STRUCTURE AND CONJUGACY FOR AUTOMORPHISMS OF FREE GROUPS I

Martin Lustig

0. Introduction and statement of results

The purpose of this paper and the subsequent Part II (= sections 5 - 9)[1] is to present a complete algorithmic solution of the following two problems, where F_n denotes a (non-abelian) free group F_n of finite rank $n \geq 2$:

Conjugacy problem for automorphisms of F_n.
(a) Given automorphisms $\alpha, \alpha' \in Aut(F_n)$, does there exist an automorphism $\beta \in Aut(F_n)$ such that $\alpha' = \beta\alpha\beta^{-1}$?
(b) Given $\alpha, \alpha' \in Aut(F_n)$, does there exist $\beta \in Aut(F_n)$ such that $\alpha' = \beta\alpha\beta^{-1}$ up to inner automorphisms of F_n ?

A solution of these problems has been announced in [13], and solutions for special cases have been given in [4], [8], [9], [10] and [17]. Related structural informaton about automorphisms of free groups has been derived in [2] and in [18]. The solution presented here is a combinatorial one, and can (not just in principle) be programmed on a computer of standard nowadays capacities. For small cases it can well be performed by hand: It is efficient in the human sense, i.e. in most cases it requires few time, while we also can concoct some rare cases where all of the more sophisticated invariants presented here have to be computed in order to show that the difficulty in finding a conjugator is not just due to bad luck or insufficient persistence.

As automorphisms of free groups are enumerable, a "positive" machine, i.e. an algorithm which stops if a conjugator β as above is found, and keeps running otherwise, is given directly. All the difficulty stays in finding an analogous "negative" machine: This amounts to the exhibition of sufficiently many structural data which are both,

(1) invariant under conjugacy, and
(2) algorithmically computable.

---

[1] also to appear in the Max-Planck preprint series. It is presently being brought into final form, compare [14] (M.L., December 15, 2000)

The main such structural data are minimal actions of F_n on a $\mathbb{R}$-tree T by isometries which are *(projectively) α-invariant* in the following sense: There exists a homothety $H : T \to T$ with stretching factor $\lambda \geq 1$ such that $\alpha(w)H = Hw : T \to T$ for all $w \in F_n$. The existence of such actions has been shown in [1], [12] and [19]. A general and geometric construction of such T has been presented in [6] and is rewiewed here in section 1. It is crucially based on the existence of a train track map $f : \Gamma \to \Gamma$ which represents the given $\alpha \in Aut(F_n)$, and T is completely described by means of a Perron-Frobenius eigenvector of the non-negative *geometric transitions matrix* associated to f. We call such actions *Perron-Frobenius actions* of F_n.

The first goal of this paper is to show that there are no other α-invariant $\mathbb{R}$-tree actions with stretching factor $\lambda > 1$. In the case where α admits an absolute train track representative, this is done in sections 1 and 2. In order to treat the general case we introduce and study in section 3 train tracks with 2-cells (called *Nielsen faces*), which can be derived algorithmically from Bestvina-Handel's relative train track representatives [3]. In section 4 we show the following result, which also contributes to a better understanding of the action of α on Culler-Vogtmann's outer space CV_n and its boundary.

Theorem I. *Every α-invariant action of F_n on an $\mathbb{R}$-tree T with stretching factor $\lambda > 1$ is Perron-Frobenius, for any $\alpha \in Aut(F_n)$.*

In particular, any such $\mathbb{R}$-tree is given by a non-negative eigenvector with eigenvalue λ on any partial train track representative $f : \mathcal{G} \to \mathcal{G}$ with Nielsen faces of α, relative to a subspace of polynomial α-growth. The converse is also true.

We then define a special type of such actions, and we show that they are given by special graphical representatives of α which we call *relative pseudo-Anosov* maps. For every exponentially growing automorphism α there exists a non-empty finite family, derived algorithmically, of relativ pseudo-Anosov train track maps which represent all of the special invariant α-invariant actions. In section 5 we show that, if two such relative pseudo-Anosov maps $f : \mathcal{G} \to \mathcal{G}$ and $f' : \mathcal{G}' \to \mathcal{G}'$ define the same $\mathbb{R}$-tree action of F_n, then there exists a conjugating map $h : \mathcal{G} \to \mathcal{G}'$, i.e. $hf = f'h$ up to differences in the vertex spaces and up to certain controlled homotopies involving the 2-cells. The crucial algorithmic insight here is that we can bound the length of the images of the edges a priory, depending only on data derived from f and f'.

In the polynomially growing case the situation is similar, but easier, as there is a unique canonical α-invarinat action on an $\mathbb{R}$-tree, this time with stretching factor $\lambda = 1$. It comes from a graph of groups decomposition of F_n as follows (see section 7 and 8):

Theorem II. *Let $\alpha \in Aut(F_n)$ be polynomially growing, say of degree $d \geq 1$, given by a relative train track representative. We can algorithmically derive a canonical graph-of-groups decomposition $F_n = \pi_1\mathcal{G}$ which is α-invariant. It is uniquely determined by the property that the vertex groups of $\mathcal{G}$ consist precisely of those elements which determine conjugacy classes in F_n which have polynomial* *growth of degree strictly smaller than d, under iteration of α. If $d = 1$, then all edge groups of $\mathcal{G}$ are cyclic, while in the case of quadratic or higher growth, $d \geq 2$, they are trivial.*

A major technical effort in our approach is needed to keep track of all data when passing from one stratum of $\mathcal{G}$ to the next lower one (given by the restriction of α to the stabilizers of the branch points of T), where each time the rank of the free group in question decreases strictly. For this purpose we introduce *fix point pattern* in section 6 and show in section 6 for exponentially growing α and in sections 7 and 8 for polynomially growing α how conjugacy of (orbits) of fixed point pattern can be decided algorithmically, modulo introducing new (orbits) of fixed point pattern for the next level of our recursion procedure. The details of this algorithm are presented in section 9.

It turns out that, after the right tools are developped, the exponentially growing case is easier to decide than the polynomial one, and the hardest among the polynomially growing is the periodic-linear case, treated in [9]. The stronger growth forces more rigidity, which makes it easier to read off computible invariants of the automorphism. For example, it follows from the methods presented here that the centralizer of an outer automorphism is virtually cyclic (i.e. essentially the subgroup generated by the automorphism itself), except for additional factors which are derived on the lowest, the linear level of our recursion procedure. In particular, a direct consequence of the work presented here is:

Corollary III. *If for an outer automorphism $\hat{\alpha} \in Out(F_n)$ no conjugacy class $[w] \subset F_n$ is fixed by $\hat{\alpha}^t$, $(t \geq 1)$, then the centralizer $Cen(\hat{\alpha}) \subset Out(F_n)$ is virtually abelian.*

Here the rank of the finite index abelian subgroup is given by the number of connected components of the canonical graph representing the p.o. set of the relative pseudo-Anosov strata of α with respect to the inclusion.

More generally, the methods and results developped in this paper have some direct applications to other questions on automrphisms: For example, one can derive rather easily a structural result about all virtually abelian subgroups of $Out(F_n)$ much in the spirit of Corollary III above. Other applications concern the complete determination of the dynamics on ∂F_n induced by an automorphism α (joint work with G. Levitt) and a fast proof of P. Brinckmann's result that α is hyperbolic (as defined by Gromov) if no conjugacy class in F_n has a periodic α-orbit.

Acknowledgments. *I would like to thank Gilbert Levitt for many useful conversations which have directly and indirectly helped shaping this work. My thanks go also to Mladen Bestvina, Marshall Cohen and Karen Vogtmann for their interest and encouragement, and to University of Utah, Salt Lake City, Université P.Sabatier, Toulouse, Université S.Antipolis, Nice, and the Max-Planck Institut, Bonn, for their support and hospitality during several extended research stays, where most of this paper has been written. Finally I'd like to thank Yoav Moriah and Karen Vogtmann for motivating me properly to take on the final writing effort.*

Part I: PERRON-FROBENIUS ACTIONS OF FREE GROUPS AND A STRUCTURE THEOREM FOR FREE GROUP AUTOMORPHISMS

The purpose of this Part I is a structural analysis for outer automorphisms of free groups which is in close analogy to (but more complicated than) the Nielsen Thurston classification of surface homeomorphisms. Algorithmic aspects play only a minor role in this Part I; we have tried to put the emphasis in our presentation primarily on the structural results and on the tools needed to understand them.

1. Invariant $\mathbb{R}$-trees through iteration of absolute train track maps

1.1. Let F_n be a free group of rank $n \in \mathbb{N}$, and let $\alpha \in Aut(F_n)$ be an automorphism of F_n. Let T be an $\mathbb{R}$-tree with a non-trivial minimal action of F_n by isometries, and let $H : T \to T$ be a homothety with stretching factor $\lambda > 0$. If for every $w \in F_n$ one has

$$\alpha(w)H = Hw : T \to T$$

F_n-action on T α-invariant.

In [12] it has been shown that for every $\alpha \in Aut(F_n)$ there is an α-invariant minimal F_n-action on some non-trivial $\mathbb{R}$-tree, with trivial arc stabilizers and with $\lambda \geq 1$.

The main goal of this and the following two sections is to determine all such α-invariant actions on $\mathbb{R}$-trees with small arc stabilizers and with $\lambda > 1$. We first recall a particular construction which leads to such actions, built on Perron-Frobenius eigenvectors of the transition matrix of a train track representative of α, and which are hence called here *Perron-Frobenius actions.* We prove in section 2:

Theorem 1.2. *For any automorphism $\alpha \neq 1$ of F_n with an absolute train track representative every α-invariant small minimal action of F_n on an $\mathbb{R}$-tree with stretching factor $\lambda > 1$ is Perron-Frobenius.*

In particular this shows that automorphisms which are represented by a train track map with irreducible transition matrix and eigenvalue $\lambda > 1$ have, up to rescaling, precisely one α-invariant $\mathbb{R}$-tree action with stretching factor strictly bigger than 1. This includes all irreducible automorphisms with irreducible powers.

The iteration tree

Throughout sections 1 and 2 of the paper we use the following notation:

Notation 1.3. Let α be an automorphism of F_n. Let Γ be a graph with $\pi_1\Gamma = F_n$, a free group of finite rank $n \geq 2$, and let $f : \Gamma \to \Gamma$ be a p.l. homotopy equivalence which induces the outer automorphism of F_n defined by α. Let $\tilde{\Gamma}$

denote the universal covering of Γ (a tree !), equipped with the usual action of F_n by covering translations, and let $\tilde{f} : \tilde{\Gamma} \to \tilde{\Gamma}$ denote the lift of f which represents the automorphisms $\alpha \in Aut(F_n)$: For every $w \in F_n$ one has

$$\alpha(w)\tilde{f} = \tilde{f}w : \tilde{\Gamma} \to \tilde{\Gamma}\,. \tag{1}$$

1.4. For a graph Γ we define an edge length function L to be a map $L : Edges(\Gamma) \to \mathbb{R}_{\geq 0}$. Every such function defines an F_n-equvariant pseudo-metric d on $\tilde{\Gamma}$, i.e. a function $d : \tilde{\Gamma} \times \tilde{\Gamma} \to \mathbb{R}_{\geq 0}$ which differs from a metric only in that distinct points may have distance 0. Other pseudo-metrics considered later in this paper are given by pulling back the distance from a metric space X under a continuous map $f : \tilde{\Gamma} \to X$.

In [6] the following construction has been introduced (with slightly more generality):

1.5. Assume that the tree $\tilde{\Gamma}$ is given an F_n-equivariant pseudo-metric d which is δ-hyperbolic. Let $\lambda > 1$ be a number which satisfies

$$d(\tilde{f}(x), \tilde{f}(y)) \leq \lambda d(x, y) \tag{2}$$

for all $x, y \in \tilde{\Gamma}$. We define a new distance function d_∞ on $\tilde{\Gamma}$ through

$$d_\infty(x, y) = lim_{k\to\infty}\; d(\tilde{f}^k(x), \tilde{f}^k(y))/\lambda^k \tag{3}$$

and observe that d_∞ also defines a pseudo-metric on $\tilde{\Gamma}$, which satisfies

$$d_\infty(x, y) \leq d(x, y) \tag{4}$$

and

$$d_\infty(\tilde{f}(x), \tilde{f}(y)) = \lambda d_\infty(x, y) \tag{5}$$

for all $x, y \in \tilde{\Gamma}$. It is not hard to see that, as $\lambda > 1$, the pseudo-metric d_∞ is uniquely determined by the value of d on the vertices of $\tilde{\Gamma}$. In particular it suffices that (2) is satisfied for all vertices $x, y \in \tilde{\Gamma}$.

It follows that there is an $\mathbb{R}$-tree T with an action of F_n by isometries, and a surjective d_∞-distance preserving F_n-equivariant map $i : \tilde{\Gamma} \to T$. This iteration tree $T = T(\Gamma, f, L(\cdot), \lambda)$ is uniquely determined, up to F_n-equivariant isometry, by the graph Γ, the map f, the edge length function $L(\cdot)$ and by λ. In particular there is a homothety $H : T \to T$ with stretching factor λ which satisfies

$$\alpha(w)H = Hw : T \to T \tag{6}$$

for all $w \in F_n$, and

$$i\tilde{f} = Hi : \tilde{\Gamma} \to T\,. \tag{7}$$

We want to state the uniqueness of the iteration tree $T = T(\Gamma, f, L(\cdot), \lambda)$ precisely, in the following form:

Lemma 1.6. *Let T' be any $\mathbb{R}$-tree with F_n-action, and let $i' : \tilde{\Gamma} \to T'$ be an F_n-equivariant map which preserves the d_∞-metric on $\tilde{\Gamma}$. Then there is a (uniquely determined) F_n-equivariant isometric embedding $j : T(\Gamma, f, L(\cdot), \lambda) \to T'$ with $i' = ji$.* □

With the above generality, of course, the tree T may well consist of a single point only. We also want to point out that in general the map i may well "crumble" the edges of $\tilde{\Gamma}$: For any edge e the image $i(e)$ does not need to be a geodesic segment.

Perron-Frobenius metrics and train track maps

Let $f : \Gamma \to \Gamma$ be as in 1.3. We use the following terminology:

Notation 1.7. Let $M(f) = (m_{e,e'})_{e,e' \in Edges(\Gamma)}$ denote the *geometric transition matrix* for f, i.e. the entry $m_{e,e'}$ denotes the number of times the path $f(e')$ crosses or partially crosses over the edge e, disregarding the orientations. As f is assumed to be p.l., each $m_{e,e'}$ is finite. Let $v = (v_e)_{e \in Edges(\Gamma)}$ be a Perron-Frobenius row eigenvector of $M(f)$, i.e. v has all entries $v_e \geq 0$ and eigenvalue $\lambda \geq 1$. Such an eigenvector exists always, by the Perron-Frobenius theorem for non-negative matrices, since $M(f)$ is integer and f is not homotopic to the trivial map. We use $v = (v_e)$ to define the Perron-Frobenius length function $L_v(e) = v_e$ for any edge e of Γ, and notice that the induced pseudo-metric $d : \tilde{\Gamma} \times \tilde{\Gamma} \to \mathbb{R}_{\geq 0}$ and λ satisfy inequality (2) of 1.5 for all vertices x, y of $\tilde{\Gamma}$.

Let Γ be a graph, X any topological space and $f : \Gamma \to X$ a continuous map. As in [6] we call a path $\beta : [0,1] \to \Gamma$ an f-backtracking path if β is locally injective and if $f\beta$ is a contractible loop in X.

Definition 1.8. A homotopy equivalence $f : \Gamma \to \Gamma$ of a finite graph Γ is called *train track map* if f preserves the set $V(\Gamma)$ of vertices of Γ and if there is no f^k-backtracking path in the interior of any edge of Γ, for all $k \geq 1$.

Train track maps $f : \Gamma \to \Gamma$ have the remarkable property that for any edge $\tilde{e}$ of the universal covering $\tilde{\Gamma}$ the image $\tilde{f}^k(e)$, for any $k \geq 1$, is again a geodesic (with respect to any metric on the tree $\tilde{\Gamma}$!). If we provide $\tilde{\Gamma}$ with a Perron-Frobenius length function $L_v(\cdot)$ with eigenvalue $\lambda > 1$, it follows directly that $L_v(\tilde{f}^k(\tilde{e})) = \lambda^k L_v(\tilde{e})$. Hence the length of $\tilde{e}$ in the associated d_∞-pseudo-metric on $\tilde{\Gamma}$ is equal to $L_v(e)$, and thus every edge $\tilde{e} \subset \tilde{\Gamma}$ is mapped by the map $i : \tilde{\Gamma} \to T$, defined as in 1.5, to a geodesic segment of length $L_v(\tilde{e})$. With respect to the pseudo-metric d_∞ this restriction $i_{\tilde{e}} : \tilde{e} \to i(\tilde{e})$ is an isometry, and we will think from now on of the metric d defined by means of the length function $L_v(\cdot)$ as being "reparametrized" along the edges (without changing the total edge length) so that it agrees with d_∞ on every edge of $\tilde{\Gamma}$.

We will call a map from a simplicial tree $\tilde{\Gamma}$ to an $\mathbb{R}$-tree *edge geodesic* if it maps every edge to a geodesic segment. Notice in particular that if $\tilde{\Gamma}$ carries a pseudo-metric which is non-zero, and if every edge is mapped to a geodesic segment of same length, as is the case above, then it follows in particular that the tree T is not just a single point.

Definition 1.9. If $f : \Gamma \to \Gamma$ is a train track map, with edge lengths $L_v(e)$ defined by a non-trivial Perron-Frobenius eigenvector v of $M(f)$ with eigenvalue $\lambda > 1$, then we call the associated iteration tree $T_v = T(\Gamma, f, L_v(\cdot), \lambda)$ a Perron-Frobenius tree and the F_n-action on T_v a Perron-Frobenius action, with respect to the train track map $f : \Gamma \to \Gamma$.

Proposition 1.10. *For any $\mathbb{R}$-tree T_v with Perron-Frobenius F_n-action it follows that*
(a) the tree T_v is non-trivial,
(b) the F_n-action on T_v is minimal, and
(c) the stabilizer of any non-degenerate arc in T_v is trivial (and hence in particular the action is small and even very small).

Proof. A complete proof has been given in [6]. It is assumed there that the transition matrix $M(f)$ is irreducible, but this assumption is not really used in the proof other than to guarantee the existence of some v and λ as assumed here. □

Bounded backtracking

In [6] a the following notion for a free group action on an $\mathbb{R}$-tree has been introduced and used intensively.

Definition 1.11. The action of F_n on the $\mathbb{R}$-tree T has *bounded backtracking* (*BBT*) if for every $Q \in T$ and every basis A of F_n there exists a constant $C \geq 0$ such that for any reduced words $v, w \in F_n$ with reduced product vw one has $dist(vQ, [Q, vQ]) \leq C$, where $[U, V]$ denotes the geodesic segment from U to V.

In the next sections we will use the property BBT explicitly in the following equivalent form:

Lemma 1.12. *The F_n-action on T satisfies BBT if and only if for any graph Γ with $\pi_1\Gamma = F_n$ and any F_n-equivariant map $i : \tilde{\Gamma} \to T$ there is a constant $c \geq 0$ such that the i-image of any geodesic arc $[x, y] \subset \tilde{\Gamma}$ is contained in the c-neighborhood of the geodesic segment $[i(x), i(y)] \subset T$.* □

Remark 1.13. If the F_n-action on the $\mathbb{R}$-tree T satisfies BBT, then so does the induced action of F_n on the metric completion $\hat{T}$ of T. This follows directly from Definition 1.11.

It has been observed in [6] that from a bounded cancellation lemma due to Bestvina-Feighn-Handel (see [2] or [5, Lemma 2.2.4]), the following can be derived:

Propositon 1.14. *Every F_n-action on an $\mathbb{R}$-tree T which is a limit of free simplicial actions (i.e. T represents a point in Culler-Vogtmann's outer space CV_n or its Thurston boundary ∂CV_n) satisfies BBT.* $\square$

2. Proof of Theorem 1.2

We first want to give a criterion for α-invariant trees to be Perron-Frobenius:

Propositon 2.1. *Let $f : \Gamma \to \Gamma$ be a train track map representing $\alpha \in Aut(F_n)$, and let T be an α-invariant $\mathbb{R}$-tree, with respect to a homothety H with stretching factor $\lambda > 1$, as in 1.1. Assume furthermore that T satisfies BBT. If there exists an F_n-equivariant edge geodesic map $i : \tilde{\Gamma} \to T$ which satisfies*

$$i\tilde{f} = Hi\,, \tag{8}$$

then the minimal F_n-invariant subtree $T_0 \subset T$ is F_n-equivariantly isometric to a Perron-Frobenius tree, with respect to f.

Proof. We define a length $L(\tilde{e})$ for every edge $\tilde{e}$ in $\tilde{\Gamma}$ with initial vertex P and terminal vertex Q through

$$L(\tilde{e}) = dist(i(P), i(Q))\,.$$

As i is assumed to be F_n-equivariant, this induces a well defined length function $L : Edges(\Gamma) \to \mathbb{R}_{\geq 0}$ on Γ. Notice that the assumption that i be edge geodesic together with equation (8) implies that for any edge $\tilde{e} \in E(\tilde{\Gamma})$ and any $t \geq 1$ the path $\tilde{f}^t(\tilde{e})$ is mapped by i to a geodesic in T (namely to $H^t i(\tilde{e})$). This shows $\lambda L(e) = length(f(e)) = \sum_{e' \in Edges(\Gamma)} m_{e',e} L(e)$, for every edge e of Γ. Thus the length function L on the edges of Γ and of $\tilde{\Gamma}$ is given by a Perron-Frobenius eigenvector of the transition matrix $M(f) = (m_{e',e})$, with eigenvalue equal to the stretching factor λ of H.

We claim now that i preserves the d_∞-metric on $\tilde{\Gamma}$ which is defined as explained in 1.5 through the length function L. In fact, by definition of L and by the above observation that $\tilde{f}^t(\tilde{e})$ is mapped by i to a geodesic in T, it follows that our claim is true for any two points on $\tilde{f}^t(\tilde{e})$, for any edge $\tilde{e}$ of $\tilde{\Gamma}$. As every finite edge path in $\tilde{\Gamma}$ is a finite concatenation of such T-*isometric* subpaths (i.e. subpaths contained in some $\tilde{f}^t(\tilde{e})$), it follows that $d_\infty(x, y) \geq dist(i(x), i(y))$ for any $x, y \in \tilde{\Gamma}$. Observe that the reduced arc $[x, y]$ in $\tilde{\Gamma}$ which joins x to y is a finite union of subarcs with endpoints of distance 0 with respect to the d_∞-metric, and of (possibly degenerate) T-isometric subpaths, such that the d_∞-lengths of the latter add up to $d_\infty(x, y)$.

Thus, in order to complete the argument for $d_\infty(x, y) = dist(i(x), i(y))$ by means of a proof by contradiction, we can now assume that there are points $x, y \in \tilde{\Gamma}$ with $dist(i(x), i(y)) = 0$ and $d_\infty(x, y) \neq 0$. From the above observation it follows then that there are points $x' \neq y'$ on one of those T-isometric subpaths of $[x, y]$ such that $\tilde{f}^t(x')$ and $\tilde{f}^t(y')$ are contained in the geodesic arc connecting $\tilde{f}^t(x)$ to $\tilde{f}^t(y)$, for all $t \geq 1$. But as $dist(i\tilde{f}^t(x'), i\tilde{f}^t(y')) \to \infty$ for $t \to \infty$, this contradicts the assumption that T satisfies BBT, as $dist(i\tilde{f}^t(x), i\tilde{f}^t(y)) = dist(H^t i(x), H^t i(y)) = \lambda^t dist(i(x), i(y)) = 0$ for all $t \geq 1$.

With these data we consider now the iteration tree $T(\Gamma, f, L(\cdot), \lambda)$ as defined in 1.3, and let $\hat{i} : \tilde{\Gamma} \to T(\Gamma, f, L(\cdot), \lambda)$ be the canonical d_∞-preserving map defined there. From the above claim and Lemma 1.6 it follows that there is an F_n-equivariant isometric embedding $j : T(\Gamma, f, L(\cdot), \lambda) \to T$, and as $T(\Gamma, f, L(\cdot), \lambda)$ is minimal by Proposition 1.10 (b), it follows that $j(T(\Gamma, f, L(\cdot), \lambda))$ agrees with the minimal subtree $T_0 \subset T$. Thus T_0 is Perron-Frobenius with respect to f. $\square$

We now want to construct a map i as in the previous Proposition 2.1, first for the vertices, and then extend it to the edges. We crucially use the dynamics of H to define the right such map. Notice that until the next proposition (inclusively) we do not need the train track property of f nor the fact that T satisfies BBT.

2.2. Let $H : T \to T$ be as in 1.1. Let $\bar{T}$ be the metric completion of T, and let $\bar{H} : \bar{T} \to \bar{T}$ be the (well defined) continuous extension of H. It follows that $\bar{H}$ as well as every $w\bar{H}^m : \bar{T} \to \bar{T}$ $(w \in F_n,\ m \geq 1)$ has a fixed point $Q(w, m)$ which is uniquely determined, as the stretching factor of the homotheties $H^m, \bar{H}^m$ and $w\bar{H}^m$ is equal to $\lambda^m > 1$.

2.3. As $\bar{H}$ is a continuous extension of H, it follows that $\bar{H}$ is also α-invariant: For every $w \in F_n$ one has $\bar{H}w = \alpha(w)\bar{H}$ and hence (replacing w by $\alpha^{-1}(w^{-1})$) also $\bar{H}^{-1}w = \alpha^{-1}(w)\bar{H}^{-1}$. Iterated applications of these formulae gives $\bar{H}^m w = \bar{H}^{m-1}\alpha(w)\bar{H} = \bar{H}^{m-2}\alpha^2(w)\bar{H}^2 = \ldots$ and hence

$$\bar{H}^m w = \alpha^m(w)\bar{H}^m : \bar{T} \to \bar{T} \tag{9}$$

for any $m \in \mathbb{Z}$.

Similarily, let $f : \Gamma \to \Gamma$ be as in 1.3 and let $\tilde{f} : \tilde{\Gamma} \to \tilde{\Gamma}$ be the lift of f to the universal covering $\tilde{\Gamma}$ of Γ which represents the same $\alpha \in Aut(F_n)$. Then the equation (1) in 1.3 can be applied iteratively to give

$$\tilde{f}^m w = \alpha(w)^m \tilde{f}^m : \tilde{\Gamma} \to \tilde{\Gamma}\,, \tag{10}$$

but here we have to assume $m \geq 0$ as f^{-1} is not well defined.

2.4. Let $f : \Gamma \to \Gamma$ be as in W.1, and assume that f maps the set of vertices $V(\Gamma)$ to itself. Let $V_0 \subset V(\Gamma)$ be the set of vertices which are fixed by some positive power of f. If we define $m = |V(\Gamma)|!$ then we obtain $f^m(P) = P$ for every $P \in V_0$ and $f^m(P) \in V_0$ for every $P \in V(\Gamma) - V_0$.

Let $\tilde{f}, \bar{H}$ and α be as in 2.3 As the action of F_n on $\tilde{\Gamma}$ by covering translations is free, it follows that for every vertex P in the full preimage $\tilde{V}_0$ of V_0 there is precisely one element $wP \in F_n$ which satisfies

$$w_P \tilde{f}^m(P) = P\,. \tag{11}$$

We use 2.2 to define a map $i : V(\tilde{\Gamma}) \to \bar{T}$, given for all $P \in \tilde{V}_0$ through

$$i(P) = Q(w_P, m)$$

and for all other vertices $P \in V(\tilde{\Gamma}) - \tilde{V}_0$ with $P_0 = \tilde{f}^m(P)$ through

$$i(P) = \bar{H}^{-m} i \tilde{f}^m(P) = \bar{H}^{-m}(Q(w_{P_0}, m)\,.$$

Proposition 2.5. *The above defined map $i : V(\tilde{\Gamma}) \to \bar{T}$ satisfies*
(a) $vi = iv$ for all $v \in F_n$, and
(b) $i\tilde{f} = \bar{H}i$.

Proof. (I) We first consider the case $P \in \tilde{V}_0$, and let $P' = vP$. In order to prove (a) we have to show that $Q(w'_P, m) = vQ(w_P, m)$, for $m \geq 1$ as in 2.4:

From the above equations (10) and (11) one derives $\tilde{f}^m(vP) = \alpha^m(v)\tilde{f}^m(P) = \alpha^m(v)w_P^{-1}P = \alpha^m(v)w_P^{-1}v^{-1}(vP)$. This shows $w_{P'} = vw_P\alpha^m(v^{-1})$. On the other hand, for $Q = Q(w_P, m)$ the definition of $Q(w_P, m)$ gives $w_P\bar{H}^m(Q) = Q$, and thus we derive from (9) just as above $\bar{H}^m(vQ) = \alpha^m(v)\bar{H}^m(Q) = \alpha^m(v)w_P^{-1}Q = \alpha^m(v)w_P^{-1}v^{-1}(vQ)$. This proves $Q(vw_P\alpha^m(v^{-1}), m) = vQ$ and hence $Q(w_{P'}, m) = vQ(w_P, m)$.

In order to prove (b) (for the case $P \in \tilde{V}_0$) we have to show that $Q(w_P, m) = Q(w_{\tilde{f}(P)}, m)$:

From (11) and (10) it follows $\tilde{f}(P) = \tilde{f}(w_P\tilde{f}^m(P)) = \alpha(w_P)\tilde{f}^{m+1}(P) = \alpha(w_P)$ $\tilde{f}^m(\tilde{f}(P))$ and hence $w_{\tilde{f}(P)} = \alpha(w_P)$. But then the definition of $Q(w_P, m)$ together with (9) gives $w_{\tilde{f}(P)}\bar{H}^m(\bar{H}(Q(w_P, m))) = \alpha(w_P)\bar{H}(\bar{H}^m(Q(w_P, m))) = \alpha(w_P)\bar{H}(w_P^{-1}Q(w_P, m)) = \alpha(w_P)\ \alpha(w_P^{-1})\bar{H}(Q(w_P, m)) = \bar{H}(Q(w_P, m))$. This implies $\bar{H}Q(w_P, m) = Q(w_{\tilde{f}(P)}, m)$.

(II) Assume now $P \in V(\tilde{\Gamma}) - \tilde{V}_0$, and let $P_0 = \tilde{f}^m(P) \in \tilde{V}_0$. In order to show (a) we use (in this order) (10), the above derived equation (a) for P_0, (9) and the definition of $i(P)$ to compute $i(vP) = \bar{H}^{-m}i\tilde{f}^m(vP) = \bar{H}^{-m}i\alpha^m(v)\tilde{f}^m(P) = \bar{H}^{-m}i(\alpha^m(v)P_0) = \bar{H}^{-m}\alpha^m(v)i(P_0) = v\bar{H}^{-m}i\tilde{f}^m(P) = vi(P)$.

In order to show (b) for P we use (b) for P_0 to compute $\bar{H}i(P) = \bar{H}\bar{H}^{-m}i\tilde{f}^m(P) = \bar{H}^{-m}\bar{H}i(P_0) = \bar{H}^{-m}i\tilde{f}(P_0) = \bar{H}^{-m}i\tilde{f}^{m+1}(P) = i\tilde{f}(P)$. The last equation, in case $\tilde{f}(P) \in V(\tilde{\Gamma}) - \tilde{V}_0$, follows from the definition of the map i in 2.4, and in case $\tilde{f}(P) \in \tilde{V}_0$ it follows from equation (b) as proved above for vertices in $\tilde{V}_0$. $\square$

We can now extend the map i geodesically to the edges:

Lemma 2.6. *Let $f : \Gamma \to \Gamma$ be a train track representative of $\alpha \in Aut(F_n)$, and let $\bar{T}$ be a (not necessarily minimal) α-invariant $\mathbb{R}$-tree which satisfies BBT. Let $\bar{H} : \bar{T} \to \bar{T}$ be a homothety with stretching factor $\lambda > 1$, and let $i : \tilde{\Gamma} \to \bar{T}$ be a geodesic F_n-equivariant map which satisfies $i\tilde{f}(P) = \bar{H}i(P)$ for all vertices $P \in V(\tilde{\Gamma})$. Then one obtains (after properly reparametrizing f and i along the edges of Γ)*

$$i\tilde{f} = \bar{H}i\,.$$

Proof. We observe:

(i) As i maps by assumption any edge $\tilde{e} \subset \tilde{\Gamma}$ to a geodesic segment, it follows that $\bar{H}i(\tilde{e}) \subset i\tilde{f}(\tilde{e})$.

(ii) As f is a train track map, then no edge e in Γ contains an f^k-backtracking path for any $k \geq 1$. Hence any lift $\tilde{f}^k(\tilde{e})$ of $f^k(e)$ is a geodesic arc in $\tilde{\Gamma}$. As $\bar{T}$ satisfies BBT, it follows from Lemma 1.12 that there is a constant $c \geq 0$ such that $i\tilde{f}^k(\tilde{e})$ is contained in the c-neighborhood of a geodesic arc in $\bar{T}$.

(iii) No edge $\tilde{e} \subset \tilde{\Gamma}$ can possibly contain a non-degenerated $i\tilde{f}$-backtracking path $\tilde{e}_0$, as we can show that $i\tilde{f}(e_0)$ must be a single point in $\bar{T}$: Otherwise the diamater of $i\tilde{f}^k(\tilde{e}_0)$ grows arbitrary large for increasing k, as $\bar{H}$ stretches $\bar{T}$ by $\lambda > 1$ and $\bar{H}^{k-1}i\tilde{f}(\tilde{e}) \subset i\tilde{f}^k(\tilde{e})$, by (i) above. But this would contradict (ii) above. This shows $c = 0$, and actually a little more: The map $i\tilde{f}^k$ maps the points of $\tilde{e}$ in a "non-strictly monotonious" fashion to $i\tilde{f}^k(e)$, i.e. small segments of $\tilde{e}$ may be collapsed to single points, but there is no backtracking along $i\tilde{f}^k(\tilde{e})$.

(iv) Let $\tilde{e}$ be an edge in $\tilde{\Gamma}$ with initial vertex P and terminal vertex Q. From observation (iii) it follows that i maps the path $\tilde{f}(\tilde{e})$ to a geodesic segment which connects $i\tilde{f}(P)$ to $i\tilde{f}(Q)$. By assumption one has $i\tilde{f}(P) = \bar{H}i(P)$ and $i\tilde{f}(Q) = \bar{H}i(Q)$; hence it follows that $i\tilde{f}(\tilde{e})$ and $\bar{H}i(\tilde{e})$ are both equal to the (unique) geodesic segment connecting these two points. Thus it follows that a reparametrization of f and i along edges gives $i\tilde{f}(x) = \bar{H}i(x)$ for all points $x \in \tilde{e}$. □

We can now prove Theorem 1.2, stated in the following slightly more explicite form:

Theorem 1.2'. *If α admits an absolute train track representative $f : \Gamma \to \Gamma$ then every α-invariant F_n-action on some small $\mathbb{R}$-tree T with stretching factor $\lambda > 1$ is Perron-Frobenius, i.e. $T = T_v$ for some Perron-Frobenius eigenvector v of the transition matrix $M(f)$. In particular, if $M(f)$ is irreducible, it follows that T is unique up to uniform rescaling of the metric.*

Proof. Let $f : \Gamma \to \Gamma$ be a train track representative of $\alpha \in Aut(F_n)$, and let $H : T \to T$ commute with α as in 1.1. We also assume that the action on T is minimal and small. As T is α-invariant with stretching factor $\lambda > 1$, this implies by 1.10 that T has trivial arc stabilizers, and hence is the limit of free actions (see [1]). It follows in particular that T satisfies BBT.

We replace T by its metric completion $\bar{T}$ and H by the canonical extension to $\bar{H}$, as in 2.2. We then apply Lemma 2.5 and Proposition 2.6 to obtain a edge geodesic F_n-equivariant map $i : \tilde{\Gamma} \to \bar{T}$. Thus Lemma 2.1 shows that there is an F_n-equivariant isometric embedding $j : T_v \to \bar{T}$ for some Perron-Frobenius tree T_v, with respect to $f : \Gamma \to \Gamma$. But as T is the minimal subtree of $\bar{T}$, it follows that $j(T_v) = T$, which shows that T is Perron-Frobenius. □

3. Partial train tracks with 2-cells

The following notions will be basic for the rest of this paper. Earlier versions of some of them have been used in [6] and [16]. In this section we present the new concepts with a slightly higher degree of generality than needed just for free group automorphisms, to include other naturally occuring cases e.g. from Nielsen-Thurston surface theory.

Partial train track maps

A *graph of spaces* $\mathcal{G}$ is a finite family of path-connected *vertex spaces* X_v, which we sometimes assume for simplicity to be $K(\pi, 1)$-spaces, with a collection $\hat{\Gamma}$ of finitely many edges attached to them ("the *edges of* $\mathcal{G}$"). We do not require that $\mathcal{G}$ is path-connected, but in order to simplify notation we only treat the path-connected case; the extension to the general case is canonical. One obtains from $\hat{\Gamma}$ a graph $\Gamma = \Gamma_{\mathcal{G}}$ by quotienting every vertex space X_v to a single vertex v. The analoguous operation on the universal covering $\tilde{\mathcal{G}}$ of $\mathcal{G}$ gives a tree $\tilde{\Gamma}_{\mathcal{G}}$ (which should not be confused with the universal covering of the graph $\Gamma_{\mathcal{G}}$) with a canonical action of $\pi_1\mathcal{G}$ by "deck transformations". Let $f : \mathcal{G} \to \mathcal{G}$ be a map of pairs, i.e. the union X of all vertex spaces satisfies $f(X) \subset X$, so that f induces maps $f_\Gamma : \Gamma \to \Gamma$ and $f_v : X_v \to X_{f_\Gamma(v)}$ for any vertex $v \in \Gamma$. We assume that f induces an automorphism f_* of $\pi_1\mathcal{G}$, which is the free product of the $\pi_1 X_v$ with $\pi_1\Gamma$. In particular it follows for every *essential* vertex v, i.e. with non-trivial $\pi_1 X_v$, that f_Γ acts periodically on v and that a suitable power of f induces an automorphism of $\pi_1 X_v$. Furthermore every homotopy class of non-trivial closed loops in X has precisely one preimage in X, and every homotopy class rel. boundary of edge paths which cross over precisely 2 edges of $\mathcal{G}$ has only finitely many f-preimages with endpoints on edges of $\mathcal{G}$.

Another consequence of the hypothesis that f_* be an automorphism is that f has bounded backtracking: There is an upper bound to the diameter of the image of any f-backtracking path, i.e. any reduced path in $\mathcal{G}$ which is mapped by f to a contractible loop. Here *length* means the simplicial length of the path, i.e. the number of edges of $\mathcal{G}$ crossed by the path (edges in X, if there are any, do not count !). A path in $\mathcal{G}$ is *reduced* if any "lift" to $\tilde{\Gamma}_{\mathcal{G}}$ is reduced. For any path γ in $\mathcal{G}$ or in $\tilde{\mathcal{G}}$ we denote by $[\gamma]$ the reduced path obtained from γ by contracting all backtracking subpaths to a point or into the subspace X. We will not distinguish between homotopic paths if the homotopy only takes place in the relative part.

A map of pairs $f : \mathcal{G} \to \mathcal{G}$ as above has the *partial train track property (relative to $X \subset \mathcal{G}$)* if any positive power f^t of f maps every f^t-backtracking path contained in an edge of $\mathcal{G}$ into the invariant subspace X. More generally, any path γ in $\mathcal{G}$ is called *legal* if any f^t-backtracking subpath of γ, for any $t \geq 1$, is mapped by f^t into X. Generalizing a notion of Bestvina-Handel [3] we say that two edges e, e' of $\mathcal{G}$, which initiate at the same vertex space X_v, together with a *connecting path* β in X_v, form an *illegal turn* if some positive iterate f^t of f maps β to a connecting path which is a contractible loop in $f^t(X_v)$, and if f^t identifies initial segments of both edges by mapping them to a path of non-zero simplicial length (i.e. $e^{-1}\beta e'$ is reduced, but $f^t(e^{-1}\beta e')$ is not !). The path β is called the *tip* of the illegal turn. The illegal turns group the edges raying out of a vertex in $\tilde{\Gamma}_{\mathcal{G}}$ into equivalence classes called the *gates* of f or of $\mathcal{G}$. If, as assumed above, $\Gamma_{\mathcal{G}}$ is finite and f acts as π_1-isomorphism on the essential vertex spaces, it follows that all gates are finite. Notice that a path γ in $\mathcal{G}$ is legal if and only if it runs over legal turns only, and furthermore that the f-image of any legal path is again legal. We sometimes refer to X as the *relative part* and to $\hat{\Gamma}$ as the *train track part* of f or of $\mathcal{G}$.

Definition 3.1. A map $f : \mathcal{G} \to \mathcal{G}$ is called a *partial train track map (relative to $X \subset \mathcal{G}$)* if $\mathcal{G}$ is a finite (possibly non-connected) graph of spaces, f maps the union X of the vertex spaces to itself, f induces an automorphism of $\pi_1\mathcal{G}$, and f has the partial train track property relative to X. In this case $\mathcal{G}$ (together with the gate structure induced by f on $\tilde{\mathcal{G}}$) is sometimes called a *partial train track*.

A reduced path η in $\mathcal{G}$ is called an *INP* ("indivisible Nielsen path") for f if it consists of precisely two legal subpaths, and if $f(\eta)$ is homotopic relative endpoints to η (compare [3]). INP's for positive powers of f are called *periodic INP's.* Notice that a periodic INP may well have its endpoints in the interior of an edge of $\mathcal{G}$; in this case it is sometimes useful to subdivide the edge, by which we add the subdivision vertex as a new inessential component to the relative part X.

We say that f has *no periodic edges* if no edge of Γ has a periodic orbit with respect to the induced map f_Γ. Similarly, if there are no edges which are contracted by an iterate of f_Γ to a single point, we say that f *no zero edges.*

Lemma 3.2. *If $f : \mathcal{G} \to \mathcal{G}$ is a partial train track map relative to an f-invariant subspace X, and if f has no periodic or zero edges, then for every finite path γ in $\mathcal{G}$ there exits an integer $t \geq 0$ such that $[f^t(\gamma)]$ is "pseudo-legal": It is a legal concatenation of legal paths and of periodic INP's.*

Proof. As the number of illegal turns in $[f^t(\gamma)]$ cannot increase with t, we can assume that for sufficiently large t the number of illegal turns becomes stationary. We can then define for each of these illegal turns an *image turn* as follows: It is the illegal turn in $[f^{t+1}(\gamma)]$ which has its tip at the "inner endpoint" of the backtracking subpath in the non-reduced image $f^{t+1}(\gamma)$ which has at its "outer endpoint" the f-image of the tip of the illegal turn we started out with.

For any two adjacent illegal turns of $[f^t(\gamma)]$ let γ' denote the legal subpath of $[f^t(\gamma)]$ connecting their tips. We consider the points P, P' on γ' which cut off the *eventually cancelled boundary segments*, i.e. the boundary segments of γ' that contain precisely those points which are under iteration of f eventually mapped into the backtracking subpaths at the illegal turns at the two endpoints of γ'. It can occur that $P = P'$, but the two eventually cancelled segments can not overlap, as otherwise the number of illegal turns would strictly decrease, contrary to our assumption.

Let us rename $\gamma' = \gamma'(t)$, $P = P(t)$, $P' = P'(t)$ and, for $t' \geq t$ let $\gamma'(t')$, $P(t')$, $P'(t')$ denote the corresponding legal subpaths and points of $[f^{t'}(\gamma)]$ which lie between the image illegal turns. Then it follows directly from our definitions that $f(P(t)) = P(t+1)$ and $f(P'(t)) = P'(t+1)$.

Claim 1: The simplicial length (i.e. the number of edges or edge segments from $\hat{\Gamma}$) of an eventually cancelled segment is bounded.

This follows directly from the above observed fact that f has bounded backtracking (i.e. the backtracking subpath at the tip of the image of an illegal turn has bounded length), together with the observation that by the finiteness of $\Gamma_{\mathcal{G}}$ there is a $t_0 \geq 0$ with $length(f^{t_0}(e)) \geq 2$ for any edge of $\Gamma_{\mathcal{G}}$, since by assumption f has no periodic or zero edges.

Let now $\gamma''(t)$ be the maximal legal subpath of $[f^t(\gamma)]$ adjacent to $\gamma'(t)$, so that the terminal eventually cancelled boundary segment $\gamma'_{term}(t)$ of $\gamma'(t)$ cancels eventually against the analogously defined initial eventually cancelled boundary segment $\gamma''_{init}(t)$ of $\gamma''(t)$.

Claim 2: In $\gamma'_{term}(t)$ and in $\gamma''_{init}(t)$ any connecting path β (in X) between two adjacent edges (in $\mathcal{G}$) has to belong, up to homotopy relative endpoints, to a given finite family of paths.

Let us first assume that f maps β into the backtracking subpath at the image of the illegal turn between $\gamma'_{term}(t)$ and $\gamma''_{init}(t)$. Then either $f(\beta)$ cancels against the image of some subpath in the interior of an edge which is mapped to the relative part, or else there is a connecting path β' in $\gamma''_{init}(t)$ such that $f(\beta')$ cancels against $f(\beta)$. As $\mathcal{G}$ has only finitely many edges, in the first case there are only finitely many choices for β, as follows from the observation at the beginning of this section, that every edge path of length 2 in $\mathcal{G}$ has only finitely many preimages with endpoints on edges. But the same is true in the second case, as at least one of β or β' must belong to a non-periodic (and hence inessential) component of X, because f induces an automorphism on $\pi_1\mathcal{G}$, so that β and β' cannot belong to the same component X_0 of X: otherwise the subpath of $\gamma'_{term}(\hat{t})\,\gamma''_{init}(\hat{t})$ bounded by β and β' could be closed up in X_0 to give a non-trivial loop with contractible image.

Next let us consider any connecting path β which is not the first connecting path in $\gamma'_{term}(t)$ or the last in $\gamma''_{init}(t)$): For those β the statement of the Claim 1 together with the fact that edges are eventually expanded by f imply that there is an upper bound $\hat{t} \geq 1$, such that β is mapped by $f^{\hat{t}}$ into the backtracking subpath at the illegal turn. Hence Claim 2 follows for any such β from the above considered first case and from the above quoted observation at the beginning of this section.

It remains to prove Claim 2 for the first connecting path of $\gamma'_{term}(t)$ and the last of $\gamma''_{init}(t)$. If those two will eventually cancel against each other (or against another connecting path), then again one of them must belong to a non-periodic vertex space, and Claim 2 follows as above. If they do not cancel against each other or against any other connecting path, they must eventually cancel against the image of a subpath of an edge. But then for one of them this edge is neither the first on $\gamma'_{term}(t)$ nor the last of $\gamma''_{init}(t)$, and hence we deduce (in two steps) from the finiteness of the choices for the other connecting paths considered above, that there is again a bound as the above $\hat{t}$, for the remaining two connecting paths considered now. This proves Claim 2.

The finiteness of the edges of $\mathcal{G}$ together with Claims 1 and 2 shows that there are only finitely many choices for the paths $\gamma'_{term}(t)$ and $\gamma''_{init}(t)$, and hence it follows that some choice will eventually repeat for some iterate $f^{\hat{t}}$. In this case it follows that even the precise position of the initial point $P'(t)$ of $\gamma'_{term}(t)$ (and similarly of the terminal point of $\gamma''_{init}(t)$), if that is contained in the interior of an edge, has to repeat, as $P'(t)$ only depends on the finite repeating edge paths which contain all of the eventually cancelled boundary segments. But this show precisely that $\gamma'_{term}(\hat{t})\,\gamma''_{init}(\hat{t})$ is a periodic INP. □

Remark 3.3. A careful analysis of the last proof shows that there exists a *critical exponent* t^* such that for any path γ with at most one illegal turn the path $f^{t^*}(\gamma)$ is pseudo-legal.

It is not difficult to see that, in the absence of periodic and zero edges, there can be only finitely many periodic INP's (as their length is bounded in terms of the cancellation bound of f and of the minimal exponent t such that f^t honestly expands every edge of $\mathcal{G}$, and as there are only finitely many choices for the connecting paths in any f-backtracking path, see the above proof), and that they can be detected by a relatively easy algorithm.

Let $f : \mathcal{G} \to \mathcal{G}$ be a partial train track map. An infinite reduced path R in $\mathcal{G}$ is called an *eigenray* of f if $f^t(R) = R$ for some $t \geq 1$ (as usually, up to homotopy which only takes place in the relative part). One verifies easily the following:

Facts 3.4. If $f : \mathcal{G} \to \mathcal{G}$ has no periodic or zero edges, then:

(1) An eigenray is always legal.
(2) Every eigenray starts either at a periodic point in the interior of an edge, or at a *periodic gate*, i.e. a gate which consists of edges e for which some $f^t(e)$ starts (up to an initial path in the relative part) at this gate.
(3) Every periodic gate defines precisely one eigenray which starts at that gate. Similarly, every periodic point in the interior of an edge defines in each direction an eigenray.
(4) Every pair of eigenrays which agrees up to a finite initial path of non-zero length defines a periodic INP which is built from these initial paths.
(5) Every periodic INP is obtained as in (4) from two eigenrays.

□

From Facts 3.4 (4) and (5) it follows that there is a natural equivalence among periodic INP's: η and η' are equivalent if and only if maximal legal subpaths γ of η and γ' of η', together with a suitable finite segment of the canonical prolongation of γ and γ' by the corresponding eigenray, form again a periodic INP (which would then be equivalent to both η and to η'). The union of all such suitable prolonged maximal legal subpaths of equivalent periodic INP's will sometimes be called a *generalized periodic INP*.

The most important example of partial train tracks are given by Bestvina-Handel's relative train track representatives for automorphisms of free groups (see [3]). Here the relative part is given by the union of all non-maximal strata.

A partial train track does not change essentially if one replaces the relative part by a homotopy equivalent space and adapts the train track map in the canonical way given by the homotopy equivalence. A particular such case, which plays an important role in the sequel, is described now:

Remark 3.5. Let $f : \mathcal{G} \to \mathcal{G}$ be a partial train track map relative X. If X' is an f-invariant union of some of the connected components of X and if the restriction of f to a selfmap of X' is a homotopy equivalence, then we can construct a new partial train track map in the following way: We fix an integer $t \geq 1$. Then we unglue every endpoint of an edge e in $\mathcal{G}$ which is attached to a point v of X' and reglue it to $f^t(v)$. We redefine the image $f(e)$ for any edge e of $\mathcal{G}$ by replacing any subpath γ' of $f(e)$ contained in X' by $f^t(\gamma')$. Similarly, if f maps a connected component X_v of $X - X'$ into X', we redefine f on X_v to be replaced by the map f^t. It follows directly from these definitions that this gives a well defined map $f_1 : \mathcal{G}_1 \to \mathcal{G}_1$, and it is easy to verify that this is again a partial train track map relative the subspace X_1 which is a copy of X. The maps f and f_1 commute via the map $h : \mathcal{G} \to \mathcal{G}_1$ which "fixes" every edge of $\mathcal{G}$ as well as every point of $X - X'$, and maps every point $x \in X'$ to the point $f^t(x)$ in the subspace X'_1 of the relative part X_1 of $\mathcal{G}_1$ which is the copy of X'. It is easy to show that h is a homotopy equivalence.

Notice that the assumption that f restricts to a homotopy equivalence of X' is necessary in order to get a homotopy equivalence h as above, with $hf = f'h$. If one is content with a more general map h which satisfies this equation, but is only an isomorphism on π_1, then the weaker assumption suffices that f induces a π_1-isomorphism on each connected component of X'. It is, however, unavoidable that X' contains no inessential components of X: otherwise $\pi_1\mathcal{G}_1$ would be different from $\pi_1\mathcal{G}$.

Nielsen faces

We now consider maps $f : \mathcal{G}^2 \to \mathcal{G}^2$, where $\mathcal{G}^2$ differs from the previously considered spaces $\mathcal{G}$ in that, other than a relative part X and a train track part $\hat{\Gamma}$ as before, it also contains finitely many 2-cells, called *Nielsen faces*, which are attached to $\mathcal{G}^1 = X \cup \hat{\Gamma}$ along their boundary. The restriction of f to $\mathcal{G}^1$ is a partial train track map as before. The Nielsen faces, however, must satisfy certain additional conditions:

In case of an *elementary Nielsen face* Δ the boundary of the 2-cell Δ forms a closed loop in $\mathcal{G}^1$ which must contain precisely one periodic INP η, while the *complementary boundary subpath* $\partial\Delta - \overset{\circ}{\eta}$ must be contained in X. To describe the general case we subdivide the above elementary Nielsen face Δ into two *fins* by introducing a *spine* which cuts Δ from the tip of η through the interior of Δ down to some point on the interior of the complementary boundary subpath $\partial\Delta - \mathring{\eta}$. A *general Nielsen face* $\hat{\Delta}$ now consists similarly of a spine with fins attached to it, but in general there may be more than two fins. In this case any two together of this finite collection $\hat{\Delta}$ of fins must form an elementary Nielsen face Δ as described before. In particular, the union of all boundary INP's of a general Nielsen face constitute precisely of a generalized INP as defined above. Below we will often treat only the case of an elementary Nielsen face, if the extension to general Nielsen faces is canonical.

We formally define a path γ in $\mathcal{G}^1$ to be *illegal* if some positive iterate $f^t(\gamma)$ contains an *formally illegal turn*, i.e. a subpath γ_0 which is homotopic relative endpoints in X to the complementary boundary subpath γ_0' of a Nielsen face Δ, such that γ proceeds on both sides of γ_0 by edges of $\mathcal{G}^1$ which lie in the same gate as the initial segments of the periodic INP $\eta = \partial\Delta - \mathring{\gamma}'$. Thus the two gates in $\mathcal{G}^1$ defined by the initial edges of the boundary INP of any Nielsen face Δ will be assembled together into a new, formally enlarged gate in $\mathcal{G}^2$; similarly one proceeds with their f-preimage gates. We say that any path which contains all of the boundary INP of a Nielsen face is non-reduced and we reduce it by homotopying the path over the Nielsen face, thus replacing the INP by the complementary boundary subpath in X. Similarly, any path with a subpath γ_0 that runs once around the whole boundary of a Nielsen face Δ, starting at an arbitrary point $x \in \partial\Delta$, is unreduced and can be reduced by homotopying the subpath γ_0 over the Nielsen face Δ to replace it by the single point x. It follows, just as for partial train tracks without Nielsen faces, that a path is legal in $\mathcal{G}^2$ if and only if all of its positive iterates are reduced. We'd like to point out that in particular the two maximal legal subpaths of the boundary INP of any Nielsen face must be legal in $\mathcal{G}^2$, not just in $\mathcal{G}^1$: they cannot cross over any formally illegal turn as defined above.

We also require that f *permutes the Nielsen faces*, i.e. the image of a Nielsen face is a Nielsen face, up to homotopy in X and up to the fact that the image of a (periodic) INP is not reduced but contains a backtracking subpath at its tip. In particular it follows that f permutes the boundary INP's in precisely the same way as the homotopy classes rel. endpoints of the complementary boundary subpaths. The purpose of Nielsen faces, obviously, is to improve pseudo-legal paths by a reduction in $\mathcal{G}^2$ to paths which are honestly legal.

However, since above we formally enlarged the definiton of an illegal turn, a priori one has to admit the possibility that the backtracking subpath at the tip of the image of a periodic INP η is not backtracking in $\mathcal{G}^1$, for example if the illegal turn at the tip of η belongs to one of the formally enlarged gates as described above. In order to avoid an undesired iteration phenomenon which could lead to infinitely many INP's, we work in this paper with the following definition:

Definition 3.6. A selfmap $f : \mathcal{G}^2 \to \mathcal{G}^2$ of a graph of spaces with finitely many attached 2-cells is a *partial train track map with Nielsen faces* relative to the subspace $X \in \mathcal{G}$, if the following conditions are satisfied:

(1) The restriction $f^1 : \mathcal{G}^1 \to \mathcal{G}^1$ of f to the 1-skeleton $\mathcal{G}^1$ of $\mathcal{G}^2$ is a partial train track map rel. X as in Definition 3.1. The train track property of f^1, however, must be satisfied also with respect to the formally illegal turns defined by the Nielsen faces.
(2) every 2-cell Δ is an elementary Nielsen face (i.e. $\partial\Delta$ is the union of a periodic INP and a path in X, and f permutes the Nielsen faces) or part of a general Nielsen face (i.e. the union of elementary Nielsen faces),
(3) every finite path has an f-iterate which can be reduced in $\mathcal{G}^2$ to a legal path,
(4) there is no non-trivial legal loop outside X which is contractible, and
(5) for every vertex space X_v of X the inclusion into $\mathcal{G}^2$ is π_1-injective.

The map f is called *partial train track map with a collection of some Nielsen faces* if properties (1), (2), (4) and (5) hold.

The most important example (implicitely) in the literature are Thurston's invariant train tracks on surfaces for pseudo-Anosov homeomorphisms: After a slight isotopy one achieves that vertices of such train tracks are mapped to vertices, and then we can subdivide each complementary region, by cutting along arcs which joins the fixed points on adjacent boundary leaves, into triangular shaped Nielsen faces, and into a left-over disk which represents the vertex space for this case. More interesting examples are obtained similarly if the homeomorphism is pseudo-Anosov only on some invariant subsurfaces.

Some of the notions introduced above become more natural if one passes to the 2-complex $\tilde{\Gamma}^2_{\mathcal{G}^2}$ which is the analogue of the tree $\tilde{\Gamma}_{\mathcal{G}}$ in the case of train tracks without Nielsen faces. One obtains $\tilde{\Gamma}^2_{\mathcal{G}^2}$ from the universal covering $\tilde{\mathcal{G}}^2$ by contracting every connected component of the full preimage $\tilde{X}$ of the relative part X to a point. This transforms every Nielsen face into a bigon with legal sides. One may feel tempted to identify equivariantly the two sides of this bigon (properly), but for many examples the resulting quotient is not any more simplicial: In fact, if the corners of the bigons are the only illegal turns in $\mathcal{G}^2$, then the quotient resulting from identifying isometrically with respect to a length function coming from an eigenvector is precisely the corresponding Perron-Frobenius limit tree, see section 4.

In this paper all partial train track maps with Nielsen faces occur from the following special construction:

Definition 3.7. Let $f : \mathcal{G} \to \mathcal{G}$ be a partial train track map, possibly without Nielsen faces or with a collection of some Nielsen faces, and let $\eta = \gamma_1\gamma_2^{-1}$ be an INP with legal subpaths γ_1, γ_2, which is not contained as boundary INP in any of the given Nieslen faces. Then the following operation is called *expansion of an (elementary) Nielsen face at the INP* η: First we check if a boundary point of η is contained in the interior of an edge of $\mathcal{G}$, and in this case we subdivide this edge. Then we add a triangular shaped 2-cell Δ to $\mathcal{G}$ by glueing two of its sides to γ_1 and to γ_2, and define the third side to be a new *auxiliary edge* e of $\mathcal{G}$ which connects the initial point of γ_1 to that of γ_2. We define $f(e) = e$ and $f(\Delta) = \Delta \cup f(\eta)$, and add e to the relative part of $\mathcal{G}$. Similarly, in case of a periodic INP we introduce a periodic orbit of 2-cells and a periodic orbit of corresponding auxiliary edges.

In case of a collection of m eigenrays which agree all up to finite initial subpaths, i.e. the case of a generalized periodic INP, we can *expand a general Nielsen face* by an analogous but more complex expansion process which adds a new *auxiliary* vertex and m new *auxiliary half edges* to the relative part as well as m fins to $\mathcal{G}$ such that any two of the latter form a 2-cell which could be obtained alternatively by the expansion of an elementary Nielsen face as described above.

It is important, however, that before doing such an expansion of a general Nielsen face, we check that *none* of the elementary periodic INP's belonging to this generalized INP is already the boundary INP if a given (elementary or generalized) Nielsen face. If this is the case, then we check which of the eigenrays are already contained in a fin, and only the other ones are expanded by attaching a new fin to $\mathcal{G}$ (along the same spine as the given fins) and a new auxiliary half edge to X (using the given auxiliary vertex as boundary vertex).

We would like to point out that, if $f : \mathcal{G} \to \mathcal{G}$ has no periodic or zero edges, then the expansion of all (periodic or general) Nielsen faces is an algorithmic procedure, based on the above described algorithm which determined all periodic INP's.

Lemma 3.8. *(a) Let $f : \mathcal{G} \to \mathcal{G}$ be a partial train track map without Nielsen faces and with no periodic or zero edges, and let $\mathcal{G}^2$ be obtained from $\mathcal{G}$ by expanding Nielsen faces at every periodic or generalized INP. Then the resulting map $f : \mathcal{G}^2 \to \mathcal{G}^2$ is a partial train track map with Nielsen faces. In particular, every finite edge path γ in $\mathcal{G}^2$ has an iterate $f^t(\gamma)$ which reduces in $\mathcal{G}^2$ to a legal path $[f^t(\gamma)]$.*
(b) Let $f : \mathcal{G} \to \mathcal{G}$ be a partial train track map with no periodic or zero edges, but possibly with a collection of some Nielsen faces which all belong to an f-invariant subspace $\mathcal{G}'$, where we assume that the restriction $f' : \mathcal{G}' \to \mathcal{G}'$ of f satisfies the properties (1) - (5) of Definition 3.6. Let $\mathcal{G}^2$ be obtained from $\mathcal{G}$ by expanding Nielsen faces at every periodic or generalized INP which is not contained in $\mathcal{G}'$. Then again the resulting map $f : \mathcal{G}^2 \to \mathcal{G}^2$ is a partial train track map with Nielsen faces.

Proof. (a) The properties (1), (2) and (5) of Definition 3.6 follow directly from the construction of $\mathcal{G}^2$ or from the analogous properties of the train track map $f : \mathcal{G} \to \mathcal{G}$, since the image of any edge of $\mathcal{G}$ does not cross over any of the newly introduced auxiliary edges (or half edges) in the enlarged relative part X. To show property (3) we consider a maximal subpath γ_i of γ which does not cross over any auxiliary edge. By Lemma 3.2 there is an iterate $f^{t_i}(\gamma_i)$ which can be reduced in $\mathcal{G}$ to a pseudo-legal path: in particular it also does not cross over any of the auxiliary edges. For $t' = max\{t_i\}$ this gives a path $f^{t'}(\gamma)$ which reduces to a path γ' that runs over no illegal turns, except at the tip of some periodic INP which is contained in a pseudo-legal subpath, and except perhaps for the in $\mathcal{G}^2$ formally declared illegal turns at the auxiliary edges. But in the latter case a further iterate allows a reduction in $\mathcal{G}^2$ which replaces the auxiliary edge by an illegal turn in $\mathcal{G}$. Thus one can start again the original iteration procedure, but on a path with strictly less auxiliary edges. This leads eventually to an iterate $f^t(\gamma)$ of γ which reduces in $\mathcal{G}$ to a path γ' with no formally illegal turns, and where every maximal subpath in $\mathcal{G}$ is pseudo-legal. Hence we can reduce γ' in $\mathcal{G}^2$ to a legal path $[f^t(\gamma)]$.

Finally, to show property (4) of Definition 3.6, we use a similar argument to produce an iterate of the given legal path γ where all auxiliary edges can be isotoped over the corresponding Nielsen faces to give a loop which reduces in $\mathcal{G}$ to a pseudo-legal loop γ'. As f eventually expands every edge of $\mathcal{G}$, the length of the legal path γ, if not contained in X, grows unboundedly if one iterates f, while the number of auxiliary edges in $f^t(\gamma)$ stays constant. As the possible cancellation between a periodic INP η and two adjacent long legal paths is bounded, unless the auxiliary edge which was replaced by η would have given rise to a formally illegal turn, it follows that γ' cannot be trivial.

(b) The proof of this part is similar to the above and is left as an exercise to the reader. For the final argument of the proof of (4) one observes that, as $\pi_1\mathcal{G}$ is a free product of the π_1 of the connected components of $\mathcal{G}' \cup X$ and of a free group, the reduced path γ' in $\mathcal{G}$ can be contractible unless it is contained in $\mathcal{G}' \cup X$, since (4) holds by hypothesis on $\mathcal{G}'$. □

We would like to point out here that a reduced path in a train track with Nielsen faces $\mathcal{G}^2$ has much weaker properties than reduced paths in train tracks $\mathcal{G}$ without Nielsen faces: In general their lifts to $\tilde{\mathcal{G}}$ are far from being geodesic or quasigeodesic, and the train track map $f : \mathcal{G}^2 \to \mathcal{G}^2$ will not have bounded backtracking with respect to arbitrary reduced paths. In fact, even for train tracks with Nielsen faces which arise as in Lemma 3.8 it easy to think up examples of reduced paths which can be homotoped over Nielsen faces to give unreduced paths in the 1-skeleton which contain backtracking subpaths in the 1-skeleton of arbitrary length.

On the other hand, geodesics in a partial train track map with Nielsen faces are difficult to control; we restrict ourselves here to the following special case:

Let $v = (v_e)_{e \in Edges(\Gamma_{\mathcal{G}})} \in \mathbb{R}_{\geq 0}^{Edges(\mathcal{G})}$ be a non-negative eigenvector of the transition matrix $M(f)$. This defines as in section 1 a length function $L_v : Edges(\mathcal{G}) \to \mathbb{R}_{\geq 0}$, $e \mapsto v_e$ on $\Gamma_{\mathcal{G}}$ and hence an F_n-equivariant pseudo-metric d_v on the universal covering $\tilde{\mathcal{G}}$ (or on $\tilde{\Gamma}_{\mathcal{G}}$).

Lemma 3.9. *Let $f : \mathcal{G} \to \mathcal{G}$ be a partial train track map with Nielsen faces and no periodic edges. For any non-negative eigenvector v of $M(f)$ with eigenvalue $\lambda > 1$ all legal paths in $\tilde{\mathcal{G}}$ are geodesics with respect to the pseudo-metric d_v.*

Proof. We first observe that every periodic INP has the property that its two maximal legal subpaths have the same v-length, as v is an eigenvector and the two subpaths have f^t-images which agree eventually everywhere, for increasing t. Next we notice that, if γ and γ' are two legal paths which are homotopic relative endpoints, then by property (3) of Definition 3.6 the path $[f^t(\gamma^{-1}\gamma')]$ is legal for sufficiently large t and thus, by property (2) of Definition 3.6, it must be contained in X, so that we can think of the unreduced path $f^t(\gamma^{-1}\gamma')$ as one long backtracking path. Hence the difference between $f^t(\gamma)$ and $f^t(\gamma')$ can only stay in the fact that they run over different maximal legal subpaths of some periodic INP's, so that their length is equal, by the above observation. But f multiplies the length of legal paths by λ, so that γ and γ' must have the same length.

Now consider a possibly non-legal path γ_* in $\mathcal{G}$. If we apply f sufficiently often, we obtain by property (3) of Definition 3.6 a legal path $[f^t(\gamma_*)]$. Now assume that γ is a legal path homotopic relative endpoints to γ_*. Then $f^t(\gamma)$ is again legal, and by the above argument we have $f^t(\gamma)$ and $[f^t(\gamma_*)]$ have the same length. But f^t stretches both γ and γ^* by the factor λ, and by definition of a reduction the length of $[f^t(\gamma_*)]$ is less or equal to that of $f^t(\gamma_*)$. Hence the length of γ^* cannot have been shorter than that of γ, which proves the statement of the lemma. $\square$

For later purposes we need the following sharpening of property (3) of Definition 3.6. The proof is not difficult and hence left to the reader:

Lemma 3.10. *Let $f : \mathcal{G} \to \mathcal{G}$ be a partial train track map with Nielsen faces and no periodic edges, and let τ be a finite tree. Then for any map $j : \tau \to \mathcal{G}$ there exists a $t \geq 0$ and a quotient tree $\tau' = q(\tau)$ such that $f^t j$ is homotopic relative endpoints to a composed map $j'q : \tau \to \tau' \to \mathcal{G}$ with the property that for any two endpoints of τ the reduced path which connects their q-images in τ' is mapped by j' to a legal path.* $\square$

It turns out that often it is convenient if all endpoints of edges in a partial train track map $f : \mathcal{G} \to \mathcal{G}$ are mapped by positive powers of f eventually to a periodic orbit. This is the case if f satisfies the following:

Initial segments condition 3.11. The image of any edge of $\mathcal{G}$ is a path which starts and ends in an edge of $\mathcal{G}$.

Any given partial train track map can be easily modified to a partial train track map which satisfies this condition: One just needs to subdivide all edges which violate that condition and to add to the relative part X of $\mathcal{G}$ the maximal initial and terminal segments which are eventually mapped under iteration of f to X.

We now want to use these notions to show the reader how to transform a given "2-strata train track" into an honest partial train track map with Nielsen faces. Here we explicitely include the (notationally so far suppressed) case that partial train tracks have more than one connected component.

Proposition 3.12. *Let $f : \mathcal{G} \to \mathcal{G}$ be a partial train track (without Nielsen faces), possibly with periodic or zero edges, such that the restriction to the relative part $X = \mathcal{G}'$ is a partial train track map $f' : \mathcal{G}' \to \mathcal{G}'$ with Nielsen faces but without periodic or zero edges, relative to a subspace X'. Then one can modify $\mathcal{G}$ in an algorithmic way to obtain a partial train track map $f_* : \mathcal{G}_* \to \mathcal{G}_*$ relative to X_* with Nielsen faces and without periodic or zero edges. Here X_* is obtained from X' by iteratively adding periodic or zero edges, $\pi_1\mathcal{G}_*$ and $\pi_1\mathcal{G}$ are canonically isomorphic,* *and f_* as well as f induce automorphisms on π_1 which commute via this canonical isomorphism.*

In fact, it follows from the proof given below that, if all connected components of X', of X and of $\mathcal{G}$ are $K(\pi, 1)$-spaces, as will be the case in our applications in the next sections, then $\mathcal{G}_*$ and $\mathcal{G}$ are canonically homotopy equivalent and f and f_1 are homotopic via this homotopy equivalence. Furthermore $\mathcal{G}_*$ is again a $K(\pi, 1)$-space.

Proof. We can assume without loss of generality that the geometric transition matrix $M(f)$ is irreducible; the reducible case follows by iterative application of this special case to f-invariant sub-train-tracks with irreducible transition matrix.

CASE 1: $\mathcal{G} - X$ consists precisely of one f-orbit of periodic edges modulo X.

We distinguish two subcases: If f acts on the given edge orbit periodically not just up to homotopy in X, but also up to homotopy only in X', then we can simply add all edges of that orbit to X', to get an invariant subspace X_*. Now $f_* = f$ and $\mathcal{G}_* = \mathcal{G}$ is the desired partial train track map relative to X_*.

If relative to X' the given edge orbit of $\mathcal{G} - X$ is not periodic, then we apply Remark 3.5 and property (3) of Definition 3.6 to ensure that the image of any edge of $\mathcal{G} - X$ is legal relative to X'. This makes $\mathcal{G}$ into a partial train track relative X' and we can apply Lemma 3.8 (b). This gives a partial train track map $f_* : \mathcal{G}_* \to \mathcal{G}_*$ relative to $X_* = X'$ with the desired properties.

CASE 2: $\mathcal{G} - X$ consists precisely of one zero edge modulo X.

We proceed precisely as in case 1, following the two subcases described there.

CASE 3: $\mathcal{G}$ has no periodic or zero edges relative X.

We first check if f satisfies the initial segments condition (for the train track part of $\mathcal{G}$, i.e. for the edges of $\mathcal{G} - \mathcal{G}'$). If not, we subdivide the edges of $\mathcal{G}$ as explained after 3.11 above to achieve the initial segments condition. This gives rise to new periodic or zero edges (the initial and terminal segments of the original edges of $\mathcal{G} - X$). Hence the transition matrix $M(f)$ may not be any more irreducible, and we may have to use iteratively finitely many times the above two cases before proceeding with this case 3, but now with the hypothesis that the initial segments condition is satisfied:

Let $\mathcal{G}'_0$ be the union of all essential components of $\mathcal{G}'$, so that the restriction $f'_0 : \mathcal{G}'_0 \to \mathcal{G}'_0$ of f' induces a π_1-isomorphism on each connected component (or is actually a homotopy equivalence, as in our applications in the next sections).

Consider now any edge e of $\mathcal{G}$ and any maximal subpath γ of $f(e)$ which is contained in $\mathcal{G}'$. From property (3) of Definition 3.6 we know that there is some $t(\gamma)$-th iterate of γ which reduces to a legal path. Similarly, we know from Lemma 3.10 that for any inessential component $X'_{v'}$ of $\mathcal{G}'$ which is mapped by f to an essential component of $\mathcal{G}'$ there is a quotient $X''_{v'}$ of $X'_{v'}$ as well as a $t(v')$-th iterate which is homotopic to a map where the image of any reduced path in X' is legal. Every inessential component of $\mathcal{G}'$ will be replaced by homeomorphic copies of the corresponding quotient $X''_{v'}$.

Let t be the maximum of all such $t(\gamma)$ and $t(v')$. We now apply Remark 3.5 and obtain a new map $f_* : \mathcal{G}_* \to \mathcal{G}_*$ which commutes with f via a homotopy equivalence (or at least a map which induces an isomorphism on π_1).

We claim that this map is a partial train track map relative to X' rather than just relative to $X = \mathcal{G}'$: We show that all iterates of any edge are legal paths. For the edges of the essential components of $\mathcal{G}'$ this is clear, as (in applying Remark 3.5) we haven't changed the original train track map. For the inessential components of X' it follows directly from the above described modification. Finally, for an edge e of $\mathcal{G}$ we use the initial segments condition to guarantee that in any iterate of e the composition of the above constructed maximal legal subpaths in $\mathcal{G}'$ or their (legal !) iterates with adjacent edges from $\mathcal{G}$ form indeed a legal path.

Thus all hypotheses of Lemma 3.8 (b) are satisfied, and hence this lemma gives us the desired result.

The reader verifies easily that, indeed, the above procedure is algorithmic, and, second, that indeed the expansion of INP's and the corresponding addition of auxiliary edges to the relative part keeps the iterative property of X_* as claimed. $\square$

4. The structure theorem for outer automorphisms of F_n

We now want to show that every automorphism $\alpha \in Aut(F_n)$ is represented by a particular partial train track map with Nielsen faces $f : \mathcal{G} \to \mathcal{G}$. By this we mean that there is an isomorphism $\pi_1\mathcal{G} \tilde{=} F_n$, such that f induces on $\pi_1\mathcal{G}$ an outer automorphism which, via this isomorphism $\pi_1\mathcal{G} \tilde{=} F_n$, is equal to the outer automorphism $\hat{\alpha}$ induced by α.

For any 2-complex Γ^2 and any map $g : \Gamma^2 \to \Gamma^2$ which represents α we say that the map g or the automorphism a is *of polynomial growth* if all irreducible diagonal block submatrices of the geometric transition matrix $M(g)$ are periodic or 0. The existence of such a map g is equivalent (compare [3] or [6]) to the fact that the growth (with respect to any fixed basis of F_n) of any conjugacy class $[\alpha^t(w)]$ in F_n is bounded above by a polynomial in t. In this case, *any* relative train track representative $g : \Gamma \to \Gamma$ of α (as introduced in [3] and derived there algorithmically from α) is of polynomial growth.

Proposition 4.1. *Every automorphism α of a free group can be represented by a partial train track map $f : \mathcal{G} \to \mathcal{G}$ with Nielsen faces and no periodic or zero edges, such the restriction to the vertex groups is of polynomial growth. Moreover, $\mathcal{G}$ and f can be derived algorithmically from a relative train track representative of α.*

Proof. We prove this by induction over the rank n of the free group F_n: Let $g : \Gamma \to \Gamma$ be a relative train track representative of α. As discussed before Remark 3.5, we can reinterprete this as a partial train track map without Nielsen faces where the top stratum of Γ is the train track part and the union of all lower strata is the relative part. (Alternatively, one can use directly Corollary 5.3 of [16] to get a partial train track representative of α). Without loss of generality the top stratum is not a 0-stratum (otherwise we may erase it and start all over with the remainder), so that all essential components of the relative part have free groups of strictly smaller rank than n as fundamental groups. By inductive hypothesis the restricted automorphisms can be represented by a partial train track map $f' : \mathcal{G}' \to \mathcal{G}'$ which satisfies the claim in question. We can replace the union of the lower strata by these partial train tracks and contract the inessential components to single points. This defines a 2-strata partial train track $f : \mathcal{G} \to \mathcal{G}$ relative $X = \mathcal{G}'$ as in Proposition 3.12 which represents the same automorphism α, and our claim is given directly by that proposition. □

As a consequence we can extend the facts proved in [6] and in the previous sections from absolute or relative train tracks to partial train tracks with Nielsen faces. In fact, the construction of a projectively α-invariant $\mathbb{R}$-tree T from a train track representative of α is done in [6, section 2] for δ-hyperbolic spaces. One only has to prove that the d_∞-pseudo-metric defined on $\tilde{\mathcal{G}}$ is indeed given by the Perron-Frobenius eigenvector v used there. This follows directly from the fact that legal paths are geodesics with respect to the metric defined by the length function $L_v(\,\cdot\,)$, as shown in the above Lemma 3.9.

For the converse direction (Theorem 1.2') we first need to show:

Proposition 4.2. *Let T be an $\mathbb{R}$-tree with very small minimal isometric action of F_n, and let $\mathcal{G}$ be a pseudo-metric 2-complex with $\pi_1 \mathcal{G} = F_n$. Let $i : \tilde{\mathcal{G}} \to T$ be an equivariant map which maps every subset of $\tilde{\mathcal{G}}$ that has diameter 0 (with respect to the lift of the pseudo-metric) to a single point of T. Then this map i satisfies BBT: There is a constant $BBT(i) > 0$ such for that every geodesic segment $[x, y] \subset \tilde{\mathcal{G}}$ the image $i([x, y])$ is contained in the $BBT(i)$-neighborhood of the geodesic segment $[i(x), i(y)] \subset T$.*

Proof. Let d be the pseudo-metric on $\tilde{\mathcal{G}}$, and let d' be any metric obtained from d by equivariantly "blowing up" the maximal subspaces of diameter 0. Notice that $d(x, y) \leq d'(x, y)$ for all $x, y \in \tilde{\mathcal{G}}$.

Notice also that by assumption there exists a constant $K > 0$ such that $dist(i(x), i(y)) \leq K d(x, y)$ for all $x, y \in \tilde{\mathcal{G}}$.

We use the fact that the fundamental group of $\mathcal{G}$ is free, and hence there is a homotopy equivalence $h : \mathcal{G} \to \Gamma$ onto some connected finite metric graph Γ. Notice that its F_n-equivariant lift $\tilde{h} : \tilde{\mathcal{G}} \to \tilde{\Gamma}$ is a quasi-isometry, say with quasi-isometry constants $\lambda > 0$ and $\epsilon \geq 0$, with respect to d' and to the metric on $\tilde{\Gamma}$ lifted from Γ. We choose an F_n-equivariant map $j : \tilde{\Gamma} \to T$ and observe that there is a constant $C > 0$ such that $i(x)$ and $j\tilde{h}(x)$ are at distance at most C for all $x \in \tilde{\mathcal{G}}$.

Now let $\gamma : [0, 1] \to \tilde{\mathcal{G}}$ be a d-geodesic i-backtracking path with endpoints $x, y \in \tilde{\mathcal{G}}$, and consider the image path $\tilde{h}\gamma$ in $\tilde{\Gamma}$ as well as its reduction relative endpoints $[\tilde{h}\gamma]$ which is a geodesic as $\tilde{\Gamma}$ is a tree. As T satisfies BBT, there is a bounded backtracking constant $BBT(j) > 0$, such that $j([\tilde{h}\gamma])$ is contained in the $BBT(j)$-neighborhood of the geodesic segment $[j\tilde{h}(x), j\tilde{h}(y)]$ and hence in the $BBT(j) + C$-neighborhood of the point $i(x) = i(y)$. Thus any point of γ which is mapped by i outside the $BBT(j) + 2C$-neighborhood N of $i(x) = i(y)$ must be contained in a maximal $\tilde{h}$-backtracking subpath γ_0 of γ. The latter has its endpoints mapped by $\tilde{h}$ to $[\tilde{h}\gamma]$ and hence by i into $N \subset T$. But the d'-length of γ_0 is bounded by ϵ, and so is its d-length, which implies that $i(\gamma_0)$ is contained in the $\frac{1}{2}K\epsilon$-neighborhood of its common endpoint in T. This shows that $BBT(j) + 2C + \frac{1}{2}K\epsilon$ is a backtracking bound for i. □

We can now extend the proof given in section 2 from absolute train tracks to train track maps with Nielsen faces $f : \mathcal{G} \to \mathcal{G}$ as considered in Proposition 4.1:

We first define a map i on every connected component $\tilde{X}_v$ of the full preimage $\tilde{X}$ in $\tilde{\mathcal{G}}$ of the relative part $X \subset \mathcal{G}$ in precisely the same way as explained before Proposition 2.5, which proof extends directly. As the automorphism grows polynomially on X, while the given homothety $H : T \to T$ with stretching factor $\lambda > 1$ reflects exponential growth, any map i which satisfies $i\tilde{f} = Hi$ must map all of any component $\tilde{X}_0$ to a single point.

We then proceed, as in Lemma 2.6, by extending i geodesically to the edges of the train track part of $\mathcal{G}$. However, as $\tilde{\mathcal{G}}$ is not necessarily a tree and reduced paths may not be geodesic, we have to use a little care in extending the proof of Lemma 2.6:

We proceed iteratively by using the strata structure of $\mathcal{G}$ from the proof of Proposition 4.1 by first considering the lowest stratum, which has already the train track property before adding any 2-cells. Hence the proof of Lemma 2.6 applies to this stratum.

We then define a length on the edges of this stratum by pulling back the i-distance of their endpoints as in the beginning of the proof of Proposition 2.1 and argue as there to show that this length function determines an eigenvector of the corresponding submatrix of the transition matrix $M(f)$. This enables us to use Lemma 3.9 to argue that, after expanding all Nielsen faces of this stratum, legal paths are geodesics.

We can now pass to the next stratum, again first without 2-cells. As we know that the f^t-images of any edge of this stratum are legal, for all $t \geq 1$, the previous argument ensures that they are also geodesic, so that again the proof of Lemma 2.6 applies: The latter is based in the property BBT (for geodesics !), which is shown in Proposition 4.2 above, in the form needed here.

The same arguments apply iteratively to all higher strata, which completes the proof of the analogue of Lemma 2.6. The final step, the proof of Proposition 2.1, extends canonically.

This shows Theorem I of the Introduction, restated here in the following more explicit form:

Theorem 4.3. *Let $\alpha \in Aut(F_n)$, and let $f : \mathcal{G} \to \mathcal{G}$ be a partial train track representative of α as in Proposition 4.1 (i.e. with Nielsen faces and no periodic or zero edges, relative to a subspace on which f has polynomial growth).*

(a) Every non-negative eigenvector v with eigenvalue $\lambda > 1$ of the geometric transition matrix $M(f)$ defines a non-trivial minimal $\mathbb{R}$-tree T_v with small isometric F_n-action and a homothety $H_\alpha : T_v \to T_v$ with stretching factor λ which satisfies $\alpha(w)H_\alpha = H_\alpha w : T_v \to T_v$ for all $w \in F_n$. Furthermore, the F_n-action on T_v has trivial arc stabilizers.

(b) Conversely, given a non-trivial minimal $\mathbb{R}$-tree T with small isometric F_n-action and a homothety $H : T \to T$ with stretching factor $\lambda > 1$ which satisfies $\alpha(w)H = Hw : T \to T$ for all $w \in F_n$, there exists a non-negative eigenvector v of $M(g)$ with eigenvalue λ and an F_n-equivariant isometry $j : T_v \to T$ which satisfies $Hj = jH_\alpha$.

(c) For every eigenvector v there is an F_n-equivariant map $i : \tilde{\mathcal{G}} \to T_v$ such that $H_\alpha i = i\tilde{f}$, where $\tilde{f}$ is the lift of f to $\tilde{\mathcal{G}}$ which satisfies $\alpha(w)\tilde{f} = \tilde{f}w : \tilde{\mathcal{G}} \to \tilde{\mathcal{G}}$ for all $w \in F_n$. Here i maps every edge e and more generally any legal path in $\tilde{\mathcal{G}}$ isometrically (with respect to the d_v-metric, given by the length function L_v defined on $Edges(\mathcal{G})$) to a geodesic segment in T_v. □

The set of non-negative eigenvectors v with eigenvalues strictly bigger than 1, for the non-negative matrix $M(f)$, is a subset of the non-negative cone of $\mathbb{R}^{Edges(\mathcal{G})}$. Hence it carries a linear structure and has (finitely many) extremal points. It is possible to show that these properties are intrinsic properties of the set of α-invariant $\mathbb{R}$-trees and thus independent of the particular train track map used. Notice the fact (see [7]) that the stabilizer of a branchpoint in T is always of smaller rank than n. Such a branchpoint is given as the i-image of a connected component of full preimage of the *0-part* of v, i.e. the union of the relative part of $\mathcal{G}$ together with all edges e with $v_e = 0$. This gives the possibility for an effective structural analysis of α by decomposing F_n into a family of iteratively defined canonical (up to finitely many choices) α-invariant subgroups of strictly smaller rank, where the α-action on each "stratum" is reflected by an α-invariant $\mathbb{R}$-tree action with $\lambda > 1$, except for the "bottom stratum", where α has polynomial growth. For example, if α is *geometric*, i.e. induced by a homeomorphism h of a surface with boundary, then the Nielsen-Thurston classification of h is a special case of this structural analysis.

However, for the purpose of this paper we do not need all of this structural analysis, and hence we will here not persue this further. In Part II we only need the following:

Corollary 4.4. *Let $f : \mathcal{G} \to \mathcal{G}$ and $f' : \mathcal{G}' \to \mathcal{G}'$ be two partial train track maps with Nielsen faces and no periodic edges, both relative to subspaces on which f and f' are polynomially growing. Assume that both, f and f' represent the same automorphism $\alpha \in Aut(F_n)$: there is a homotopy equivalence $h : \mathcal{G} \to \mathcal{G}'$ which commutes up to homotopy with f, f'. Then to every non-negative eigenvector v of $M(f)$ there is canonically associated a non-negative eigenvector v' of $M(f')$, and the homotopy equivalence h restricts to a homotopy equivalence between the essential components of the 0-parts on $\mathcal{G}$ and $\mathcal{G}'$ defined by v and v'. This restriction of h commutes up to homotopy with the corresponding restrictions of f and of f'.*

□

In the 1:1 correspondence between eigenvectors from this corollary clearly the partial order of $\mathbb{R}$-tree actions is preserved, which is defined by the inclusion on the associated set of conjugacy classes of elements $w \in F_n$ that have translation length 0. It is easy to see that the maximal $\mathbb{R}$-trees T_v with respect to this partial order are precisely the ones which have exactly one non-periodic irreducible diagonal block in the transition matrix, modulo the 0-part of v. In particular, up to scalar multiples such an eigenvector v is uniquely determined by its 0-part among all non-negative eigenvectors of $M(f)$ with eigenvalue strictly bigger than 1. In analogy to the situation of a surface homeomorphism which is on a subsurface a pseudo-Anosov self-map, we will call such an action on T_v a *relative pseudo-Anosov* action and T_v a *relative pseudo-Anosov* $\mathbb{R}$-tree. For any such eigenvector v of the transition matrix $M(f)$, for f as in Corollary 4.4, we define a new partial train track map, called a *relative pseudo-Anosov train track map*, by adding all edges e of $\mathcal{G}$ in the 0-part of v to the relative part X. Hence, if $f : \mathcal{G} \to \mathcal{G}$ is relative pseudo-Anosov, then there is (up to rescaling) precisely one eigenvector with eigenvalue $\lambda > 1$, and this Perron-Frobenius eigenvector is strictly positive on all edges of $\mathcal{G}$.

As by Theorem 4.3 there are only finitely many relative pseudo-Anosov $\mathbb{R}$-tree actions (up to rescaling), for any $\alpha \in Aut(F_n)$, we have proved:

Theorem 4.5. *Every automorphism α of F_n either is polynomially growing, or else there exists a non-empty finite family of relative pseudo-Anosov $\mathbb{R}$-trees T_i, and one can algorithmically derive from α for each of the T_i a relative pseudo-Anosov train track map $f_i : \mathcal{G}_i \to \mathcal{G}_i$ which represents α and satisfies $T_i = T_{v_i}$ for the (up to scalar multiplication uniquely determined) eigenvector v_i of $M(f_i)$ with $\lambda > 1$. If $f' : \mathcal{G}' \to \mathcal{G}'$ is another relative pseudo-Anosov representative of α with relative pseudo-Anosov $\mathbb{R}$-tree T', and if there is an equivariant isometry $T_i \to T'$, then there is a homotopy equivalence $h : \mathcal{G}_i \to \mathcal{G}'$ which is a map of pairs and commutes with f, f' up to homotopy of pairs.* □

This theorem is the essential tool which we use in Part II to define an inductive procedure which determines the conjugacy class of α. It is noteworthy that without a further normalization process (see [15]) the properties of the transition matrix of such a pseudo-Anosov $\mathbb{R}$-tree action are not very well determined by the $\mathbb{R}$-tree: For example, it may well occur that two relative pseudo-Anosov train track maps f and f' determine the same $\mathbb{R}$-tree, but the matrix $M(f)$ is irreducible, while $M(f')$ is reducible.

We finish this section with a lemma on relative pseudo-Anosov train track maps which is needed in the next section. First we need to recall some notation from [6]:

Consider a point Q in an $\mathbb{R}$-tree T with $wH^t(Q) = Q$ for some $t \geq 1$, $w \in F_n$, for some homothety H with stretching factor $\lambda > 1$. In [6] it was pointed out that every direction at Q contains precisely one *eigenray*: This is an infinite geodesic path ρ in T which starts at Q and is wH^t-invariant. It is easy to see that F_n-translates and the H-images of eigenrays are again eigenrays in T. The proof of the following is a direct consequence of our discussion of eigenrays in $\mathcal{G}$ in the last section, and of the fact that $M(f)$ has only one diagonal block which is non-periodic, and none which is 0.

Lemma 4.6. *For any eigenray ρ in a pseudo-Anosov $\mathbb{R}$-tree T there is a legal path $\tilde{R}$ in $\tilde{\mathcal{G}}$, the lift of an eigenray R in $\mathcal{G}$, which is mapped by i isometrically onto ρ. Any two such paths $\tilde{R}$ and $\tilde{R}'$ agree up to finite initial segments $\tilde{R}_0 \subset \tilde{R}$ and $\tilde{R}'_0 \subset \tilde{R}'$, and if these have v-length > 0 then their images in $\mathcal{G}$ form an INP. All INP's in $\mathcal{G}$ come from an eigenray ρ in T in such a way. Every edge of $\mathcal{G}$ is contained in the image R of any such $\tilde{R} \subset \tilde{\mathcal{G}}$.* □

References

1. M.Bestvina, M.Feighn, *Outer limits*, preprint.
2. M.Bestvina, M.Feighn, M.Handel, *The Tits Alternative for $Out(F_n)$ I, II*, preprints.
3. M. Bestvina, M. Handel, *Train tracks for automorphisms of the free group*, Annals of Math. **135** (1992), 1–51.
4. M.M.Cohen, M.Lustig, *The conjugacy problem for Dehn twist automorphisms of free groups*, Comment. Math. Helv., to appear.
5. W. Dicks, E. Ventura, *The group fixed by a family of injective endomorphisms of a free group*, Contemporary Mathematics **195**, Providence, RI: AMS, (1996).
6. D. Gaboriau, A. Jaeger, G. Levitt, M. Lustig, *An index for counting fixed points of automorphisms of free groups*, Duke Math. Jour. **93** (1998), 425–452.
7. D. Gaboriau, G. Levitt, *The rank of actions on* $\mathbf{R}$*-trees*, Ann. Sci. Ec. Norm. Super., IV. Ser. **28 (5)** (1995), 549–570.
8. S. Kristic, *A uniqueness decomposition theorem for actions of finite groups on free groups*, J. Pure Appl. Algebra **61** (1989), 29–48.
9. S.Krstic, M.Lustig, Vogtmann, *An equivariant Whitehead algorithm and conjugacy for roots of Dehn twists*, Proc. Edinburgh Math. Soc., to appear.
10. J.Los, *On the conjugacy problem for automorphisms of free groups. With an addendum.*, Toplogy **53** (1996), 779–808.
11. M. Lustig, *Automorphismen von freien Gruppen*, Habilitation thesis (1992).
12. M. Lustig, *Automorphisms, train tracks and non-simplicial* $\mathbf{R}$*-tree actions*, unpublished preprint 1993.
13. M. Lustig, *Prime decomposition and conjugacy problem for $Out(F_n)$*, unpublished survey preprint 1994.
14. M. Lustig, *Structure and conjugacy for automorphisms of free groups*, http://homepage.ruhr-uni-bochum.de/Martin.Lustig, April 1999.
15. M. Lustig, private notes.
16. M. Lustig, U. Oertel, *Invariant Laminations for Diffeomorphisms of Handlebodies*, Quart. J. Math. Oxford, to appear.
17. Z.Sela, *The isomorphism problem for hyperbolic groups. I.*, Ann. Math., II. Ser. **141 (2)** (1995), 217–283.
18. Z.Sela, *The Nielsen-Thurston classification and automorphisms of a free group. I.,II.*, preprints.
19. R.Skora, unpublished preprint.

Laboratoire de Mathématiques, Faculte des Sciences (St. Jérome), Avenue Escadrille Normandie-Niemen, Université d'Aix-Marseille III, 13397 Marseille Cedex 20, France

E-mail address: `Martin.Lustig@math.u-3mrs.fr, lustig@mpim-bonn.mpg.de`