose that $\mathscr{P}$ is a partition of $2g$ and $W \in GR^*(\mathscr{P})$, then the dilatation of $W(\mathscr{P})$ is less than or equal to the dilatation of W.*

Proof. Suppose that $W \in GR^*(\mathscr{P})$, then as W is complete, we can delete appropriate letters to be left with a permutation PW of $W(\mathscr{P})$. Proposition 14 asserts that the spectral radius of the incidence matrix of PW does not exceed the spectral radius of the incidence matrix of W. Lemma 4(b) assures that PW has normal form $W(\mathscr{P})$, and as $W(\mathscr{P})$ corresponds to a p.A. map (by Corollary 11(b)) so does $W(\mathscr{P})$. We conclude from Lemma 12(b) that the dilatation of PW does not exceed the dilatation of W. As the dilatation of a word is invariant by changing the word to normal form, we deduce that the dilatation of PW equals the dilatation of $W(\mathscr{P})$. This finishes the proof of the lemma. □

The smallest dilatation arising from a word in $\mathscr{GR}^*$. We remark that for the proofs of Lemma 16, Lemma 17, and Corollary 19 below, it is more convenient to use the track $\overline{\tau}(\mathscr{P})$ introduced earlier. To do that it is necessary to prove a result analogous to Lemma 12, as is done in [B1]. The advantage is that we do not need to distinguish between the cases $m_i = 1$ and $m_1 > 1$.

We conclude from the last section that for a given partition $\mathscr{P}$ of $2g$ the smallest dilatation in $GR^*(\mathscr{P})$ is realized by $W(\mathscr{P})$. We denote this smallest dilatation by $\lambda(\mathscr{P})$. To solve the second of the problems posed in §3 we need to find the partition $\mathscr{P}$ of $2g$ for which $\lambda(\mathscr{P})$ is minimal.

Lemma 16. *If* (n, m) *is a partition of* $2g$, *then* $\lambda(n, m) \leq \lambda(\mathscr{P})$, *for every partition* $\mathscr{P}$ *of* $2g$.

Sketch of proof. Clearly it is enough to show that for $\mathscr{P} = (m_1, \dots, m_k)$, with $k \geq 3$ and $\mathscr{Q}^{(i)} = (m_1, \dots, m_i + m_{i+1}, \dots, m_k)$, where $1 \leq i \leq k-1$, we have $\lambda(\mathscr{Q}^{(i)}) \leq \lambda(\mathscr{P})$.

Denote by P and $Q^{(i)}$ the incidence matrix of $W(\mathscr{P})$ and $W(\mathscr{Q}^{(i)})$, respectively. By explicitly computing P and $Q^{(i)}$ we can see that there is a matrix R such that $Q^{(i)}$ is a principal minor of R and $R \leq P$. The claim then follows from Lemma 1(b) and (c). □

We are left to prove

Lemma 17. $\lambda(g, g) \leq \lambda(m_1, m_2)$, *where* (m_1, m_2) *is a permutation of* $2g$.

Sketch of proof. By analyzing the incidence matrix corresponding to $W(m_1, m_2)$, we find that the dilatation $\lambda(m_1, m_2)$ of $W(m_1, m_2)$ is the solution of maximum modulus of

$$\text{(viii)} \qquad x^{2g+2} - x^{2g+1} + 1 - x = 2x^{m_1+1} + 2x^{m_2+1}.$$

An analysis of this equation yields that the solution of maximum modulus is minimal if $m_1 = m_2 = g$, where we minimize over all permutations (m_1, m_2) of $2g$. □

We summarize

Theorem 18. *The smallest dilatation arising from a word in* $\mathscr{GR}^*$ *is realized by* $W(g, g) = S_2 S_1$.

Proof. Lemmas 15, 16, and 17. □

Remark. It is interesting that for every partition $\mathscr{P}$ of $2g$ the smallest dilatation arising from a word in $GR^*(\mathscr{P})$ is realized by a map that arises in knot theory (see [GK]).

Asymptotic behavior of λ_g^*. We denote the dilatation of the map $W(g, g)$ by λ_g^*, for $g \geq 1$.

We can show the following asymptotic behavior of λ_g^*.

Corollary 19. (a) λ_g^* *satisfies* $x^{2g+2} - x^{2g+1} + 1 - x = 4x^{g+1}$.

(b) $\lambda_g^* \leq (2g-1)^{1/g}$, *for* $g > 1$ *and hence* $\lim_{g\to\infty} \lambda_g^* = 1$.

Sketch of proof. The proof of part (a) follows from the proof of Lemma 17.

For the proof of part (b) we choose $g > 1$ and define $\mathscr{P} = (g, g)$ and $W = W(g, g)$. Lemma 8(c) implies that $\tau^*(W, \mathscr{P}) = \tau(\mathscr{P})$. We denote the incidence matrix of W by M and deduce from Lemma 9(d) that $M \in \mathrm{Mat}_l(\mathbb{R})$ is P.F. Note that the fact that M is P.F. implies that M^t and $(M^t)^g$ are also P.F. Moreover, as $\lambda = \lambda_g^*$ is the spectral radius of M, λ^g is the spectral radius of $(M^t)^g$. We conclude from Theorem B(b) that

$$\lambda^g \leq \max_{1\leq i\leq l} \frac{((M^t)^g x)_i}{x_i},$$

for any $x \in \mathbb{R}^l$, $x > 0$. For judicious choice of $\overline{x} \in \mathbb{R}^l$ we compute

$$\frac{((M^t)^g \overline{x})_i}{\overline{x}_i} \leq (2g-1), \quad \text{for } i = 1, \dots, 6g, \text{ and } g > 1.$$

The claim of part (b) follows from the last two inequalities. □

We remark that it was already known that the smallest dilatation l_g^* arising from a p.A. homeomorphisms of a surface $F(g, 0, 0)$ converges to 1. In fact [P2] showed that $l_g^* \leq 11^{1/g}$. Combining ideas from [P2] and the present investigation we can improve the estimate to $l_g^* \leq 6^{1/g}$ (see [B2]).

6. Relations in $S(\mathscr{P})$

In this section we consider again a fixed partition $\mathscr{P} = (m_1, \dots, m_k)$ of $2g$.

For the purpose of this section, we will need to distinguish between a word in the semigroup of formal words $S(\mathscr{P})$ and the mapping classes they represent. In particular, there are no nontrivial relations among the words in $S(\mathscr{P})$. We define

$$S(\mathscr{P})/\sim = \{\text{words in } S_1, \dots, S_k\}/\sim,$$

where we mod out by the relations (v). As these relations also hold for the mapping classes $[\phi_i]$ that correspond to the S_i, for $i = 1, \dots, k$, we can define a semigroup homomorphism

$$\begin{aligned} \Psi\colon S(\mathscr{P})/\sim &\longrightarrow MC(F(0, 1, 2g+1)) \\ S_i &\mapsto [\phi_i]. \end{aligned}$$

The goal of this section is to show that the map Ψ is an injection.

As in §4, we define $t(3)$ tc be the cardinality of the set of branches of $\tau(\mathscr{P})$ and $t(2)$ to be the cardinality of the set X of branches of $\tau(\mathscr{P})$.

For $1 \leq i \leq k$, we denote by $M_i \in \mathrm{Mat}_{t(2)}(\mathbb{Z})$ the incidence matrix of $S_i \in S(\mathscr{P})$ with respect to the branches X of $\tau = \tau(\mathscr{P})$. Note that we have the obvious relations

$$\text{(ix)} \qquad M_i M_j = M_j M_i, \quad \text{if } |i-j| \neq 1,$$

(compare to (v)), so we can define a semigroup homomorphism as follows

$$\begin{aligned} \psi\colon S(\mathscr{P})/\sim &\longrightarrow \mathrm{Mat}_{t(2)}(\mathbb{Z}) \\ S_i &\mapsto M_i. \end{aligned}$$

We have

Lemma 20. (a) *$\tau(\mathscr{P})$ supports a positive measure.*

(b) *For $j = 1, 2$, we define $N_j = \psi(W_j)$, where $W_j \in S(\mathscr{P})$. Of course, we can write N_j, for $j = 1, 2$ as a product of M_i, for $1 \leq i \leq k$. We claim that if $x \in \mathbb{R}^{t(2)}$ is strictly positive and satisfies $N_1 x = N_2 x$, then N_1 is related to N_2 via the relations* (ix).

Proof. Part (a) is easily established and for the somewhat tedious proof of part (b) we refer to [B1]. □

We can now prove

Theorem 21. (a) *ψ is a faithful linear representation.*

(b) *$S(\mathscr{P})/\sim$ injects into the mapping class group $MC(F(0, 1, 2g+1))$.*

Proof. The proof of part (a) follows immediately from Lemma 20, so we are left to prove part (b).

To that end, we assume that $\Psi(W_1) = \Psi(W_2) = [\phi]$, where we took ϕ so that $\tau(\mathscr{P})$ is a ϕ invariant train track. By Lemma 20(a) we can choose a positive $\mu \in V(\tau)$ and the commutative diagram (ii) implies that $\overline{W}_j(\mathscr{I}(\mu)) = \mathscr{I}(\widehat{W}_j\mu)$, for $j = 1, 2$. As W_1 is isotopic to W_2 by assumption, we know that $\overline{W}_1(\mathscr{I}(\mu)) = \overline{W}_2(\mathscr{I}(\mu))$ and hence $\mathscr{I}(\widehat{W}_1\mu) = \mathscr{I}(\widehat{W}_2\mu)$. We use now the injectivity of $\mathscr{I}$ to conclude that $\widehat{W}_1\mu = \widehat{W}_2\mu$. We next represent μ by $x \in \mathbb{R}^{t(3)}$ and denote the restriction of x onto its first $t(2)$ coordinates by x', so x' represents an assignment of numbers to the branches of X. Using Lemma 9(a) we conclude that $N_1x' = N_2x'$, where $N_i = \psi(W_i)$, for $i = 1, 2$. But $x' > 0$, hence Lemma 20(b) implies that N_1 is related to N_2 via the relations (ix). We conclude from part (a) that $W_1 = W_2$ as elements of $S(\mathscr{P})/\sim$, which finishes the proof of the theorem. □

References

[Ab] W. Abikoff, *The real-analytic theory of Teichmüller space*, Lecture Notes in Math., Springer-Verlag, vol. 820, 1980.

[B1] M. Bauer, *Examples of pseudo-Anosov maps*, Thesis, Univ. of Southern California, 1989.

[B2] ——, *An upper bound for the least dilatation*, Trans. Amer. Math. Soc. **330** (1992), 361–370.

[B3] ——, *Dilatations and continued fractions*, Linear Algebra Appl. (to appear).

[Be] L. Bernstein, *The Jacobi-Perron algorithm—Its theory and application*, Lecture Notes in Math., Springer-Verlag, 1971.

[BH] J. S. Birman and H. M. Hilden, *On isotopies of homeomorphism of Riemann surfaces*, Ann. of Math. **97** (1973).

[CB] S. Bleier and A. Casson, *Automorphisms of surfaces after Nielsen and Thurston*, London Math. Soc. Student Texts 9, Cambridge Univ. Press, 1988.

[FLP] A. Fathi, F. Laudenbach, V. Poenaru et al., *Travaux de Thurston sur les surfaces*, Asterisque **66–67**, Sem. Orsay, Soc. Math. de France, 1979.

[Ga] F. Gantmacher, *Theory of matrices* (vol. 2), Chelsea, 1960.

[Gi] J. Gilman, *On the Nielsen type and the classification for the mapping class group*, Adv. in Math. **40** (1981).

[GK] D. Gabai and W. Kazez, *Pseudo-Anosov maps on fibered 2-bridge knots* (to appear).

[HP] R. C. Penner (with J. L. Harer), *Combinatorics of train tracks*. Ann. of Math. Studies (orange series), Princeton Univ. Press., 1991.

[Hu] P. Hummel, *Continued fractions and matrices*, Tôhoku Math. J. **46** (1940).

[Ni] J. Nielsen, *Untersuchungen zur Topologie der Geschloßenen Flächen*. I, II and III, Acta Math. **50** (1927), **53** (1929), and **58** (1932).

[Pa] A. Papadopoulos, *Difféomorphismes pseudo-Anosov et automorphismes symplectiques de l'homologie*, Ann. Sci. Ecole Norm. Sup. **15** (1982), 543–546.

[PP] A. Papadopoulos and R. C. Penner, *A characterization of pseudo-Anosov foliations*, Pacific J. Math. **130** (1987).

[P1] R. C. Penner, *A construction of pseudo-Anosov homeomorphisms*, Trans. Amer. Math. Soc. **130** (1988).

[P2] ——, *Bounds on small dilatations* (to appear).

[Th] W. Thurston, *The geometry and topology of three-manifolds*, Lecture Notes, Princeton Univ., 1978.

[Zi] H. Zieschang, *On the homeotopy group of surfaces*, Math. Ann. **206** (1973).

Department of Mathematics, University of Connecticut, Storrs, Connecticut 06269-3009

Current address: Département de Mathématiques, Université de Rennes I, Campus de Beaulieu, 35042 Rennes Cedex, France