Invariant Laminations for Diffeomorphisms of Handlebodies

by

Martin Lustig[1] **and Ulrich Oertel**[2]

§0. Introduction.

The Nielsen-Thurston theory for the classification of homeomorphisms of surfaces uses the description of pseudo-Anosov homeomorphisms in terms of invariant laminations and train tracks, see [T]. An analogous theory was developed by M. Bestvina and M. Handel in [BH] for analyzing automorphisms of free groups. The purpose of this paper is to understand diffeomorphisms of handlebodies from the point of view of invariant laminations. It bridges, in some sense, the two existing theories.

We suppose that H is a handlebody of genus g and dimension $n \geq 2$. This means that H is a manifold that can be expressed as a union of 0-handles and 1-handles. If $n \geq 3$ and H is orientable and connected, then H is the boundary connected sum of g copies of $D^{n-1} \times S^1$. According our definition any connected surface with boundary is a handlebody.

Theorem 1. *Suppose $f : H \to H$ is a diffeomorphism of an n-dimensional handlebody. Then there is a measured lamination $\Lambda \subset \mathring{H}$ with transverse measure μ such that, after a suitable isotopy of f, $f(\Lambda, \mu) = (\Lambda, \lambda\mu)$ for some $\lambda \geq 1$. The leaves of Λ are open $(n-1)$-dimensional discs. If $n \neq 3, 4$, their lifts to the universal cover $\tilde{H}$ of H are proper subsets of int$(\tilde{H})$, in the sense that intersections with compact sets are compact.*

The lamination Λ of the theorem is a lamination in a non-compact manifold; in general it does not extend to a properly embedded lamination in H, in the sense that there is an embedding of the pair $(\Lambda, \partial\Lambda)$ in $(H, \partial H)$. This may disturb the reader, but we will give an argument at the end of Section 5 to show it is unavoidable.

Letting F_g denote the free group of rank g, we obtain the following application to automorphisms of free groups:

Theorem 2. *Every automorphism $\alpha \in Aut(F_g)$ is realized by a stratified train track map $h : \Gamma \to \Gamma$, where Γ is a graph with $\pi_1(\Gamma) = F_g$.*

The notion of a stratified train track map (compare also [GJLL]) is a somewhat simplified version of the notion of a relative train track map, as defined in [BH]: Suppose Γ is a graph with fundamental group F_g. We call a path $\beta : [0, 1] \to \Gamma$ a *h-backtracking path* if β is locally injective with $h\beta(0) = h\beta(1)$, and if $h\beta$ defines

[1]Supported by a Heisenberg Stipendium for the Deutsche Forschungsgemeinschaft (DFG), by the Institute for Advanced Study, Princeton, and by the Max Planck Institute, Bonn.

[2]Supported by the NSF, the Max Planck Institute, and the Université Louis Pasteur, Strasbourg.

a contractible loop in Γ. A *stratified train track map* is a map $h : \Gamma \to \Gamma$ with a stratification $\Gamma_1 \subset \Gamma_2 \subset \ldots \subset \Gamma_r = \Gamma$ of Γ, where each Γ_i is a (possibly non-connected) h-invariant subgraph, such that, for all $t \geq 1$, any h^t-backtracking path in $\Gamma_i - \Gamma_{i-1}$ is mapped by h^t to Γ_{i-1}.

Every relative train track map is a stratified train track map, see [GJLL]. Thus our Theorem 2 also follows from Theorem 5.12 of [BH]. The converse is not quite true, but it seems likely that one could derive the existence of relative train track maps from our point of view, with some additional work. In any case, for many purposes this seems unnecessary, since stratified train track maps share the main property of relative train track maps. For example, it is now possible to prove the Scott Conjecture using stratified train track maps rather than relative train track maps, see [GJLL].

To prove Theorem 1, one first represents the isotopy class of a diffeomorphism by a diffeomorphism f with "outward expanding" dynamics, i.e. a diffeomorphism which moves points out from a concentric handlebody $H_0 \subset \text{int}(H)$. Then one chooses a suitable collection $\mathcal{E}$ of essential discs in H_0 with the property, after isotopy of f, that $f(\mathcal{E})$ intersects H_0 in parallel copies of discs of $\mathcal{E}$. From the image of $\mathcal{E}$ under all positive iterates of f, one constructs a non-compact branched manifold B which branches at spheres in the boundaries of the different concentric handlebodies $f^t(H_0)$. Considering the incidence of the discs of $f(\mathcal{E})$ on discs of $\mathcal{E}$, one can find an "incidence matrix" with non-negative integer entries and an eigenvector of weights on the branched manifold, much as one finds an eigenvector of weights on an invariant train track for a homeomorphism of a surface. These weights satisfy switch conditions, and hence define a measured lamination Λ carried by B, which is invariant under f. This has all of the properties of the lamination of Theorem 1, except for the key property of properness of the leaves of $\tilde{\Lambda}$ in $\text{int}(\tilde{H})$ (see Proposition 1.4). As we shall see, the lack of properness of leaves of $\tilde{\Lambda}$ manifests itself as a property of the branched manifold B, namely, the existence of "monogons" for B. Sections 3 and 4 are devoted to modifying B and the lamination Λ in order to eliminate monogons. It turns out that the modifications reduce the eigenvalue of the incidence matrix, and this ensures that the modifications must eventually end with a branched manifold having no monogons. The facts we need about non-negative matrices are collected in Section 6.

Roughly speaking, the train track map of Theorem 2 is obtained by applying Theorem 1 and dualizing the system $\mathcal{E}$ of essential discs in H_0 to obtain a graph, with a dual map induced by f. It turns out that the train track property of the induced map on this graph is equivalent to the fact that B does not admit monogons.

There are a number of obvious questions arising from our results. We do not address them in this paper because we do not yet have satisfactory answers, although we do have partial answers. The first and most interesting question: What happens when the handlebody has dimension 3? We know that there are still invariant measured laminations in this case, and it is possible to arrange for some nice properties for the invariant laminations, but *not* the property that the lifts of leaves to the interior of the universal cover of the handlebody are properly embedded. Apparently, the theory in dimension 3 will be quite different. In dimension $n = 2$, one could ask what the relationship is between the invariant laminations of our Theorem 1 and the invariant stable lamination of the Nielsen-Thurston theory. It appears that the laminations in the two theories are essentially (but not exactly) the same. We have not established the precise relationship. For simplicity, we omitted the case of $n = 4$ in this paper; we do not know if there are any real differences between this case and the case $n \geq 5$. Another obvious question: Is the lamination (Λ, μ) of Theorem 1 unique? Under appropriate circumstances, for example in the case that the incidence matrix $M(\mathcal{E})$ is irreducible, it seems reasonable to expect that the lamination (Λ, μ) is independent of the choice of $\mathcal{E}$.

§1. Construction of invariant laminations.

Let H be an n-dimensional handlebody; in this section we do not need the assumption $n \neq 3, 4$. Let $H_0 \subset \mathring{H}$ be a closed handlebody with a product structure on its complement, $H - \mathring{H}_0 = \partial H \times I$, where $I = [0,1]$ and $\partial H \times 0$ is identified with ∂H_0. A diffeomorphism $f : H \to H$ is called *outward expanding* with respect to H_0 if $f|_{\partial H \times I} = f|_{\partial H_0} \times h$, where $h : I \to I$ is an embedding moving every point towards the fixed point 1, so that $h(1) = 1$ and $h(t) > t$ for all $t < 1$. We define $H_t = f^t(H_0)$ for all integers $t \geq 1$.

Since any diffeomorphism $f : H \to H$, after a suitabe isotopy, agrees within a collar $\partial H \times I$ of ∂H with a product diffeomorphism, a further "vertical" isotopy of f within this collar gives:

1.1. Lemma. *Every diffeomorphism of a handlebody is isotopic to an outward expanding diffeomorphism.*

Let $\mathcal{E}$ be a a finite collection of $(n-1)$-discs E_i which are properly embedded in H_0. We will not distinguish notationally between $\mathcal{E}$ and $\cup_{E_i \in \mathcal{E}} E_i$. Each disc E_i is mapped to a properly embedded disc $f(E_i) \subset H_1$.

The collection $\mathcal{E}$ is *admissible* if every $f(E_j)$ intersects H_0 in a collection of subdiscs, each isotopic to some $E_i \in \mathcal{E}$. For each j, $f(E_j) - \mathring{H}_0$ is a punctured disc, which we denote P_j.

1.2. Lemma. *After a suitable isotopy every outward expanding diffeomorphism $f : H \to H$ admits a non-empty admissible collection $\mathcal{E} \subset H_0$ as above, where no E_i is ∂-parallel.*

Proof. Let $\mathcal{E}$ be any non-empty collection of non-∂-parallel $(n-1)$-discs with the property that the union of the $E_i \in \mathcal{E}$ cut H_0 into one or more n-balls. Such a choice of essential $(n-1)$-discs corresponds to a core graph Γ of H_0, with each edge of Γ dual to some E_i. By an isotopy of f within $\mathring{H}$ we arrange that for each j the restriction map $f|_{E_j}$ is transverse to the graph Γ and does not meet the vertices of Γ. Since we can think of H_0 as a small regular neighborhood of Γ, $f(E_j)$ intersects H_0 in a collection of subdiscs, each isotopic to some E_i. Thus $\mathcal{E}$ is admissible. □

We are now ready to construct the invariant lamination. Since it depends on the choice of our $\mathcal{E}$, the lamination is not canonical, and it will be necessary later to perform modifications which replace the original invariant lamination by a better one.

From any admissible collection $\mathcal{E}$, we will first construct an $(n-1)$-dimensional branched manifold B embedded in $\mathring{H}$ with local model, for $n = 2$, as show in Figure 1a, but with arbitrarily many "branches" on the right side. A local model in the case $n > 2$ is obtained by taking a product with a ball of the appropriate dimension.

Figure 1

To define B, we isotope each $f(E_j)$ within a regular neighborhood of $\mathcal{E}$ such that after the isotopy, for every E_i, all component discs of any $f(E_j) \cap H_0$ isotopic to E_i are identified with E_i. We perform similar identifications f-equivariantly, identifying copies of $f(E_i)$ isotopic in H_1 and contained in $f^2(E_j)$, etc., to obtain a non-compact branched manifold B with ends approaching ∂H. Thus B is determined by $B \cap H_1$, and can be constructed from $B \cap H_1$ by attaching iterates $f^t(B \cap (H_1 - \mathring{H}_0))$. The *branch locus* is the set of $(n-1)$-dimensional spheres (on ∂H_t's) where B is not a manifold, each of which is called a *branch sphere*. The *sectors* are completions of components of $B -$ branch locus(B). The sectors can be enumerated, and all of them are either discs E_i, punctured discs $P_i = f(E_i) - H_0$, or punctured discs $f^t(P_i)$, $t \geq 1$, where $E_i \in \mathcal{E}$. The number of branches on the side $H - H_t$ of a branch sphere on ∂H_t is called the *branching order* at that branch sphere.

Let m_{ij} denote the number of parallel copies of E_i in $f(E_j)$, and let $M = M(\mathcal{E})$ denote the matrix (m_{ij}), which will be called the *incidence matrix* for $\mathcal{E}$ with respect to f. It is a standard fact, see Section 6, that a matrix M with non-negative integer entries has an eigenvector x with non-negative entries and that the corresponding eigenvalue $\lambda = \lambda(\mathcal{E})$ satisfies $\lambda \geq 1$, unless some positive power M^k is the zero matrix. Since M^k is the incidence matrix for f^k it follows that $M^k = 0$ implies that every $E_i \in \mathcal{E}$ is ∂-parallel. The eigenvector x now yields an infinite weight function w which assignes a weight to each sector:

$$w(E_i) = x_i \text{ and } w(f^t(P_i)) = x_i/\lambda^{t+1} \text{ for } t \geq 1 \qquad (*)$$

As is well known, a weight function defines a measured lamination $(\Lambda \, , \mu)$ carried by B if the entries of the weight vector satisfy the *switch conditions*. This means that the weight on the sector in $H_t - H_{t-1}$ adjacent to a branch sphere in ∂H_t equals the sum of weights on sectors in $H_{t+1} - H_t$ adjacent to the same branch sphere, with appropriate multiplicity if a sector abuts the branch sphere more than once. We easily verify:

1.3. Lemma. *A weight function w, defined in terms of a non-negative vector x and of $\lambda \geq 1$ as above in (*), satisfies the switch conditions (and hence yields a measured lamination (Λ, μ) carried by B) if and only if x is an eigenvector of the incidence matrix M and λ is the corresponding eigenvalue.*

In order to make Λ f-invariant, we refine the construction slightly. We choose a neighborhood $N(\mathcal{E}) = \mathcal{E} \times I$ in H_0 and isotope f further, such that we have not only admissibility, but also the property that $f(N(\mathcal{E}))$ meets H_0 only in $\mathring{N}(\mathcal{E})$ and that within this intersection f respects the product strucure. Let N be the union of all $f^t(N(E_i))$. Notice that, up to an f-equivariant isotopy, N is a neighborhood of B. Then we define Λ to be $\cap\{f^t(N(\mathcal{E})) \mid t \geq 0\}$. If we isotope f such that the intersections $f^t(N(\mathcal{E})) \cap H_0 \subset \mathring{N}(\mathcal{E})$ are sufficiently small, then the intersection $\Lambda = \cap\{f^t(N(\mathcal{E})) \mid t \geq 0\}$ will have path components with empty interior, so that Λ is a lamination. Since $f(N) \subset N$, the intersection $\Lambda = \cap\{f^t(N(\mathcal{E})) \mid t \geq 0\}$ is f-invariant. The total transverse measure of all leaves of Λ which intersect some $f^t(N(E_i))$ is given by $w(f^{t-1}(P_i)) = x_i/\lambda^t$; therefore, f multiplies the measure by λ.

We observe that the leaves of Λ are open $(n-1)$-dimensional discs. To see this, we use the characterization of $\mathbb{R}^{n-1}$ by the property that every compact set is contained in a neighborhood homeomorphic to an $(n-1)$-disc. Clearly this is true for the leaves of our lamination.

We summarize the results of our construction as follows:

1.4. Proposition. *Suppose $f : H \to H$ is diffeomorphism of an n-dimensional handlebody, $n \geq 2$. Then, after modifying f by a suitable isotopy, there is a measured lamination $\Lambda \subset \mathring{H}$ with transverse measure μ, such that $f(\Lambda, \mu) = (\Lambda, \lambda\mu)$ for some stretch factor $\lambda \geq 1$. The leaves of Λ are open $(n-1)$-dimensional discs.*

More specifically, given a system $\mathcal{E}$ and an eigenvector x of $M(\mathcal{E})$ with eigenvalue $\lambda \geq 1$, there exists a lamination (Λ, μ) as above, carried by $B(\mathcal{E})$, which is uniquely determined, up to isotopy of Λ and up to scalar multiplication of μ.

We have not yet defined what it means for a collection $\mathcal{E}$ to be a system; we give this definition together with definitions of other properties of collections $\mathcal{E}$:

- $\mathcal{E}$ is a *system* if it is non-empty and admissible.
- $\mathcal{E}$ is a *subsystem* of a system if it is a subset of the system and a system.
- $\mathcal{E}$ is *∂-parallel* if each $E_i \in \mathcal{E}$ is ∂-parallel.
- $\mathcal{E}$ is *complete* if the interior of $H_0 - \mathcal{E}$ is a union of open n-discs.
- $\mathcal{E}$ is *efficient* if no two distinct E_i are parallel and no E_i is ∂-parallel.
- $\mathcal{E}$ is *irreducible* if for each pair (i, j), there is a $t \geq 1$ such that $f^t(E_j)$ contains at least one subdisc parallel to E_i. It follows directly from 6.3 (b) that this is equivalent to the fact that the incidence matrix M is irreducible and non-zero.

We conclude with two lemmas concerning systems. We observe that Lemma 1.6, despite its innocent appearance, will be a crucial step in the proof of Theorem 1.

1.5. Lemma. *For every non-∂-parallel system $\mathcal{E}$ there is a system $\mathcal{E}'$, a subsystem of $\mathcal{E}$ up to isotopy of $\mathcal{E}'$, which is irreducible, not ∂-parallel, and satifies $M(\mathcal{E}') \neq 0$. In fact, none of the discs in $\mathcal{E}'$ is ∂-parallel.*

Proof. Consider the subset $\mathcal{E}'' \subset \mathcal{E}$ which consists of those E_i which are not ∂-parallel. By assumption $\mathcal{E}''$ is non-empty, and, after isotopying the E_i suitably, we can eliminate all intersections of $f(\mathcal{E}'')$ with H_0 which are ∂-parallel. Hence $\mathcal{E}''$ is admissible and thus a system. The indices for the elements of $\mathcal{E}''$ can be permuted such the incidence matrix is in block upper diagonal form with irreducible diagonal block matrices, see Section 6. Now let $\mathcal{E}'$ be the subsystem of $\mathcal{E}''$ corresponding to the first block submatrix. Because $\mathcal{E}'$ is not ∂-parallel, it follows that $M(\mathcal{E}')$ is not the zero matrix. $\square$

1.6. Lemma. *Suppose we are given a system $\mathcal{E}$ which is irreducible and not ∂-parallel, but non-efficient. Then there exists an irreducible, efficient system $\mathcal{E}'$ with $\lambda(\mathcal{E}') \leq \lambda(\mathcal{E})$. If $\mathcal{E}$ contains ∂-parallel discs, then $\lambda(\mathcal{E}') < \lambda(\mathcal{E})$.*

Proof. If there are parallel discs E_a and E_b in $\mathcal{E}$, then we can divert all occurrences of an E_a in any $f(E_j)$ to E_b, or vice versa, by isotopy of f and discard E_a from $\mathcal{E}$. After finitely many such diversions we obtain a system $\mathcal{E}''$ which does not contain distinct discs which are parallel. Each diversion affects the incidence matrix as in the hypotheses of Lemma 6.7. This lemma then guarantees that $\lambda(\mathcal{E}'') \leq \lambda(\mathcal{E})$.

Now, using Lemma 1.5, we can replace $\mathcal{E}''$ by a subsystem $\mathcal{E}'$ of itself to obtain an irreducible subsystem without ∂-parallel discs. Since $\mathcal{E}' \subset \mathcal{E}''$ also does not contain parallel discs, it is efficient. Clearly one has $\lambda(\mathcal{E}') \leq \lambda(\mathcal{E}'')$.

Notice that, if some of the original discs in $\mathcal{E}$ are ∂-parallel, then, as a very first step, we can isotope the $f(E_j)$'s to eliminate those. Since entries of the incidence matrix relative to $\mathcal{E}$ are strictly reduced, this strictly reduces the Perron-Frobenius eigenvalue λ, see Lemma 6.5 (b). $\square$

§2. Monogons.

In this section we shall describe a geometric phenomenon, the existence of monogons for a branched manifold, which, as we shall see, is responsible for a non-minimal eigenvalue λ in the statement of Propositon 1.4, and is also responsible for non-properness of lifts of leaves to the universal cover of $\mathring{H}$.

Before explaining monogons, we present some notation related to branched manifolds. Branched manifolds were first defined by R. Williams in [W]; branched 1-manifolds or train tracks were used to study laminations in surfaces by W. Thurston in various settings, e.g. in [T]; branched surfaces were used to study incompressible surfaces in 3-manifolds in [FO], and were then used to study laminations in 3-manifolds in [O] and [GO].

Suppose B is the codimension-1 branched manifold as constructed in Section 1, embedded in a handlebody H of dimension n, then a *fibered neighborhood* $N(B)$ of B in H is a closed regular neighborhood of B foliated by interval fibers, as shown in Figure 1 for dimension $n = 2$, with the frontier of $N(B)$ being the union of the *horizontal boundary*, $\partial_h N(B)$, and the *vertical boundary* $\partial_v N(B)$ as shown. The model of the fibered neighborhood in Figure 1 can be used to obtain a model of the fibered neighborhood of an $(n-1)$-dimensional branched manifold in a handlebody of dimension n by taking the product with an $(n-2)$-ball, but we must again allow arbitrarily many branches on the right side of the model. The vertical boundary $\partial_v N(B)$ is a trivial I-bundle over an $(n-2)$-manifold, a collection of spheres as shown, where the $(n-2)$-manifold is called the *maw manifold* whose components are *maw spheres*. The maw spheres correspond to the spaces between branches in the right side of Figure 1a. There is a *projection map* $\pi : \mathring{H} \to \mathring{H}/\sim$ which maps $\mathring{H}$ to a quotient space in which fibers of $N(B)$ are contracted to points of $\mathring{H}/\sim$. Notice that $\mathring{H}/\sim$ can be identified with $\mathring{H}$, and $N(B)$ is collapsed to a branched manifold which can be identified with B. The maw spheres appear as cusps in $\mathring{H}_B$, the manifold obtained by cutting $\mathring{H}$ on B. More precisely, $\mathring{H}_B$ and $\mathring{H} - \mathring{N}(B)$ have homeomorphic interiors, and $\partial \mathring{H}_B$ can be obtained as a quotient of $\partial(\mathring{H} - \mathring{N}(B))$ as follows: Every component of $\partial_v(N(B)) \subset \partial(\mathring{H} - \mathring{N}(B))$, a product of type $S^{n-2} \times I$, is collapsed along I-fibers to give a cusp, homeomorphic to S^{n-2}, in $\partial \mathring{H}_B$.

A lamination is *carried* by B if it can be embedded in $N(B)$ transverse to interval fibers; it is *fully carried* if every interval fiber of $N(B)$ is intersected by the lamination.

A *monogon* for B is an embedding $\kappa : (K, \beta, \alpha) \to (\mathring{H} - \mathring{N}(B), \partial_h N(B), \partial_v N(B))$ of a disc K, see Figure 2a, where α and β are complementary closed arcs in ∂K. The arc $\alpha \subset \partial_v N(B)$ is called the *tip* of the monogon. A triple (K, β, α) as above is called a *half-disc*.

Figure 2

Applying the projection map π, we get another point of view: if K is a monogon for B, then $\pi(K)$, which we also denote by K, is a disc embedded in $\mathring{H}$ with $K \cap B = \partial K$, and we also call this disc K a *monogon*, see Figure 2b. We can also regard K as a disc properly embedded in $\mathring{H}_B$ with the property that ∂K intersects the cusp manifold $\pi(\partial_v N(B)) \subset \partial \mathring{H}_B$ transversely at a single point $\pi(\alpha)$, called the *tip*.

The proof of Theorem 1 will be done by performing a sequence of modifications of the initial branched manifold and lamination with the goal of eliminating monogons.

In the proof of the following lemma, we shall use a property of the fibered neighborhood of a branched manifold which is frequently useful. Suppose B is a branched manifold properly embedded in a manifold M, suppose that $N(B)$ is its fibered neighborhood, and suppose S is a compact, possibly non-connected, manifold fully carried by B. We assume S intersects each fiber at least twice; if this is not the case, then we can replace S by $\text{fr}N(S)$, which is a double cover of S embedded in M. Then by isotoping S we can ensure that $\partial_h N(B) \subset S$. This is useful because cutting M on S now yields another manifold. In the cut-open manifold $\hat{M}$, the fibered neighborhood appears as an I-bundle called the *interstitial* I*-bundle*, and $\partial_v N(B)$ becomes the "vertical" boundary of the I-bundle.

Complementary components of $N(B)$ appear in $\hat{M}$ as complementary components of the I-bundle. Notice that in our context the interstitial bundle is simply a collection of products of the form (disc $\times\, I$).

2.1. Lemma. *Suppose that $B = B(\mathcal{E})$ is obtained from a system $\mathcal{E}$, and suppose that the dimension n of H satisfies $n \neq 3, 4$. Let S be any disjoint union of discs fully carried by $B \cap H_t$, e.g. $S = f^t(\mathcal{E})$. Finally suppose that there is a map κ from a half-disc K to H_t, $\kappa : (K, \beta, \alpha) \to (H_t, S, I)$, where I is an interval fiber of $N(B)$ and where the ends of α are mapped to distinct points of I. Then there is a monogon for B in H_t.*

Proof. Replacing S by $\mathrm{fr}N(S)$, we may assume that $\partial_h(N(B)) \subset S$, and that when one cuts open H_t on S one obtains a manifold $\hat{H}_t$ containing a collection of interstitial products of the form (disc $\times\, I$) corresponding to the cut-open fibered neighborhood.

We begin by simplifying the intersection pattern $\kappa^{-1}(S)$. We first homotope $\kappa|_\alpha$ relative endpoints such that the arc α is mapped to an embedded arc. We then homotope κ such that it becomes transverse to S; the map can be made transverse at ∂K since it intersects S only along an arc. There are two cases to consider, depending on whether an orientation on S induces transverse orientations which agree (Figure 3a) or disagree (Figure 3b) on I. In the first case, the pull-back pattern $\kappa^{-1}(S)$ must spiral, contradicting the compactness of S, so we conclude that the pull-back pattern must be as in Figure 3b with opposite induced orientations at the ends of α. Next we remove closed curves of $\kappa^{-1}(S)$ by a further homotopy. This is easily done, since S is simply connected and $\pi_2(H_t)$ is trivial. Finally, we notice that there may be further arcs of $\kappa^{-1}(S)$, each having both endpoints on α. Replacing the disc K by a half-disc cut from K by an innermost arc, and restricting κ to this smaller half-disc, we may assume that the image of α intersects S only at its endpoints.

Figure 3

Next we simplify intersections of K with $\partial_v N(B)$. Notice that now κ can be regarded as a map to $\hat{H}_t$, the manifold obtained by cutting H_t open on S. In $\hat{H}_t$ we see the interstitial I-bundle; this is the completion of the $\mathring{I}$-bundle $N(B) - S$, and $\partial_v N(B)$ corresponds to the vertical boundary of this I-bundle. Now we remove intersections $\kappa^{-1}(\partial_v N(B))$. Each closed curve of intersection $\gamma \subset K$ is mapped to a component of $\partial_v N(B)$, so we can eliminate the curve of intersection as follows: The curve γ bounds a disc R in K which is mapped to H. Now $\kappa|_\gamma$ can be extended to a map of R to $\partial_v N(B)$, since $\partial_v N(B)$ has trivial π_1 when $n \neq 3,4$. We replace $\kappa|_R$ by this map of R to $\partial_v N(B)$, pushed off of $\partial_v N(B)$ to the appropriate side. This removes the closed curve of intersection. We remove all closed curves of intersection in this way.

Each arc γ, bounding a half-disc R in K disjoint from the tip of K, is mapped to a path in $\partial_v N(B)$ representing an element of relative π_1 of $\partial_v N(B)$. Now $\partial_v N(B)$ is homeomorphic to a disjoint union of products $S^{n-2} \times I$, and we know that $\pi_1(S^{n-2} \times I, S^{n-2} \times \partial I)$ is generated by an interval fiber (for any choice of base point). If $\kappa|_\gamma$ is trivial in relative π_1, we can therefore replace $\kappa|_R$ by a map of R to $\partial_v N(B)$ pushed off of $\partial_v N(B)$ as above. This eliminates an arc of intersection, reducing the total number of arcs in K.

After removing all arcs trivial in the relative π_1 of $\partial_v N(B)$, we are left with κ such that $\kappa^{-1}(\partial_v N(B))$ consists of arcs in K mapped to vertical fibers in $\partial_v N(B)$, up to homotopy. Choosing an innermost arc not containing the arc α and bounding a half-disc K', the restriction to the monogon K' yields a map to $H - \mathring{N}(B)$, which can be made an embedding because $n \neq 3,4$. The new monogon K' cannot be mapped to the interstitial I-bundle since this is a collection of products of the form (disc $\times I$). $\square$

Next we prove that the lack of monogons is what is needed to prove Theorem 1.

2.2. Proposition. *Let $n \neq 3,4$, let $\mathcal{E}$ be a system, and suppose that (Λ, μ) is a measured lamination carried by $B = B(\mathcal{E})$. If B has no monogons, then the leaves of the lift $\tilde{\Lambda}$ of Λ to the universal cover int$(\tilde{H})$ of int(H) are properly embedded.*

Proof. We suppose that Λ is embedded in $N(B)$ transverse to interval fibers. If a leaf $\tilde{\ell}$ of $\tilde{\Lambda}$ is not properly embedded, then for some sector $\tilde{Z}$ of $\tilde{B}$ it passes through $\pi^{-1}(\text{int}(\tilde{Z}))$ of $\tilde{B}$ infinitely often. Consider two distinct points of $\tilde{\ell}$ which project under π. to the same point in $\tilde{B}$. They are joined by an arc $\tilde{b}$ in $\tilde{\ell}$. Let $\tilde{a}$ be a short arc in a fiber of $N(B)$ joining the ends of $\tilde{b}$. Then $\tilde{a} \cup \tilde{b}$ is a closed curve, which is null-homotopic in $\tilde{H}$. Projecting to H, we have a closed curve $a \cup b$, null-homotopic in H, with a embedded in a fiber of $N(B)$ and b contained in a leaf ℓ. Furthermore, the arc b is contained in a component of $\ell \cap H_t$ for some t. But such a component of $\ell \cap H_t$ must be isotopic to $f^t(E_i)$ for some $E_i \in \mathcal{E}$. Thus we have a map $\kappa : (K, \beta, \alpha) \to (H_t, S, I)$ to H_t where $S = f^t(\mathcal{E})$ is exactly as in the hypotheses of Lemma 2.1. We conclude that B has an embedded monogon. $\square$

§3. Modifications.

From the results of Sections 1 and 2, we know that for any system $\mathcal{E}$, we can construct a branched manifold $B = B(\mathcal{E})$ carrying an invariant lamination Λ with incidence matrix $M = M(\mathcal{E})$ having Perron-Frobenius eigenvalue $\lambda = \lambda(\mathcal{E})$. To prove Theorem 1 we will need to show that the lamination Λ constructed in Section 1, starting from a diffeomorphism f of a handlebody H, can be replaced by one having the property that the lifts of leaves to the univeral cover of $\mathring{H}$ are properly embedded. The goal is to obtain $\mathcal{E}$ such that $B(\mathcal{E})$ has no monogons. The proof is done by identifying desirable properties of $\mathcal{E}$ and showing that these are preserved by a sequence of modifications designed to eliminate monogons. We will show that each modification reduces or leaves unchanged the eigenvalue $\lambda(\mathcal{E})$. The eigenvalue will serve as a complexity which can be reduced by performing a sequence of modifications whenever $B(\mathcal{E})$ has a monogon.

3.1. Lemma. *If a system $\mathcal{E}$ is irreducible, then it determines a lamination $\Lambda(\mathcal{E}) = \Lambda$ carried by $B(\mathcal{E})$, with a transverse measure μ, where the lamination is unique up to isotopy and the measure is unique up multiplication by a scalar. The lamination is f-invariant, $f(\Lambda, \mu) = (\Lambda, \lambda\mu)$, with stretch factor $\lambda \geq 1$ equal to the Perron-Frobenius eigenvalue $\lambda(\mathcal{E})$. The measured lamination (Λ, μ) is fully carried by $B(\mathcal{E})$.*

In this section, when we say (Λ, μ) is unique we will always mean that Λ is unique up to isotopy and that μ is unique up to multiplication by a scalar.

Proof of Lemma 3.1. Since $\mathcal{E}$ is irreducible, the incidence matrix $M(\mathcal{E})$ is irreducible and non-zero. Hence there is positive eigenvector x for $M(\mathcal{E})$, and, up to scalar multiples, x is the unique non-negative eigenvector of $M(\mathcal{E})$, see Section 6. This determines the weight function w on B as in Lemma 1.3 and hence determines, up to isotopy, the lamination $\Lambda = \Lambda(\mathcal{E})$ carried by B, and determines, up to scalar multiplication, a transverse measure μ for Λ. The equation $f(\Lambda, \mu) = (\Lambda, \lambda(\mathcal{E})\mu)$ is proved in Proposition 1.4. The measured lamination (Λ, μ) is fully carried by B since x and hence w are positive and not just non-negative. $\square$

Some of our modifications will leave not only the eigenvalue λ unchanged, but also the lamination (Λ, μ). Others will change both λ and Λ. The first modification of our system $\mathcal{E}$, which we shall use throughout this section, is called splitting.

3.2. Definition. We choose an E_v in the system $\mathcal{E}$ for a diffeomorphism f, and we suppose that the branch sphere of $B = B(\mathcal{E})$ at ∂E_v has branching order $r \geq 2$. At the branch sphere, there are r punctured discs attached in a specified order, say $P_{j_1} \dots P_{j_r}$, each P_{j_m} being attached on an inner boundary σ_{j_m}. We do not rule out repetitions among the P_{j_m}'s, but there are no repetitions among the σ_{j_m}'s. We form a new system $\mathcal{E}'$ by replacing E_v by two parallel discs E_{v_1} and E_{v_2}. We choose an integer $s, 1 \leq s \leq r-1$, and we attach $P_{j_1}, \dots, P_{j_s}$ to E_{v_1} while attaching $P_{j_{s+1}}, \dots, P_{j_r}$ to E_{v_2}, see Figure 4. To preserve equivariance, we split P_v and $f^t(P_v), t \geq 1$ in the same way to get the new branched manifold $B' = B(\mathcal{E}')$. Clearly, the choice of splitting is actually determined by the choice of a maw sphere at ∂E_v, namely the maw between P_{j_s} and $P_{j_{s+1}}$. We say that the new system $\mathcal{E}'$ or the new branched manifold B' is obtained from the original system or branched manifold by a *simple splitting* (at the chosen maw). There is a projection map $p : \mathcal{E}' \to \mathcal{E}$ or $p : B' \to B$, which maps both E_{v_1} and E_{v_2} to E_v and which commutes with f.

A composition of a finite sequence of simple splittings is called a *splitting*.

It may happen that we wish to split E_i, but that the branching order at ∂E_i is 1. This means that a simple splitting on E_i is impossible. It is then necessary to perform a sequence of simple splittings as follows. The outer boundary of E_i is

Figure 4

attached to exactly one inner boundary of exactly one P_{j_1}. If the branching order at the outer boundary of P_{j_1} or E_{j_1} is not 1 we can perform a simple splitting on E_{j_1}, followed by a simple splitting on E_i. Otherwise, we consider the $f(P_{j_2})$ with one inner boundary attached to the outer boundary of P_{j_1}, and perform a simple splitting there if possible, followed by a splitting on E_{j_1} and E_i, etc. Thus we can often split an arbitrary sector of B using a non-simple splitting.

3.3. Lemma. *Suppose a system $\mathcal{E}$ is irreducible and not ∂-parallel. Then splitting yields a system $\mathcal{E}'$ such that*

(1) *$\mathcal{E}'$ is irreducible and not ∂-parallel,*
(2) *$\Lambda(\mathcal{E}')$ is isotopic to $\Lambda(\mathcal{E})$, and*
(3) *$\lambda(\mathcal{E}') = \lambda(\mathcal{E})$.*

Proof. For a simple splitting, $\mathcal{E}'$ differs from $\mathcal{E} = \{E_1, \ldots, E_k\}$ only in that one E_i, say E_1 is replaced by two parallel copies E_0' and E_1'.

The fact that $\mathcal{E}'$ is non-∂-parallel is immediate.

Let P_{i_0}' and P_{i_1}' be punctured discs with inner boundary on E_0' and E_1' respectively. Then $P_{i_0} = p(P_{i_0}')$ and $P_{i_1} = p(P_{i_1}')$ are punctured discs attached to E_1. To prove irreducibility, we must show that given any r, s, $f^t(E_r')$ contains a copy of E_s' for some $t \geq 1$. We know there exists k with $f^k(p(E_r'))$ containing $p(E_s')$. This proves the lemma if $s \neq 0, 1$. If $s = 0$, then we know that for some k, $f^k(E_r)$ contains an E_{i_0}, hence $f^{k+1}(E_r')$ contains an E_s'. Similarly if $s = 1$.

To prove statements (2) and (3), we examine eigenvectors before and after the splitting. An eigenvector x' with eigenvalue λ' for the incidence matrix $M(\mathcal{E}')$ defines a weight function w' on $B' = B(\mathcal{E}')$ which satisfies the switch conditions, by Lemma 1.3. Hence it determines an f-invariant measured lamination with stretch factor λ', and this lamination agrees with (Λ', μ'), by Lemma 3.1. Furthermore one has $\lambda' = \lambda(\mathcal{E}')$.

But via the projection map $p : B' \to B$ the function w' induces a weight function w on $B = B(\mathcal{E})$ which also satisfies the switch conditions. Hence, by Lemma 1.3, it comes from an eigenvector x of $M(\mathcal{E})$, and the corresponding eigenvalue is equal to λ'. By Lemma 3.1 the lamination determined by w is equal to (Λ, μ), and $\lambda' = \lambda(\mathcal{E})$. The isotopy moving (Λ', μ') to (Λ, μ) corresponding to the projection of B' to B. $\square$

3.4. Lemma. *Given a system $\mathcal{E}$, irreducible and not ∂-parallel, with incidence matrix $M(\mathcal{E})$ not a permutation matrix, then there is a splitting which yields $\mathcal{E}'$ with $B' = B(\mathcal{E}')$ having at least two distinct maw spheres at ∂H_0.*

Proof. Since the incidence matrix is not a permutation matrix, there is some E_j with branching order ≥ 2. By a simple counting argument, there is at least one P_j with at least two inner boundary components. By irreducibility, repeated simple splitting eventually splits P_j. After P_j is split, there are at least two maw spheres. (Of course, one could obtain arbitrarily many maw spheres by further splitting.) $\square$

3.5. Lemma. *Suppose the system $\mathcal{E}$ is irreducible and not ∂-parallel, with associated incidence matrix $M(\mathcal{E})$ not a permutation matrix. Suppose that B has at least two distinct maw spheres on ∂H_0. Given any sector P of B and a maw sphere $\gamma \subset \partial H_0$, then there exists an equivariant splitting of B which splits P, but does not split at the maw sphere γ.*

Proof. If P is not already a disc E_0 of $\mathcal{E}$, the sector P corresponds to a disc $E_0 \in \mathcal{E}$, in the sense that, for some q, $P = f^q(E_0) - \mathring{H}_{q-1}$. Let $\delta \in \partial H_0$ be a maw sphere other than γ and suppose δ is the outer boundary of E_z. By irreducibility, we know that $f^t(E_z)$ contains a copy of E_0 for some $t \geq 1$. Considering the pattern of intersection with $\cup \partial H_i$ with $f^t(E_z)$, we can find a sequence of sectors $E_0 = Z_0, Z_1, Z_2, \ldots, Z_p = f^{t-1}(P_z)$ where each $Z_i \neq Z_0$ is $f^c(P_i)$ for some i, c, and where the outer boundary of Z_j is an inner boundary of Z_{j+1}. We now split B by first doing a simple splitting of Z_p at the maw $f^t(\delta)$, then doing a simple splitting of Z_{p-1} at the newly formed maw on the outer boundary of Z_{p-1}, etc., down to Z_0. Since δ is distinct from γ, and since all splittings except the first are done on newly formed maws, there is never any splitting at γ. $\square$

As we have mentioned, most of the modifications of an irreducible and non-∂-parallel system $\mathcal{E}$ representing a lamination $\Lambda = \Lambda(\mathcal{E})$ yield systems $\mathcal{E}'$ representing an isotopic lamination Λ'. The following modification is more drastic.

Figure 5

3.6. Definition. Suppose we are given a system $\mathcal{E}$ such that there is an embedded monogon K for B contained in $H_1 - \mathring{H}_0$ with tip at ∂H_0. If the tip a of the monogon joins two copies of $E_i \in \mathcal{E}$, the monogon yields an equivariant isotopy of $\mathcal{E}$ "cancelling" the two copies of E_i in some $f(E_j)$, see Figure 5.

More explicitly, we have two copies of E_i contained in $f(E_j)$, and we assume that these are adjacent. Thus the two copies of E_i may be identified with $E_i \times \partial I$ in a regular neighborhood $E_i \times I$ of E_i. The tip of the monogon is $a = x \times I$, where $x \in \partial E_i$ and the remainder of the boundary of the monogon is an arc b. We can isotope the arc b in $f(E_j)$ to a extending the isotopy to $f(E_j)$, so that a band in $f(E_j)$ now lies in $\partial E_i \times I$ as $\delta \times I$, where δ is an arc in ∂E_i. Let $\gamma = \partial E_i - \delta$. We remove $E_i \times \partial I$ and we attach $\gamma \times I$. Finally, we push the portion of $f(E_j)$ in $\partial E_i \times I$ off of H_0. Applying f^{-1} to the isotoped $f(\mathcal{E})$, we obtain a new isotopic system $\mathcal{E}''$ in H_0. The modification replacing $\mathcal{E}$ by $\mathcal{E}''$ is called a *cancellation.* Notice that though $\mathcal{E}''$ is isotopic to $\mathcal{E}$, the incidence matrix will have changed and in general the new lamination will not be isotopic to the old one.

Figure 6

3.7. Lemma. *Suppose a system $\mathcal{E}$ is irreducible and not ∂-parallel. Then cancellation yields a system $\mathcal{E}''$ with a subsystem called $\mathcal{E}'$ such that*

(1) *$\mathcal{E}'$ is irreducible and not ∂-parallel, and*
(2) *$\lambda(\mathcal{E}') < \lambda(\mathcal{E})$.*

Proof. After performing a cancellation we get a system $\mathcal{E}''$ which is isotopic to $\mathcal{E}$. In the cancellation we do not eliminate any disc E_i, even if after the cancellation the branching order at ∂E_i is zero. In particular, as $\mathcal{E}$ is not ∂-parallel, neither is $\mathcal{E}''$. The incidence matrix has one entry, m_{ij}, reduced by 2. It follows from Lemma 6.5 that the new incidence matrix has strictly smaller Perron-Frobenius eigenvalue. We can then pass to an irreducible and non-∂-parallel subsystem by Lemma 1.5. $\square$

Suppose K is a monogon for B. In high dimensions $n \geq 5$ we may assume that $K \cap H_0$ consists of the tip and we may also assume that K is transverse to $\cup \partial H_t$. We may also remove trivial curves of the intersection $K \cap (\cup \partial H_t)$ by isotopy of K. The *complexity* of a monogon is the number of arcs of intersection of $K \cap (\cup \partial H_i)$, excluding the tip. Our next modification is designed to reduce the complexity of a monogon. The half-disc in the definition will be a half-disc cut from K by an innermost arc of intersection.

3.8. Definition. Suppose V is an embedded half-disc in $H_1 - \mathring{H}_0$, with u and v complementary closed arcs in ∂V having endpoints in common, and with $v \subset$

$B, u \subset \partial H_0$, see Figure 6. Suppose that the points of $\partial u = \partial v$ lie on ∂E_a and ∂E_b, and suppose B has branching order 1 at ∂E_a and at ∂E_b. We define a *down-move on the half-disc* V changing $\mathcal{E}$ to $\mathcal{E}'$. We isotope the arc v to u and a little beyond through V, performing the isotopy equivariantly and extending the isotopy to $\mathcal{E}$. This has the effect of performing an ambient boundary connect sum (called a band sum when $n = 3$) joining E_a and E_b to yield $E_a \# E_b$, while also joining P_a to P_b, and $f^t(P_a)$ to $f^t(P_b)$. We discard E_a and E_b from $\mathcal{E}$ and add $E_c = E_a \# E_b$ to obtain $\mathcal{E}'$.

3.9. Lemma. *Suppose a system $\mathcal{E}$ is irreducible and not ∂-parallel. Then a down-move on a half-disc V yields a system $\mathcal{E}'$ such that*

1. *$\mathcal{E}'$ is irreducible and not ∂-parallel,*
2. *$\Lambda(\mathcal{E}')$ is isotopic to $\Lambda(\mathcal{E})$, and*
3. *$\lambda(\mathcal{E}') = \lambda(\mathcal{E})$.*

Proof. We begin with the proof of (1). To prove the irreducibility of the new system $\mathcal{E}'$, we must show that, given any indices r, s for discs in $\mathcal{E}'$, $f^t(E'_r)$ contains a copy of E'_s for some $t > 1$. If neither E'_r nor E'_s is E'_c, then this follows from the irreducibility of $\mathcal{E}$. If $E'_r = E'_c$, then since we know that $f^t(E_a)$ contains E'_s, it follows that $f^t(E'_c)$ contains $E'_s = E_s$. If $E'_s = E'_c$, then we consider the punctured disc P_j having E_a and E_b attached to its inner boundaries, and we suppose initially that $j \neq a, j \neq b$. We know that for some t, $f^t(E_r)$ contains E_j, hence $f^{t+1}(E'_r)$ contains E'_c. If $E_j = E_a$, say, then since $f^t(E_r)$ contains an $E_j = E_a$ it also contains an E_b. This is because ∂E_a has branching order 1 and $f(E_j)$ is the only disc of $f(\mathcal{E})$ which contains E_a. Hence $f^t(E_r)$ contains $f(E_j)$ and thus E_b. Hence $f^t(E'_r)$ contains E'_c.

Next, we prove that $\mathcal{E}'$ is not ∂-parallel. Notice that, if $\mathcal{E}'$ is ∂-parallel, then in particular $f(E'_c) = f(E_a) \# f(E_b)$ is ∂-parallel, so that $f(E_a)$ and $f(E_b)$ are isotopic: Hence either both E_a and E_b are ∂-parallel, or neither is. The system $\mathcal{E}'$ contains discs isotopic to all discs of $\mathcal{E}$ except E_a and E_b. Thus we need only worry about the possibility that every disc of $\mathcal{E}$ except E_a and E_b is ∂-parallel. Since $\mathcal{E}$ is not ∂-parallel, we know that E_a and E_b are not both ∂-parallel, hence neither is ∂-parallel. Because ∂E_a and ∂E_b have branching order one, all discs of $f(\mathcal{E})$ except possible $f(E_j)$ meet H_0 only in ∂-parallel discs. Thus there is at most one disc, $f(E_j)$, of $f(\mathcal{E})$ which is not ∂-parallel: All other discs of $f(\mathcal{E})$ meet H_0 only in ∂-parallel discs. Hence we have shown that if there are two discs in $\mathcal{E}$ (or $f(\mathcal{E})$) which are not ∂-parallel, then there is only one, a contradiction.

The proof of statements (2) and (3) is identical to that of (2) and (3) in Lemma 3.3. We replace the projection p in this proof by the f-equivariant isotopy from $B(\mathcal{E})$ to $B(\mathcal{E}')$ which is given by the isotopy on $\mathcal{E}$ described in Definition 3.8. $\square$

§4. Proof of Theorem 1.

Using the moves described in Section 3, we can now prove Theorem 1 as follows: Given a diffeomorphism f of a handlebody H, Proposition 4.1, below, gives a system $\mathcal{E}$ with $\lambda(\mathcal{E}) \geq 1$ such that the branched manifold $B(\mathcal{E})$ has no monogons. Thus, by Proposition 2.2, any measured lamination Λ carried by $B(\mathcal{E})$, lifts to the universal cover $\tilde{H}$ of H to a lamination $\tilde{\Lambda}$ with leaves that are proper subsets of $\text{int}(\tilde{H})$. By Proposition 1.4 such a lamination Λ with transverse measure μ exists and has all other properties claimed in Theorem 1.

4.1. Proposition. *Suppose $f : H \to H$ is a diffeomorphism of an n-dimensional handlebody, $n \neq 3, 4$. Then there is a branched manifold B, obtained from an irreducible, efficient system $\mathcal{E}$, such that $B = B(\mathcal{E})$ has no monogons.*

Proof. We will first give the proof for $n \geq 5$. Most of this proof applies also to the case $n = 2$, but we will make two statements labelled "Statement 1" and "Statement 2," which are not true in the case $n = 2$. Following the proof in the case $n \geq 5$ we shall indicate how the proof needs to be changed in the case $n = 2$. "Change 1" ("Change 2") will indicate an argument to replace Statement 1 (Statement 2).

We begin with an irreducible and non-∂-parallel system $\mathcal{E}$, obtained from Lemma 1.2 and Lemma 1.5, and we first consider the case that the associated Perron-Frobenius eigenvalue satisfies $\lambda > 1$. By Lemma 6.3 this is equivalent to the property that $M(\mathcal{E})$ is not a permutation matrix. Hence we can apply Lemma 3.4 to ensure that there are at least two distinct maw spheres for B at ∂H_0. If there is no monogon for B , then $\mathcal{E}$ satisfies the conclusion stated in the theorem. Assume now there is a monogon K for B.

Statement 1: If $n \geq 5$, we can assume that K intersects H_0 only at its tip, that it is transverse to $\cup \partial H_i$ without closed curves of intersection, and that all iterates $f^t(K)$ are disjointly embedded.

Let V be a half-disc, not containing the tip, cut from K by an innermost arc of intersection, and suppose $\partial V = u \cup v$ where $v \subset \partial K \subset B$ and the complementary closed arc u lies in ∂H_j for some j.

There are two cases to consider. If $V \subset H_i$, then in the first case $u \subset \partial H_i$ and in the second case $u \subset \partial H_{i-1}$.

Statement 2: In the first case, we can easily isotope $(K, \partial K)$ to eliminate an arc of intersection as follows. The arc v lies in some sector $P = f^t(P_j)$, with both ends of v on the outer boundary of P. If we assume $n \geq 4$, since P is an $(n-1)$-dimensional punctured ball, it is clear that one can isotope v within B out of H_i and a little beyond. We extend the isotopy to all of K, then we eliminate closed curves of intersection of K with $\cup \partial H_i$ which may have appeared. Thus we have reduced the complexity of K, i.e. the number of arcs of intersection $B \cap (\cup_{t \geq 0} H_t)$. This isotopy involves only K, not the system $\mathcal{E}$, and we do not need to worry about equivariance.

In the second case, the half-disc V satisfies $u \subset \partial H_{i-1}$. Applying an f^t for appropriate negative t, we may assume that V and P lie in H_1 with $u \subset \partial H_0$. The two points of $\partial v = \partial u$ lie in ∂E_a and ∂E_b, say, where possibly $E_a = E_b$. If the two ends of v lie on the same inner boundary of P, an isotopy of v extended to K

eliminates an arc of intersection, as in the first case. Otherwise, if necessary, we do a simple splitting at the maw(s) adjacent to the arc v as shown in Figure 7, to ensure that E_a and E_b are distinct, and that the branching order at ∂E_a and at ∂E_b is one. This splitting can be done, since, if the given E_a and E_b are equal, then the branching order at $\partial E_a = \partial E_b$ is necessarily bigger than 1, as the endpoints of v lie on distinct inner ∂-components of P. This gives a new system $\mathcal{E}'$, but Lemma 3.3 shows that $M(\mathcal{E}')$ is irreducible, with the same eigenvalue λ, and $B(\mathcal{E}')$ carries the same lamination up to isotopy.

Figure 7

Figure 8

There is, however one danger when doing the splitting: It is possible that the splitting occurs at a maw sphere which includes the tip of K. This means that one obtains a new monogon whose complexity is larger, since splitting at the maw introduces a new arc of intersection. In this special case, before doing the above splitting, we first use the fact that B has two distinct maws at ∂H_0 and apply Lemma 3.5 to split the sector P, see Figure 8. Then we can split at the maws of the new B which lie at the ends of the arc v (as described above). Neither splitting affects the maw containing the tip of the monogon, by Lemma 3.5. Hence the monogon K remains essentially unchanged after the splitting.

Now we are ready to do a down-move on V. By Lemma 3.9, this yields a new irreducible and non-∂-parallel system with the same Perron-Frobenius eigenvalue, which gives rise to an isotopic lamination. While doing the down move, we isotope K to eliminate the arc of intersection u, thus we obtain a monogon of smaller complexity.

Repeating these operations on the simpler monogon, we must ultimately obtain a monogon with complexity 0, i.e a monogon contained in $H_1 - H_0$. Now cancellation gives a new irreducible and non-∂-parallel system $\mathcal{E}'$ with strictly smaller eigenvalue and a new measured lamination, by Lemma 3.7. Now Lemma 1.6 applies, and we obtain a system $\mathcal{E}_1$ and a lamination (Λ_1, μ_1) with $\lambda(\mathcal{E}_1) \leq \lambda(\mathcal{E}') < \lambda(\mathcal{E})$ and with $\mathcal{E}_1$ irreducible and efficient.

Now, if the new system $\mathcal{E}_1$ is such that there is a monogon, we repeat everything we have done in the proof up to now to obtain another irreducible and efficient system $\mathcal{E}_2$ with $\lambda(\mathcal{E}_2) < \lambda(\mathcal{E}_1)$. We cannot reduce $\lambda(\mathcal{E}_i)$ infinitely often, always obtaining efficient and irreducible $\mathcal{E}$, because efficient systems have bounded cardinality, and the set of eigenvalues of non-negative integer irreducible matrices of bounded size is discrete, see Lemma 6.6. Thus we must finally obtain a system with B having no monogons, or we obtain $\mathcal{E}$ with $\lambda = 1$.

In the case that $\lambda = 1$ the matrix $M(\mathcal{E})$ is a permutation matrix, see Lemma 6.3. If the incidence matrix is a permutation matrix, clearly the branched manifold $B(\mathcal{E})$ has no branch locus and hence no monogons.

Finally we consider the case $n = 2$. It suffices to give arguments to replace Statements 1 and 2.

Change 1: Here we cannot assume that K intersects H_0 only at its tip, but iterates of K automatically have disjoint interiors. Let V and $\partial V = u \cup v$ be as above. We proceed as above, reducing the complexity of K. If $\mathring{K} \cap H_0$ is non-empty, we may come to a situation where $V \subset H_0$ and $u \subset \partial H_0$. Then $\mathcal{E}$ contains a ∂-parallel arc. In this case we can immediately apply Lemma 1.6 to obtain a new $\mathcal{E}$ with smaller eigenvalue.

Change 2: For $n = 2$ and $i \geq 1$ the "first case" $V \subset H_i$ and $u \subset \partial H_i$ cannot occur, for it would yield a P_i without inner boundary. $\square$

For $n = 2$ the f-invariant lamination Λ considered here appears to be closely related to the stable lamination for f, given by Nielsen-Thurston theory. The precise relationship, however, remains to be established; the fact that in general the two laminations do not quite agree can be easily seen in the case $\lambda = 1$.

We now generalize our basic construction as presented in Section 1. The invariant measured lamination Λ constructed in Propositions 1.4 and 4.1 from a system $\mathcal{E}$ has one important disadvantage: It may not fill $\mathring{H}$, by which we mean that it may not be the case that every essential loop meets Λ. The only way to change this is to enlarge the system $\mathcal{E}$. This can be done such that $\mathcal{E}$ defines a lamination Λ with a filtration $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_k = \Lambda$, where the leaves of $\Omega_i = \Lambda_i - \Lambda_{i-1}$ form a lamination in $\mathring{H} - \Lambda_{i-1}$, but not necessarily in $\mathring{H}$. We say that Λ has a stratified measure relative to this filtration, if each Ω_i has a transverse measure μ_i of full support. As a shorthand, we use the notation $(\Lambda, \mu_1, \ldots, \mu_k)$ to denote the lamination with a stratified measure. Here the lamination (Λ_1, μ_1) will be the same as the lamination (Λ, μ) constructed in Proposition 1.4.

To find such a lamination with stratified measure which fills H, we first consider any complete decomposing system $\mathcal{E}$ with a filtration $\mathcal{E}_1 \subset \mathcal{E}_2 \subset \cdots \subset \mathcal{E}_k = \mathcal{E}$ of subsystems, corresponding to the block decomposition of $M(\mathcal{E})$ as a possibly reducible matrix with irreducible diagonal blocks M_i , see Section 6. Here the i-th block corresponds to the discs of $\mathcal{E}_i - \mathcal{E}_{i-1}$, and the discs of $f(\mathcal{E}_i - \mathcal{E}_{i-1})$ may contain copies of discs in $\mathcal{E}_{i-1}$. Then $B = B(\mathcal{E})$ is defined as in Section 1, but has a filtration $B_1 \subset B_2 \subset \cdots \subset B_k = B$ of branched manifolds, with $B_i = B(\mathcal{E}_i)$. An eigenvector x_i of the irreducible diagonal block matrix M_i with Perron-Frobenius eigenvalue $\lambda_i = \lambda(M_i)$ gives, as before, a weight function on $B_i - B_{i-1}$, which satisfies the switch conditions on the branch spheres contained in $B_i - B_{i-1}$. Thus it defines a measured "lamination with boundary" Ω_i^* carried by $B_i - B_{i-1}$. We can embed Ω_i^* and Λ_{i-1} disjointly in $N(B_i)$ transverse to I-fibers. The leaves of Ω_i^* have punctures corresponding to $(n-1)$-discs in $f^t(\mathcal{E}_{i-1})$, which we must cap off. We can extend Ω_i^* canonically in $N(B_i) - \Lambda_{i-1}$, capping off boundary components with discs transverse to I-fibers in $\pi^{-1}(B_{i-1}) \subset N(B_i)$, to give a lamination Ω_i, with transverse measure μ_i. Clearly one has $f(\Omega_i, \mu_i) = (\Omega_i, \lambda_i \mu_i)$. Define $\Lambda_i = \Omega_1 \cup \Omega_2 \cup \ldots \cup \Omega_i$.

We now claim that we can modify B so that there is no monogon for any of the $B_i - B_{i-1}$. Here we say that $B_i - B_{i-1}$ has a monogon if B_i has a monogon whose boundary, including the tip, is disjoint from B_{i-1}.

To justify this claim, we begin with $B_1 = B(\mathcal{E}_1)$, and we perform the elementary moves as described in the proof of Proposition 4.1, to obtain a new system, also denoted $\mathcal{E}_1$, with no monogons for B_1. In the process of changing $\mathcal{E}_1$ we may have changed the decomposing discs belonging to higher strata, but we can always regard the current system as another stratified decomposing system of the type described above, with all data derived analogously.

We now consider $B_2 - B_1$. If there are monogons with boundary in $B_2 - B_1$, we apply the proof of Proposition 4.1 to $\mathcal{E}_2$, using the moves of Section 3, to eliminate the monogon and reduce λ_2. We leave it to the reader to check that all moves of Section 3 involve only discs of $\mathcal{E}_2$, thus $\mathcal{E}_1, B_1$ and Λ_1 are unaffected. By a slight adaption of the proof of Proposition 4.1 to this relative case, one can ultimately obtain a branched manifold B_2 without monogons in $B_2 - B_1$.

Repeating this process inductively, one can show:

4.2. *Suppose $f : H \to H$ is a diffeomorphism of an n-dimensional handlebody, $n \neq 3, 4$. Then in $\mathring{H}$ there is a lamination with a stratified measure $(\Lambda, \mu_1, \ldots, \mu_k)$, such that, after a suitable isotopy of f, one has*

$$f(\Lambda, \mu_1, \mu_2, \ldots, \mu_k) = (\Lambda, \lambda_1 \mu_1, \ldots, \lambda_k \mu_k)$$

for some $\lambda_i \geq 1$. The lamination Λ is the increasing union of laminations Λ_i fully carried by branched manifolds B_i, and the support of μ_i is $\Omega_i = \Lambda_i - \Lambda_{i-1}$. Further, the branched manifold $B_i - B_{i-1}$ has no monogons in H, and Λ fills $\mathring{H}$.

§5. Train track maps from invariant laminations.

Let Γ be a graph, and $f : \Gamma \to \Gamma$ a continuous map. Recall from the introduction that a path $\beta : [0, 1] \to \Gamma$ is an *f-backtracking path* if β is locally injective, satisfies

$f\beta(0) = f\beta(1)$, and $f\beta$ defines a contractible loop in Γ. This is equivalent, for the universal covering $\tilde{\Gamma}$ of Γ and for lifts $\tilde{\beta} : [0,1] \to \tilde{\Gamma}$ and $\tilde{f} : \tilde{\Gamma} \to \tilde{\Gamma}$ of f, to the condition that $\tilde{\beta}$ is injective and $\tilde{f}\tilde{\beta}$ is closed. In particular, $\tilde{\beta}$ is an $\tilde{f}$-backtracking path. If β is injective we will not distinguish between β and $\mathrm{im}(\beta)$.

5.1. Definition. (Compare [JGLL], [L].) A homotopy equivalence $f : \Gamma \to \Gamma$ is called a *partial train track map* relative to a (possibly non-connected) subgraph $\Gamma' \subset \Gamma$ if $f(\Gamma') \subset \Gamma'$ and if for any $t \geq 1$ all f^t-backtracking subpaths contained in an edge of $\Gamma - \Gamma'$ are mapped by f^t to Γ'. We further assume that no subgraph of Γ' is homotopy equivalent to Γ, and that Γ' contains all vertices of Γ.

The goal of this section is to prove Theorem II of Section 0. We show that every f-invariant lamination in H, as given by Theorem 1, canonically induces a partial train track map. At the end of the section we shall show how iterated partial train track maps define a stratified train track map.

Let $\mathcal{E}$ be a system in H_0. As before, we abuse notation by writing $\mathcal{E} = \cup_{E_i \in \mathcal{E}} E_i$. We associate to $\mathcal{E}$ a *dual graph* $\Gamma = \Gamma(\mathcal{E})$ and a quotient map $p : H_0 \to \Gamma$ as follows: Consider fibered neighborhoods $N(\mathcal{E}) = \cup(E_i \times I) \subset H_0$ and partition H_0 into subsets each of which is either an $(n-1)$-dimensional leaf of the form $E_i \times \{t\}$ for some $t \in \mathring{I}$, or is a component of $H_0 - \mathring{N}(\mathcal{E})$. Each of these complementary components of $\mathring{N}(\mathcal{E})$ is collapsed by the quotient map p to a vertex of Γ, and each product $E_i \times I$ is collapsed to an edge e_i, where every 2-dimensional fiber $E_i \times \{t\}$ is collapsed to a point on the edge e_i. Thus p collapses every element of the decomposition to a point of Γ.

As in Section 1, we assume that $f(N(\mathcal{E})) \cap H_0 \subset \mathring{N}(\mathcal{E})$ and that f respects the product structure. We claim that f^{-1} induces a map $\bar{f} : \Gamma \to \Gamma$. To see this, we observe that the homeomorphism f^{-1} maps H_0 to $f^{-1}(H_0) = H_{-1} \subset \mathring{H}_0$. Notice that the pair (H_0, H_{-1}) is homeomorphic to the pair (H_1, H_0). Thus we see that each element of the decomposition of H_0 is mapped by f^{-1} onto an element of the corresponding decomposition of H_{-1}, which is contained in a well-defined element of the decompositon of H_0. Hence f^{-1} induces a map $\bar{f} : \Gamma \to \Gamma$, which satisfies $\bar{f}p = pf^{-1}$.

If $\mathcal{E}$ is complete, then the map p is a homotopy equivalence, and hence $\bar{f}$ is also a homotopy equivalence.

We will show that the map $\bar{f}$ reflects the properties of the invariant lamination and the branched manifold B. In particular, if one embeds the graph Γ in H_0 geometrically dual to $\mathcal{E}$, then for any edge e_i of Γ one can read off the image path $\bar{f}^t(e_i)$ by simply recording the sequence of intersections of e_i with the $(n-1)$-balls of $f^t(\mathcal{E})$.

5.2. Proposition. *Let $\mathcal{E} \in H_0$ be a complete system and let $\hat{\mathcal{E}} \subset \mathcal{E}$ be a subsystem which is not ∂-parallel. Let $\Gamma = \Gamma(\mathcal{E})$ be the dual graph and let Γ' be the subgraph consisting of all vertices of Γ and all edges of the form $p(E_i \times I)$ with $E_i \in \mathcal{E} - \hat{\mathcal{E}}$. If the branched manifold $B(\hat{\mathcal{E}})$ has no monogons, then $\bar{f} : \Gamma \to \Gamma$ is a partial train track map relative to Γ'.*

Proof. We first observe that, as $\hat{\mathcal{E}}$ is not ∂-parallel, it follows that no subgraph of Γ' is homotopy equivalent to Γ.

Suppose that $\bar{f}$ is not a partial train track map relative to Γ'. Recall that this means that there is an $\bar{f}^t$-backtracking path β, for some $t \geq 1$, contained in some edge e_i of $\Gamma - \Gamma'$, with $\bar{f}^t(\beta)$ not contained in Γ'. In this case it is easy to see that by possibly shortening β we may well assume that the point $\bar{f}^t(\partial\beta)$ is not contained in Γ'.

Let $\{x\} \times I \subset E_i \times I = N(E_i)$ be mapped by $p : H_0 \to \Gamma$ homeomorphically to e_i. Let $b \subset \{x\} \times I$ be the inverse image of β under $p|_{\{x\}\times I}$, and let y and z be the endpoints of b. Since $\bar{f}^t(\partial\beta)$ lies in in $\Gamma - \Gamma'$, it follows that $f^{-t}(y)$ and $f^{-t}(z)$ are contained in $\mathring{N}(\hat{\mathcal{E}})$. The assumption that β is $\bar{f}^t$-backtracking implies that $\bar{f}^t(\beta)$ is closed. Hence $f^{-t}(y)$ and $f^{-t}(z)$ lie in the same $(n-1)$-dimensional fiber $E_j \times \{s\}$ of the product $\mathring{N}(\hat{\mathcal{E}}) = E_i \times \mathring{I}$. This is equivalent to the statement that y and z lie in the same $(n-1)$-dimensional fiber of the product $f^t(\mathring{N}(\hat{\mathcal{E}}))$.

Let c be a path in this fiber which connects y to z. Then $p(c) = p(y) = p(z)$, and hence p maps the loop $f^{-t}(b \cup c)$ to the loop $\bar{f}^t(\beta)$. But the latter is contractible as β is $\bar{f}^t$-backtracking, and, since p is a homotopy equivalence, it follows that the path $b \cup c$ is also contractible. In terms of the branched manifold $B = B(\hat{\mathcal{E}})$, we have b contained in an I-fiber of $N(B)$ and c contained in an $(n-1)$-disc parallel to a disc of $f^t(\hat{\mathcal{E}})$. This gives a map of a half-disc as in the hypotheses of Lemma 2.1. Hence we conclude from that lemma that B has an embedded monogon, a contradiction. □

For later use, we point out the following technical property of the map $\bar{f}$: An edge $e \subset \Gamma - \Gamma'$ is not mapped to a "tight" path in Γ. On the contrary, any point x of e which is mapped by $\bar{f}$ to a vertex of Γ is contained in a non-degenerate subsegment of e on which is $\bar{f}$ is constant. This applies in particular to the endpoints of e. We call a map with this property *end-stationary* relative to Γ'.

In order to prove Theorem 2 in Section 0, we will first derive the following consequence of Proposition 5.2. Notice that in [GJLL] this statement has been derived alternatively as corollary of the existence of relative train track maps, proved in [BH, Theorem 5.12].

5.3. Corollary. *If $h_0 : \Gamma_0 \to \Gamma_0$ is a homotopy equivalence of a graph Γ_0 then there is a graph Γ and a partial train track map $h : \Gamma \to \Gamma$ relative to some subgraph $\Gamma' \subset \Gamma$, as well as a homotopy equivalence $l : \Gamma_0 \to \Gamma$ such that hl and lh_0 are homotopic. Here Γ_0 and hence Γ may well have more than one connected component, and h can assumed to be end-stationary.*

In particular, every automorphism of a free group F_g is realized by a partial train track map.

Proof. We first find a homeomorphism $f : H \to H$ for some n-dimensional handlebody H with $n \geq 5$, and a homotopy equivalence $j : \Gamma_0 \to H$ with $fj = jh_0$ up to homotopy. These exist since every automorphism of a free group can be realized by a handlebody homeomorphism. By Lemma 1.1 we can assume that f is outward expanding with respect to some concentric subhandlebody $H_0 \subset H$. By Proposition 4.1 we find a system $\hat{\mathcal{E}} \subset H_0$ which is not ∂-parallel and for which the associated branched manifold $B(\hat{\mathcal{E}})$ does not have monogons. We define $\mathcal{E}$ to be any extension of $\hat{\mathcal{E}}$ to a complete system.

Now we apply Proposition 5.2 to obtain a graph Γ dual to $\mathcal{E}$, a partial train track map $h : \Gamma \to \Gamma$ relative to some subgraph Γ' and a homotopy equivalence $p : H \to \Gamma$ with $hp = pf$. Then the composition pj is the desired homotopy equivalence l. The extension to the case where Γ_0 has more than one component is trivial, since neither the connectedness of H nor that of Γ has been used in this paper up to now. Notice that the map h is by construction end-stationary. $\square$

We now prove Theorem 2 from Section 0. We would like to point out that an alternative proof, analogous to that of Proposition 5.2, can be derived from statement 4.2, if one extends Lemma 2.1 and Proposition 2.2 to laminations with stratified measure.

Recall that a map $h : \Gamma \to \Gamma$ is a stratified train track map if there is an h-invariant filtration $\Gamma_1 \subset \Gamma_2 \subset \cdots \subset \Gamma_r = \Gamma$ such that $h|_{\Gamma_i}$ is a partial train track map relative to Γ_{i-1}.

5.4. Theorem. *Every homotopy self-equivalence $h : \Gamma \to \Gamma$ of a (possibly disconnected) graph Γ is conjugate, up to homotopy, to a stratified train track map.*

In particular, every automorphism of F_g is realized by a stratified train track map.

Proof. We prove the claim by induction over $-\chi(\Gamma_*)$, where Γ_* is the subgraph of Γ obtained by deleting all contractible components. Notice that the statement of our theorem follows directly if $\chi(\Gamma_*) = 0$.

By Corollary 5.3 the given homotopy self-equivalence $h : \Gamma \to \Gamma$ is conjugate, up to homotopy, to a partial train track map $h_1 : \Gamma_1 \to \Gamma_1$ relative to some subgraph Γ_1'. We observe that h_1 restricts to a homotopy equivalence $h_1' : \Gamma_1' \to \Gamma_1'$. (This follows from Whitehead's theory for free group automorphisms, or alternatively, from a combination of the strong version of Kurosh's Theorem, see [C, p. 175], and of Grushko's Theorem.) Since no subgraph of Γ_1' is homotopy equivalent to Γ_1, we can use our induction hypothesis and find a stratified train track map $h_2 : \Gamma_2 \to \Gamma_2$ relative to some stratification $\Gamma^1 \subset \Gamma^2 \subset \cdots \subset \Gamma^r = \Gamma_2$, which is conjugate, up to homotopy via some homotopy equivalence g, to the map h_1': One has $h_2 g = g h_1'$ up to homotopy. For any point $x \in \Gamma_1'$ let γ_x be the path connecting $gh_1'(x)$ to $h_2 g(x)$ which is defined by this homotopy.

We use the map g to glue, for each edge e of $\Gamma_1 - \Gamma_1'$, a homeomorphic copy $\hat{e}$ along its boundary $\partial\hat{e}$ to $g(\partial e) \subset \Gamma_2$. We denote the graph obtained in this way $\hat{\Gamma}$ and extend the map g to a map $\hat{g} : \Gamma_1 \to \hat{\Gamma}$ by the canonical homeomorphism $e \to \hat{e}$ on every edge $e \subset \Gamma_1 - \Gamma_1'$. A map $\hat{h} : \hat{\Gamma} \to \hat{\Gamma}$ is defined by $\hat{h}(x) = h_2(x)$ if $x \in \Gamma_2$ and by $\hat{h}(x) = \hat{g} h_1 \hat{g}^{-1}(x)$ for all points $x \in \hat{e}_* \subset \hat{\Gamma} - \Gamma_2$. Here $\hat{e}_*$ denotes all of $\hat{e}$ except for small boundary segments on which $\hat{g} h_1 \hat{g}^{-1}$ is constant with value $\hat{g} h_1 \hat{g}^{-1}(\partial\hat{e})$; these exist since h_1 is end-stationary. For any such boundary segment with endpoint $x \in \partial e$ we define the $\hat{h}$-image to be the path $\gamma_{\hat{g}^{-1}(x)}$. Thus $\hat{h}$ satisfies $\hat{h}\hat{g} = \hat{g} h_1$ up to homotopy, and hence it is conjugate, up to homotopy, to our original homotopy self-equivalence $h : \Gamma \to \Gamma$. In particular $\hat{h}$ is a homotopy equivalence. Note that for all points $x \in \Gamma_1 - \Gamma_1'$ with $\hat{h}\hat{g}(x) \in \hat{\Gamma} - \Gamma_2$ we actually have that $\hat{h}\hat{g}(x) = \hat{g} h_1(x)$ exactly (not just up to homotopy).

We claim that $\hat{h}$ inherits from h_1 the property of being a partial train track map relative to Γ_2. This implies that $\hat{h}$ is a stratified train track map, relative to the stratification $\Gamma^1 \subset \Gamma^2 \subset \cdots \subset \Gamma^r = \Gamma_2 \subset \hat{\Gamma}$, which establishes the desired conclusion.

To show that $\hat{h}$ is a partial train track map relative to Γ_2, let $\hat{\beta} \subset \hat{e}$ be an $\hat{h}^t$-backtracking path for some $t \geq 1$ and some edge $\hat{e}$ of $\hat{\Gamma} - \Gamma_2$. If $\hat{h}^t(\hat{\beta})$ is not contained in the subgraph $\Gamma_2 \subset \hat{\Gamma}$, we can assume, replacing $\hat{\beta}$ by a suitable subpath as in the proof of Proposition 5.2, that it has endpoints $\hat{y}$ and $\hat{z}$ which are mapped by $\hat{h}^t$ to $\hat{\Gamma} - \Gamma_2$: Define $\hat{x} = \hat{h}^t(\hat{y}) = \hat{h}^t(\hat{z}) \in \hat{\Gamma} - \Gamma_2$.

Recall that $\hat{g}$ is injective on $\Gamma_1 - \Gamma_1'$. The fact that $\hat{h}\hat{g} = \hat{g}h_1$, up to homotopy, implies $\hat{h}^t\hat{g} = \hat{g}h_1^t$ up to homotopy for any $t \geq 1$, where, as above, the homotopy does not involve points in $\hat{\Gamma} - \Gamma_2$ which are mapped by $\hat{h}^t\hat{g}$ to $\hat{\Gamma} - \Gamma_2$. Therefore, since $\hat{x} \in \hat{\Gamma} - \Gamma_2$, for the points $x = \hat{g}^{-1}(\hat{x}), y = \hat{g}^{-1}(\hat{y})$ and $z = \hat{g}^{-1}(\hat{z})$ in $\Gamma_1 - \Gamma_1'$, one has $x = h_1^t(y) = h_1^t(z)$.

Let $\beta = \hat{g}^{-1}(\hat{\beta}) \subset e \subset \Gamma_1 - \Gamma_1'$. Then $\partial\beta = \{x, y\}$, and $h_1^t(\beta)$ is closed. Since $\hat{g}$ is a homotopy equivalence, it follows from the assumption that $\hat{\beta}$ is $\hat{h}^t$-backtracking that $h_1^t(\beta)$ is a contractible loop in Γ_1. Thus β is h_1^t-backtracking, and hence the partial train track property of h_1 implies that $h_1^t(\beta)$ is contained in the subgraph Γ_1'. But this contradicts the fact that $h^t(\partial\beta) = \{x\} \subset \Gamma_1 - \Gamma_1'$.

Hence $\hat{\beta}$ must be contained in Γ_2, and thus $\hat{h}$ is a partial train track map relative to Γ_2. $\square$

We now give an explanation, promised in the introduction, to show why it would be unreasonable to expect to obtain a compact f-invariant measured lamination in the statement of Theorem 1. Suppose that we had a properly embedded f-invariant measured lamination (Λ, μ), and suppose further, as is reasonable, that the lift $(\tilde{\Lambda}, \tilde{\mu})$ to the universal cover were dual to an $\mathbb{R}$-tree $\mathcal{T}$. Then the tree would be projectively invariant under the action of the automorphism f_* of F_g induced by f. It is proved in [GJLL] that for most automorphisms f_*, the action is "non-geometric," which implies that it is not dual to the lift of a measured lamination carried by a compact branched manifold. (It is possible to prove indirectly, using Proposition 5.2 of this paper and results in [GJLL] that for our f-invariant lamination, $(\tilde{\Lambda}, \tilde{\mu})$ is in fact dual to an $\mathbb{R}$-tree.)

§6. Non-negative matrices.

The theory of non-negative matrices is well-developed and by now classic; several expository books are available. We will use [G] as standard reference; we also recommend the short treatment in [DV].

A matrix $A = (a_{ij})$ is called *non-negative* if all entries are real numbers which satisfy $a_{ij} \geq 0$. Such a matrix is called *reducible* if after a suitable permutation of indices, the matrix takes the shape of an upper triangular block matrix with more than one block. Otherwise it is called *irreducible.* For a reducible matrix, by refining the block structure if necessary, we can always assume that the diagonal blocks are irreducible.

The most basic fact in the theory is that for a non-negative matrix A is that there is always a real eigenvalue $\lambda = \lambda_A \geq 0$, called the *Perron-Froebenius eigenvalue*, which realizes the spectral radius of A, and that there is always a corresponding eigenvector x, called a *Perron-Frobenius eigenvector*, which has non-negative entries. If A is irreducible, then among eigenvectors x with non-negative entries, x is uniquely determined up to scalar multiplication.

We would like to remind the reader of the following characterization of the Perron-Frobenius eigenvalue for any (possibly reducible) non-negative matrix A. If A is irreducible this is shown, for example, in [G], §13 (40); otherwise, use the fact that every reducible matrix is the limit of irreducible non-negative matrices:

6.1.(a) $\lambda_A = \inf\{\max\{(Aw)_k/w_k\} \mid w = (w_k) \text{ a vector with positive entries}\}$

(b) $\quad = \sup\{\min\{(Aw)_k/w_k\} \mid w = (w_k) \text{ a vector with positive entries}\}$

The following fact follows directly from the definition of irreducibility, see [G].

6.2. Lemma. *Each of the following properties of the matrix $A \neq 0$ is equivalent to the irreducibility of A.*

(1) *For every index pair (i,j) there exists a $t \geq 1$ such that the (i,j)-th entry of A^t is positive.*

(2) *For any two entries $a_{i_1 j_1}$, $a_{i' j'}$ there exists a sequence of index pairs $\{(i_k, j_k)\}$, $k = 2, \dots, t$, such that $i_t = i', j_t = j'$ and for all $k \geq 1$*

(a) $i_k = j_{k-1}$, and

(b) $a_{i_k j_k} > 0$.

From 6.1 (b), considering $w = (1, 1, \dots, 1)$, and from the definition of λ_A as spectral radius, one deduces (compare also [DV], Proposition II.1.10):

6.3. Lemma. *Let $A = (a_{ij})$ be a non-zero, irreducible matrix with non-negative integer entries. Then the Perron-Frobenius eigenvalue satisfies $\lambda \geq 1$. If $\lambda = 1$, then A is a transitive permutation matrix.*

This implies the following:

6.4. Lemma. *Every non-negative matrix A (possibly reducible) with $\lambda_A = 0$ satisfies $A^k = 0$ for some $k \geq 1$.*

6.5. Lemma. *Let $A = (a_{ij})$ and $A' = (a'_{ij})$ be two non-negative matrices of same size, with Perron-Frobenius eigenvalues λ_A and $\lambda_{A'}$.*

(a) If $a_{ij} \geq a'_{ij}$ holds for all indices i, j, then it follows $\lambda_A \geq \lambda_{A'}$.

(b) Assume furthermore that A is irreducible and that there is at least one pair of indices with strict inequality $a_{ij} > a'_{ij}$, then it follows $\lambda_A > \lambda_{A'}$.

Proof. (a) This follows directly from the definition of the Perron-Frobenius eigenvalue as spectral radius.

(b) If A' is irreducible, this is proved in [G], §13.2 , Satz 6. If A' is reducible, then let A'_* be the irreducible diagonal block submatrix of A' with $\lambda_{A'}$ as eigenvalue, and let A_* be the corresponding submatrix of A. For $0 < \mu < 1$ let A_μ be the matrix which agrees with A on A_* , and with μA everywhere else. Thus A_μ is irreducible, and from the case A' irreducible, considered previously, we obtain the inequalities $\lambda_{A_\mu} < \lambda_{A_{\mu'}} < \lambda_A$ if $0 < \mu < \mu' < 1$. Since $\lambda_{A'} = \lambda_{A'_*} \leq \lambda_{A_*} \leq \lambda_{A_\mu}$, this proves our claim. $\square$

6.6. Lemma. *For any infinite set of non-negative irreducible integer matrices $A(n)$ the set of Perron-Frobenius eigenvalues $\lambda_{A(n)}$ is unbounded.*

Proof. Since $\{A(n)|n \in \mathbb{N}\}$ is infinite, there must be at least one index pair (i, j) such that the entry $a_{i,j}(n)$ of $A(n)$ is unbounded for $n \in \mathbb{N}$. Let $A'(n)$ be obtained from A(n) by replacing every non-zero entry distinct from $a_{i,j}(n)$ by 1. Then $A'(n)$ is irreducible, and Lemma 6.5 (a) implies $\lambda_{A(n)} \geq \lambda_{A'(n)}$. By further replacing some of the entries equal to 1 by 0's we can change each $A'(n)$ into a matrix $A''(n)$ which is obtained from a permutation matrix by changing the (i, j)-entry from 1 to $a''_{i,j}(n) = a_{i,j}(n)$. It satisfies $\lambda_{A'(n)} \geq \lambda_{A''(n)}$. A suitable power $A''(n)^k$ with $k \leq n!$ is a diagonal matrix with at least one diagonal entry bigger or equal to $a_{i,j}(n)$. As these are unbounded, it follows that the set $\lambda_{A''(n)}$ and hence the set $\lambda_{A(n)}$ is unbounded. □

We consulted M. Boyle, M. Newmann, R. Thompson, and R. Williams about the following lemma. They confirmed that the statement was correct and offered two proofs.

6.7. Lemma. *Let A be an irreducible non-negative matrix, and, for indices $i \neq j$, let A_i be obtained from A by replacing the i-th column by the sum of the i-th and the j-th column and then erasing the j-th column and row. Similarly, let A_j be obtained from A by replacing the j-th column by the sum of the i-th and the j-th column, and erasing the i-th column and row. Then one has*

$$\lambda_{A_i} \leq \lambda_A \qquad \text{or} \qquad \lambda_{A_j} \leq \lambda_A .$$

Proof (essentially Boyle's version). Let A'_i be obtained from A by replacing the j-th row by an identical copy of the i-th row. Let N differ from the identity matrix only in that the entry (j, i) is equal to 1. Then $N^{-1}A'_iN$ has as i-th column the sum of the i-th and the j-th column of A, except that all entries of its j-th row are equal to 0. Everywhere else it agrees with A. We have $\lambda_{A_i} = \lambda_{A'_i}$.

As a consequence it suffices to prove the above inequalities for A'_i and A'_j (defined analogously) rather than A_i and A_j: Let the column v be a Perron-Frobenius eigenvector of A, and suppose without loss of generality that its entries satisfy $v_i \leq v_j$. Let v' be obtained from v by replacing the entry v_j with v_i. This gives $A'_iv = \lambda_A v'$. As the entries of v' and v satisfy $v'_k \leq v_k$, it follows $(A'_iv')_k \leq (A'_iv)_k = \lambda_A v'_k$. Since A is irreducible, the eigenvector v and hence the vector v' has positive entries, and thus 6.1 (a) directly gives $\lambda_{A'_i} \leq \lambda_A$. □

BIBLIOGRAPHY

[BH] M. Bestvina and M. Handel, *Train tracks and automorphisms of free groups*, Annals of Math. **135** (1992), 1-51.

[C] D. Cohen *Combinatorial Group Theory: a topological approach*, Cambridge U. Press, 1989.

[DV] W. Dicks and E. Ventura, *The Group Fixed by a Family of Injective Endomorphisms of a Free Group*, Contemporary Math. **195** (1996).

[FLP] A. Fathi, F. Laudenbach and V. Poenaru, *Travaux de Thurston sur les surfaces*, Asterisque 66-67.

[FO] W. Floyd and U. Oertel, *Incompressible surfaces via branched surfaces*, Topology **23** (1984), 117-125.

[G] F. Gantmacher, Matrizentheorie, Springer Verlag, Berlin, 1986.

[GO] D. Gabai, U. Oertel, *Essential laminations in 3-manifolds*, Annals of Math. **130** (1989), 41-73.

[JGLL] A. Jäger, D. Gaboriau, G. Levitt and M. Lustig, *An index for counting fixed points of automorphisms of free groups*, preprint 1996.

[L] M. Lustig, *Perron-Frobenius action of free groups on* $\mathbb{R}$*-trees*, preprint 1996.

[O] U. Oertel, *Measured laminations in 3-manifolds*, Transactions AMS **305**, No.2, (1988) 531-573.

[T] W. P. Thurston, *On the geometry and dynamics of diffeomorphisms of surfaces*, Bull. A.M.S. **19**, (1988), 417-431.

[W] R. Williams, *Expanding attractors*, Pub. Math. I.H.E.S. **43** (1974), 169-203.