Dynamics on free-by-cyclic groups

SPENCER DOWDALL
ILYA KAPOVICH
CHRISTOPHER J LEININGER

Given a free-by-cyclic group $G = F_N \rtimes_\varphi \mathbb{Z}$ determined by any outer automorphism $\varphi \in \mathrm{Out}(F_N)$ which is represented by an expanding irreducible train-track map f, we construct a $K(G, 1)$ 2–complex X called the *folded mapping torus* of f, and equip it with a semiflow. We show that X enjoys many similar properties to those proven by Thurston and Fried for the mapping torus of a pseudo-Anosov homeomorphism. In particular, we construct an open, convex cone $\mathcal{A} \subset H^1(X; \mathbb{R}) = \mathrm{Hom}(G; \mathbb{R})$ containing the homomorphism $u_0 \colon G \to \mathbb{Z}$ having $\ker(u_0) = F_N$, a homology class $\epsilon \in H_1(X; \mathbb{R})$, and a continuous, convex, homogeneous of degree -1 function $\mathfrak{H} \colon \mathcal{A} \to \mathbb{R}$ with the following properties. Given any primitive integral class $u \in \mathcal{A}$ there is a graph $\Theta_u \subset X$ such that:

(1)   The inclusion $\Theta_u \to X$ is π_1–injective and $\pi_1(\Theta_u) = \ker(u)$.
(2) $u(\epsilon) = \chi(\Theta_u)$.
(3) $\Theta_u \subset X$ is a section of the semiflow and the first return map to Θ_u is an expanding irreducible train track map representing $\varphi_u \in \mathrm{Out}(\ker(u))$ such that $G = \ker(u) \rtimes_{\varphi_u} \mathbb{Z}$.
(4)   The logarithm of the stretch factor of φ_u is precisely $\mathfrak{H}(u)$.
(5)   If φ was further assumed to be hyperbolic and fully irreducible then for every primitive integral $u \in \mathcal{A}$ the automorphism φ_u of $\ker(u)$ is also hyperbolic and fully irreducible.

1   Introduction

Given a group K and $\varphi \in \mathrm{Out}(K)$ represented by the automorphism $\Phi \colon K \to K$, we denote the *mapping torus group* of φ as the semidirect product

$$G = G_\varphi = K \rtimes_\varphi \mathbb{Z} = \{w, t \mid t^{-1}wt = \Phi(w) \text{ for all } w \in K\}.$$

The group G and the *associated homomorphism* $u_0 \colon G \to \mathbb{Z}$ given by $u_0(t^n w) = n$ depend only on φ, and not on Φ. Conversely, given a *primitive integral* $u \in \mathrm{Hom}(G, \mathbb{R}) = H^1(G; \mathbb{R})$ (a homomorphism with $u(G) = \mathbb{Z}$), there is an *associated*

splitting $G = \ker(u) \rtimes_{\varphi_u} \mathbb{Z}$ with $\varphi_u \in \mathrm{Out}(\ker(u))$ defining the splitting and with associated homomorphism u. Namely, if $t_u \in G$ is such that $u(t_u) = 1 \in \mathbb{Z}$, then $\varphi_u \in \mathrm{Out}(\ker(u))$ is the outer automorphism class of the automorphism Φ_u of $\ker(u)$ defined as $\Phi_u(g) = t_u^{-1} g t_u$, where $g \in \ker(u)$. Of particular interest is the case when $\ker(u)$ is finitely generated. According to a result of Neumann [49], as well as more general results of Bieri, Neumann and Strebel [10] and Levitt [43], for finitely generated G there exists an open cone $\mathscr{C}_G \subseteq H^1(G; \mathbb{R})$ (a subset invariant by positive scalar multiplication), which we call the *fg cone of G*, such that a primitive integral element $u \in H^1(G; \mathbb{R})$ has finitely generated kernel if and only if $u \in \mathscr{C}_G$.

When $G = \pi_1(M)$, where M is a closed oriented 3–manifold, Stallings [58] proved that every primitive integral element $u \in \mathscr{C}_G$ is induced by a fibration $M \to \mathbb{S}^1$. The fiber $S_u \subset M$ is then a surface with $\pi_1(S_u) = \ker(u)$, and φ_u is induced by a homeomorphism $F_u \colon S_u \to S_u$ called the *monodromy* for u. Suppose that for some $u_0 \in \mathscr{C}_G$, the monodromy F_{u_0} is *pseudo-Anosov* and let $\mathscr{C}_0 \subset \mathscr{C}_G$ denote the component containing u_0. Through a detailed analysis of M, Thurston [60] (see also Fathi, Laudenbach, and Poénaru [26]) and Fried [29; 30] proved that the monodromies for all primitive integral elements in $\mathscr{C}_0$ are intimately related to one another in important topological, geometric and dynamical ways. For example, the assignments $u \mapsto \chi(S_u)$ and $u \mapsto h(F_u)$ for primitive integral $u \in \mathscr{C}_0$, where $h(F_u)$ denotes the topological entropy of F_u, both extend to convex functions on all of $\mathscr{C}_0$.

In this paper, we provide a model for, and a similar analysis of, the situation where the 3–manifold group is replaced by a *free-by-cyclic group*, $G = F_N \rtimes_\varphi \mathbb{Z}$, where F_N is a free group of rank $2 \le N < \infty$ and $\varphi \in \mathrm{Out}(F_N)$. Free-by-cyclic groups share many features with fundamental groups of compact 3–manifolds fibering over the circle (for example, see Algom-Kfir and Rafi [4] Feighn and Handel [27] Kapovich and Lustig [40]), although the behavior in the free-by-cyclic case is generally more varied and more complicated.

Let $\varphi \in \mathrm{Out}(F_N)$ be an arbitrary element and let $f \colon \Gamma \to \Gamma$ be a *tame graph-map* which is a *topological representative* of φ; see Definitions 2.1, 2.3, 2.6, and 2.16 (such representatives always exist and are easy to build; see Remark 3.2). Let $G = G_\varphi$ be the mapping torus group of φ and let $u_0 \in H^1(G_\varphi; \mathbb{R})$ be the associated homomorphism. From f (and a folding sequence for f; see Section 3) we construct $(X, \psi, \epsilon, \mathcal{A})$, where:

- X is a $K(G, 1)$ polyhedral 2–complex.
- ψ is a semiflow on X, that is, an action of the additive semigroup $\mathbb{R}_+$.
- $\epsilon \in Z_1(X; \mathbb{R})$ is a cellular cycle.
- $\mathcal{A} = \mathcal{A}_X \subset H^1(G; \mathbb{R})$ is an open, convex cone containing u_0.

A *section* of ψ is a subset $\Theta \subset X$ such that every flow line meets Θ infinitely often at a discrete set of times. Consequently any section has a *first return map* $\Theta \to \Theta$. With this terminology and notation, our results can be summarized as follows.

Theorem A Let $N \geq 2$, let $\varphi \in \mathrm{Out}(F_N)$ be an arbitrary element represented by a tame graph-map $f \colon \Gamma \to \Gamma$ as above, and let $G = F_N \rtimes_\varphi \mathbb{Z}$. Then there exists $(X, \psi, \epsilon, \mathcal{A})$ as above such that for any primitive integral $u \in \mathcal{A}$, there exists a compact connected topological graph $\Theta_u \subset X$ with the following properties:

(1) The inclusion $\Theta_u \subset X$ is π_1–injective and, with the identification $\pi_1(X) = G$ we have $\pi_1(\Theta_u) = \ker(u)$.

(2) $u(\epsilon) = \chi(\Theta_u) = 1 - \mathrm{rk}(\ker(u))$.

(3) Θ_u is a section of the semiflow ψ and the first return map $f_u \colon \Theta_u \to \Theta_u$ represents $\varphi_u \in \mathrm{Out}(\ker(u)) = \mathrm{Out}(\pi_1(\Theta_u))$.

If we further assume that $f \colon \Gamma \to \Gamma$ is an expanding irreducible train-track map then there exists a continuous, convex, homogeneous of degree -1 function $\mathfrak{H} \colon \mathcal{A} \to (0, \infty)$ such that for every primitive integral $u \in \mathcal{A}$ we have:

(4) $f_u \colon \Theta_u \to \Theta_u$ is an expanding irreducible train-track map.

(5) $\mathfrak{H}(u) = \log(\lambda(\varphi_u)) = h(f_u)$ where $\lambda(\varphi_u) > 1$ is the "growth rate" or "stretch factor" of φ_u and $h(f_u)$ is the topological entropy of f_u.

(6) If in addition we assume that $\varphi \in \mathrm{Out}(F_N)$ is hyperbolic and fully irreducible then for every primitive integral $u \in \mathcal{A}$ the automorphism φ_u of $\ker(u)$ is hyperbolic and fully irreducible.

In the remainder of the introduction, we will discuss in more detail some generalizations and refinements of the various parts of this theorem (see Sections 1.1–1.3 and Theorems B, C, and D below, which imply Theorem A), elaborating on the similarities and differences with the 3–manifold setting. Next we will describe in Section 1.4 an application and the connection to work of Algom-Kfir and Rafi [4]. In Section 1.5 we compare our construction with related constructions of Gautero [32] and Wang [63]. Finally, in Section 1.6 we briefly explain the relation between the class ϵ and the work of McMullen [46], Dunfield [23], and Button [15] on the Alexander norm, further strengthening the analogy with the 3–manifold setting.

Remark 1.1 In our follow-up paper [19], our definition of a section of ψ is more restrictive than what is given above. It turns out that all sections we construct in this paper do satisfy the conditions of that definition, however.

1.1 Fibrations, semiflows and an Euler-like class.

In what follows, we assume that M is a closed, oriented, aspherical 3–manifold and write $G_M = \pi_1(M)$. We also assume that M fibers over the circle $\eta_0 \colon M \to \mathbb{S}^1$ with fiber S and monodromy $F \colon S \to S$, so that $M = S \times [0, 1]/(x, 1) \sim (F(x), 0)$. Write $u_0 \in H^1(M; \mathbb{R})$ to denote the induced primitive integral element $u_0 = (\eta_0)_* \colon G_M \to \mathbb{Z}$.

We sketch here a geometric proof of the existence of the open set $\mathscr{C}_{G_M}$, essentially due to Tischler [62], that motivates our construction. The starting point of the analysis is the following two observations. Firstly, there is a natural suspension flow on M determined by the local flow on $S \times [0, 1]$, given by $(x, t) \mapsto (x, t + s)$ for $x \in S$ and $s \leq 1 - t$. Secondly, via the de Rham isomorphism $H^1(M; \mathbb{R}) \cong H^1_{\mathrm{dR}}(M; \mathbb{R})$, one can represent u_0 by the closed, nowhere vanishing 1–form $v_0 = \eta_0^*(\mathrm{vol}_{\mathbb{S}^1})$, where $\mathrm{vol}_{\mathbb{S}^1}$ is the unit volume form on $\mathbb{S}^1$. Any class $u \in H^1(M; \mathbb{R})$ *projectively* close to u_0 (that is, close up to scaling by positive real number), is also represented by a closed nowhere vanishing 1–form v projectively close to v_0.

This 1–form v provides a crucial tool for studying properties of u: If u is a primitive integral class represented by such a v, then integrating v defines a fibration $\eta_v \colon M \to \mathbb{R}/\mathbb{Z} = \mathbb{S}^1$ so that $v = \eta_v^*(\mathrm{vol}_{\mathbb{S}^1})$ and $u = (\eta_v)_* \colon G_M \to \mathbb{Z}$. Thus every primitive integral $u \in H^1(M; \mathbb{R})$ projectively close to u_0 is also induced by a fibration over $\mathbb{S}^1$, and hence has finitely generated kernel equal to the fundamental group $\pi_1(S_v)$ of the fiber S_v. Furthermore, the fiber S_v is a section of the suspension flow on M, and the first return map to S_v is precisely the monodromy F_v of the fibration. The Euler characteristic $\chi(S_v)$ of any fiber S_v may then be calculated by pairing the homology class of S_v with the Euler class of the 2–plane bundle $\ker(v)$; equivalently, one may pair u, the Poincaré dual of S_v, with the Poincaré dual e_v of this Euler class. On the other hand, for any v projectively close to v_0, $\ker(v)$ and $\ker(v_0)$ are homotopic 2–plane subbundles of TM and so have equivalent Euler classes. Therefore, there is a single class $e = e_{v_0}$ such that $\chi(S_v) = u(e)$ for all u projectively close to u_0; see Thurston [60].

Now consider a free-by-cyclic group $G = G_\varphi = F_N \rtimes_\varphi \mathbb{Z}$ defined by $\varphi \in \mathrm{Out}(F_N)$ with $2 \leq N < \infty$ and associated primitive integral $u_0 \colon G \to \mathbb{Z}$ with $\ker(u_0) = F_N$. We would like to obtain a similar geometric understanding of those $u \in H^1(G; \mathbb{R})$ which are projectively near to u_0. Taking a topological representative $f \colon \Gamma \to \Gamma$ of φ (that is, a graph Γ together with a topological graph map f and an isomorphism $\mathfrak{m} \colon F_n \to \pi_1(\Gamma)$ for which $\mathfrak{m}^{-1} f_* \mathfrak{m}$ represents $\varphi \in \mathrm{Out}(F_N)$; see Section 2.8), one may build the corresponding mapping torus $Z_f = \Gamma \times [0, 1]/(x, 1) \sim (f(x), 0)$. Analogous to the 3–manifold above, Z_f has fundamental group isomorphic to G and admits a map to the circle $Z_f \to \mathbb{S}^1$ inducing $u_0 \colon G \to \mathbb{Z}$. However, it is unclear how to

"perturb" this map to the circle to obtain useful information about primitive integral elements $u \in H^1(G; \mathbb{R})$ which are near u_0, up to scaling.

To circumvent this problem, for any $\varphi \in \mathrm{Out}(F_N)$, given a tame graph map $f\colon \Gamma \to \Gamma$ which is a topological representative of φ (which always exists and is easy to find) and a folding sequence for f, we introduce a new 2–complex called the *folded mapping torus X_f* for φ. The complex X_f is obtained as a quotient of the full mapping torus Z_f of f by following the "folding line" from Γ to $\Gamma\varphi$ in the Culler–Vogtmann outer space determined by the folding sequence. See Section 2 for background, Section 3 for a description of the folding line, and Section 4 for the construction of X_f.

The local flow on $\Gamma \times [0, 1]$ given by $(x, t) \mapsto (x, t + s)$ descends to a semiflow ψ on X_f (see Section 4.2). Moreover, X_f is a $K(G, 1)$–space (see Corollary 4.9), and the homomorphism $u_0\colon G \to \mathbb{Z}$ is induced by a map $\eta_0\colon X_f \to \mathbb{S}^1$.

Although X_f is not a manifold, and so there is no notion of smooth 1–forms on X_f in the classical sense, the combinatorial structure of X_f nevertheless provides a way to perturb the maps η_0 and u_0. In Section 4.4, we construct a *trapezoid cell structure* on X_f (with a choice of orientation for the 1–cells) so that the restriction of η_0 to every semiflow line and every 1–cell is an orientation preserving local homeomorphism. We then define the *positive cone* of X_f to be the subset $\mathcal{A} = \mathcal{A}_{X_f} \subset H^1(X_f; \mathbb{R})$ of cohomology classes represented by cellular 1–cocycles that are positive on every positively oriented 1–cell. Moreover, this cell structure allows for the definition of a cellular 1–chain ϵ on X_f that serves as a "combinatorial" analogue of the Poincaré dual of the Euler class of the 2–plane subbundle $\ker(\nu_0)$ of TM considered above; see Remark 5.10 for more on this analogy. We summarize these basic properties of X_f in the following.

Proposition 1.2 Let $\varphi \in \mathrm{Out}(F_N)$ (where $N \geq 2$) be arbitrary, let $f\colon \Gamma \to \Gamma$ be a tame graph-map which is a topological representative of φ, and let X_f be a folded mapping torus for f, with the trapezoid cell structure, and let ψ be the associated semiflow. Let $u_0 \in H^1(G_\varphi; \mathbb{R})$ be the associated homomorphism for the splitting $G_\varphi = F_N \rtimes_\varphi \mathbb{Z}$. Then:

(a) The 2–complex X_f is a $K(G_\varphi, 1)$–space.

(b) The set $\mathcal{A}_{X_f} \subset H^1(X_f; \mathbb{R})$ is an open convex cone containing u_0.

(c) The singular 1–chain ϵ is a 1–cycle and thus defines an element also denoted $\epsilon \in H_1(X_f; \mathbb{R})$, which we call the E_0–class of X_f.

Remark 1.3 The construction of X, ψ, and $\mathcal{A}$ all depend on the choice of a representative f and a corresponding folding line for f. For any triple $(X', \psi', \mathcal{A}')$ obtained by making other choices, $\mathcal{A}' \cap \mathcal{A}$ is an open convex cone containing u_0 and X' is homotopy equivalent to X. The E_0–class ϵ is independent of these choices.

Working in the framework of X_f, we obtain the following results that precisely mirror the phenomenon described above for the 3–manifold M.

Theorem B Let $G_\varphi = F_N \rtimes_\varphi \mathbb{Z}$ be the mapping torus group of $\varphi \in \text{Out}(F_N)$ (where $N \geq 2$), let $f \colon \Gamma \to \Gamma$ be a tame graph-map which is a topological representative of φ, and let X_f be a folded mapping torus for f with the associated semiflow ψ, positive cone $\mathcal{A}_{X_f} \subset H^1(X_f; \mathbb{R})$, and E_0–class $\epsilon \in H_1(X_f; \mathbb{R})$. Given a positive cellular 1–cocycle $z \in Z^1(X_f; \mathbb{R})$ representing a primitive integral element $u \in \mathcal{A}_{X_f} \subset H^1(X_f; \mathbb{R}) = H^1(G_\varphi; \mathbb{R})$, there exists a map $\eta_z \colon X_f \to \mathbb{S}^1$ and a fiber $\Theta_z = \eta_z^{-1}(x_0)$ for some $x_0 \in \mathbb{S}^1$ so that:

(1) Θ_z is a finite connected topological graph such that the inclusion $\Theta_z \subset X_f$ is π_1–injective and such that $\pi_1(\Theta_z) = \ker(u) \leq \pi_1(X_f) = G_\varphi$.

(2) $u(\epsilon) = \chi(\Theta_z) = 1 - \text{rk}(\ker(u))$.

(3) Θ_z is a section of ψ.

(4) The first return map $f_z \colon \Theta_z \to \Theta_z$ of ψ is a homotopy equivalence with $f(V\Theta_z) \subset V\Theta_z$ and $(f_z)_*$ represents $\varphi_u \in \text{Out}(\ker(u))$.

See Section 1.5 and Section 1.6 for discussion of related constructions and results.

Remark 1.4 The map η_z depends on the choice of representative cocycle z and the graph Θ_z further depends on the point x_0. A different representative z' gives rise to a homotopic map $\eta_{z'} \simeq \eta_z$. It follows from the proof of Theorem B in Section 6.7 that for any $x_1 \in \mathbb{S}^1$, $\eta_{z'}^{-1}(x_1)$ can be "flowed onto" Θ_z producing a homotopy equivalence $\eta_{z'}^{-1}(x_1) \to \Theta_z$ homotopic to the inclusion $\eta_{z'}^{-1}(x_1) \hookrightarrow X$.

1.2 Train tracks and semiflows

We return to the motivational setting of surfaces and 3–manifolds. Thurston [61; 26] proved that a homeomorphism $F \colon S \to S$ is isotopic to a *pseudo-Anosov homeomorphism* if and only if $F_* \in \text{Out}(\pi_1(S))$ is *atoroidal*, meaning that it has no nontrivial periodic conjugacy classes. In fact, Thurston proved that this happens if and only if F_* is *geometrically atoroidal*, meaning that there are no nontrivial periodic conjugacy classes of elements represented by *simple* closed curves.

Thurston's hyperbolization theorem [52] moreover implies that the monodromy $F\colon S \to S$ of the fibration $\eta_0\colon M \to \mathbb{S}^1$ is isotopic to a pseudo-Anosov homeomorphism if and only if M is a hyperbolic manifold. Consequently, F is isotopic to a pseudo-Anosov homeomorphism if and only if the monodromy $F_\nu\colon S_\nu \to S_\nu$ of *any* fibration η_ν is isotopic to a pseudo-Anosov homeomorphism. If $F\colon S \to S$ is a pseudo-Anosov homeomorphism (not just isotopic to one), Fried [26, Exposé 14] proved a stronger statement: when the 1–form ν is sufficiently close to ν_0 (the 1–form representing u_0) then the monodromy $F_\nu\colon S_\nu \to S_\nu$ of the fibration η_ν, which is the first return map of the suspension flow of F, is already pseudo-Anosov (again, the isotopy is unnecessary). Fried's proof does not appeal to the hyperbolic structure on M.

The equivalent algebraic conditions on F_* described above have analogues for $\varphi \in \mathrm{Out}(F_N)$. However, these conditions on φ are not equivalent and have different implications. Firstly, we first say that φ is *atoroidal* if φ has no nontrivial periodic conjugacy classes in F_N. Appealing to the work of Bestvina and Feighn [7], Brinkmann [13] proved that $\varphi \in \mathrm{Out}(F_N)$ is atoroidal if and only if it is *hyperbolic*, meaning that the $G = F_N \rtimes_\varphi \mathbb{Z}$ is word-hyperbolic. It follows that the monodromy φ_u of any other primitive integral $u \in \mathscr{C}_G$ is necessarily hyperbolic as well.

Simple closed curves on a surface S provide a means of cutting S into pieces, thus the F_N–analogue of a periodic conjugacy class represented by a *simple closed curve* on S is a periodic free factor. By analogy with the geometric atoroidal condition, one therefore says that an element $\varphi \in \mathrm{Out}(F_N)$ is *fully irreducible* or *iwip* (for "irreducible with irreducible powers") if no positive power of φ preserves the conjugacy class of a proper free factor of F_N.

The fully irreducible condition and the hyperbolic (or atoroidal) condition are independent, with neither one implying the other. Thus one is often interested in outer automorphisms of F_N that are both hyperbolic and fully irreducible; such outer automorphisms provide the closest counterpart to (isotopy classes containing) pseudo-Anosov surface homeomorphisms and are also known to be generic in $\mathrm{Out}(F_N)$ (eg, see Rivin [55; 56] and Calegari and Maher [16]).

From a geometric/dynamical perspective, the analogue of a pseudo-Anosov homeomorphism is an *expanding irreducible train-track map* $f\colon \Gamma \to \Gamma$ representing $\varphi \in \mathrm{Out}(F_N)$ (see Definitions 2.3 and 2.11). Such representatives are automatically tame (see Lemma 2.12). Train-track maps were introduced by Bestvina and Handel [9], who also showed that an expanding irreducible train-track representative exists for every fully irreducible $\varphi \in \mathrm{Out}(F_N)$, and more generally for every irreducible φ of infinite order. Our next result proves that if the folded mapping torus X_f is constructed using such a f, then for every primitive integral $u \in \mathcal{A}$ represented by a positive cocycle z,

the first return map $f_z\colon \Theta_z \to \Theta_z$ from Theorem B is also an expanding irreducible train-track representative of $\varphi_u \in \mathrm{Out}(\ker(u))$; this is analogous to Fried's result.

If φ is additionally assumed to be fully irreducible and hyperbolic, then in view of the hyperbolicity of $G \cong G_\varphi$, the monodromy φ_u of any other primitive integral $u \in \mathscr{C}_G$ is also hyperbolic. However, it is not at all clear whether φ_u must also be fully irreducible. In general, proving full irreducibility for free group automorphisms is a fairly hard task; see for example Clay and Pettet [17]. Nevertheless, by using the structure of the folded mapping torus X_f, we prove that for any primitive integral $u \in \mathcal{A}$, the monodromy $\varphi_u \in \mathrm{Out}(\ker(u))$ is indeed fully irreducible:

Theorem C *Let $\varphi \in \mathrm{Out}(F_N)$ (where $N \geq 2$) be an outer automorphism represented by an expanding irreducible train track map $f\colon \Gamma \to \Gamma$, let X_f be a folded mapping torus, and ψ the associated semiflow. Let $u \in \mathcal{A}_{X_f} \subset H^1(G_\varphi; \mathbb{R}) = H^1(X_f; \mathbb{R})$ be a primitive integral element which is represented by a positive 1–cocycle $z \in Z^1(X_f; \mathbb{R})$, and let $\Theta_z \subset X_f$ be the topological graph from Theorem B so that the first return map $f_z\colon \Theta_z \to \Theta_z$ is a homotopy equivalence representing φ_u. Then:*

(1) *The first return map $f_z\colon \Theta_z \to \Theta_z$ is an expanding irreducible train-track map (with respect to an appropriately chosen linear graph structure on Θ_z).*

(2) *If we additionally assume that φ is hyperbolic and fully irreducible, then $\varphi_u \in \mathrm{Out}(\ker(u))$ is also hyperbolic and fully irreducible.*

We note that if in item (2) above φ is only assumed to be fully irreducible (and not hyperbolic), then φ_u need not be fully irreducible; see Remark 7.11.

Our definitions require that train-track maps are piecewise linear on every edge with respect to some linear structure on the graph; see Sections 2.4–2.6 for precise statements. While this property is part of the assumption on f, it becomes part of the conclusion for f_z for any positive cocycle z. The graph Θ_z does not come equipped with any natural linear structure, and thus part of the proof of Theorem C includes the construction of a natural linear structure. This structure is deduced as a consequence of a general dynamical result (which appears to be of independent interest), Theorem A.3, about topological graph maps that are "sufficiently mixing" (a notion we call *expanding on all scales*).

To prove that φ_u is fully irreducible, under the assumption that φ is hyperbolic and fully irreducible, we utilize a criterion of full irreducibility for hyperbolic automorphisms obtained by Kapovich [39]; see also Pfaff [53; 54]. That criterion says that if $f'\colon \Gamma' \to \Gamma'$ is a train-track map representative of a hyperbolic automorphism φ' then φ' is fully irreducible if and only if some power $A(f')^m$ of the transition matrix $A(f')$

is strictly positive and the Whitehead graphs of all vertices of Γ' are connected. We first obtain a general result strengthening this criterion (Theorem B.3 below) and show that the condition $A(f')^m > 0$ can be replaced by just requiring the transition matrix $A(f')$ to be irreducible. We then verify that $f_z\colon \Theta_z \to \Theta_z$ is a train-track map with an irreducible transition matrix and with connected Whitehead graphs by exploiting the properties of f and of the semiflow ψ. This allows us to conclude that φ_u is fully irreducible. As another useful application of Theorem B.3, we show in Corollary B.4 that for hyperbolic elements in $\mathrm{Out}(F_N)$ being irreducible is equivalent to being fully irreducible.

1.3 Entropy and stretch factors

For a pseudo-Anosov homeomorphism $F\colon S \to S$, the *stretch factor* or *dilatation* $\lambda(F)$ records the asymptotic growth rate for lengths of curves under iteration. That is,

$$\log(\lambda(F)) = \lim_{n\to\infty} \frac{\log(\ell(F^n(\gamma)))}{n},$$

where $\ell(\cdot)$ is the length of the geodesic representative in any fixed metric on S and γ is any homotopically nontrivial curve on S. From the work of Fathi and Shub [26, Exposé 10], $\log(\lambda(F))$ is the topological entropy $h(F)$ of F. Returning to the 3–manifold M, we continue to assume that some (hence every) monodromy $F\colon S \to S$ for a primitive integral $u_0 \in H^1(M;\mathbb{R})$ is pseudo-Anosov, and we let $\mathscr{C}_0 \subset \mathscr{C}_G$ denote the component of the fg cone containing u_0. Fried [29] proved that the function which assigns to every $u \in \mathscr{C}_0$ the topological entropy $h(F_u)$ of the monodromy $F_u\colon S_u \to S_u$, extends to a continuous, homogeneous of degree -1 function

$$\mathfrak{h}\colon \mathscr{C}_0 \longrightarrow \mathbb{R}$$

for which the reciprocal $1/\mathfrak{h}$ is concave. See Long and Oertel [44], Oertel [51] and McMullen [45] for other proofs and refinements of this result.

Any $\varphi \in \mathrm{Out}(F_N)$ also has a *stretch factor* $\lambda(\varphi) \geq 1$ defined as

$$\log \lambda(\varphi) = \sup_{w \in F_N} \lim_{n\to\infty} \frac{\log \|\varphi^n(w)\|_A}{n},$$

where $\|w\|_A$ denotes the cyclically reduced length of w with respect to some fixed free basis A of F_N. It is known that the above definition does not depend on the choice of A and that the sup in the above formula is actually a max.

An element $\varphi \in \mathrm{Out}(F_N)$ is called *exponentially growing* if $\lambda(\varphi) > 1$. It is known that every fully irreducible φ is exponentially growing. More generally, if $\varphi \in \mathrm{Out}(F_N)$

admits an expanding irreducible train-track representative $f\colon \Gamma \to \Gamma$, then φ is exponentially growing and $\lambda(\varphi) = \lambda(f)$, where $\lambda(f)$ is the Perron–Frobenius eigenvalue of the transition matrix $A(f)$ of f (see Section 2.4). We note that in this situation $\lambda(f)$ depends only on φ and not on the choice of f, and that it is alternatively calculated via topological entropy $h(f)$ as $h(f) = \log(\lambda(f))$ (see Proposition A.1 below for a more general version of this statement). This discussion applies, for example, to all fully irreducible automorphisms and, more generally, to all irreducible elements of $\mathrm{Out}(F_N)$ of infinite order.

The final part of Theorem A provides the analogue of Fried's result on entropy.

Theorem D Let $\varphi \in \mathrm{Out}(F_N)$ (where $N \geq 2$) be an element represented by an expanding irreducible train track map $f\colon \Gamma \to \Gamma$. Let X_f be a folded mapping torus, ψ the associated semiflow, and $\mathcal{A}_{X_f} \subset H^1(X_f; \mathbb{R})$ the positive cone of X_f. Then there exists a continuous, homogeneous of degree -1 function $\mathfrak{H}\colon \mathcal{A}_{X_f} \to \mathbb{R}$ such that for every primitive integral $u \in \mathcal{A}_{X_f}$

$$\log(\lambda(\varphi_u)) = \mathfrak{H}(u).$$

Moreover, $1/\mathfrak{H}$ is concave, and hence $\mathfrak{H}$ is convex.

For any positive cocycle $z \in Z^1(X_f; \mathbb{R})$ representing a primitive integral $u \in \mathcal{A}$ as above, we have $h(f_z) = \log(\lambda(f_z)) = \log(\lambda(\varphi_u))$, where $f_z\colon \Theta_z \to \Theta_z$ is the expanding irreducible train-track map provided by Theorem C. The proof of Theorem D involves showing that $[z] \mapsto h(f_z)$ extends to a function $\mathfrak{H}$ which is continuous, homogeneous of degree -1, and for which $1/\mathfrak{H}$ is concave; this argument closely follows Fried [29], with some important modifications.

The rough idea of the proof is as follows. According to the variational principle, $h(f_z) = \sup_\nu h_\nu(f_z)$, where ν runs over all f_z–invariant Borel probability measures on Θ_z and where $h_\nu(f_z)$ is the Kolmogorov–Sinai entropy with respect to ν. There is a natural bijection between ψ–invariant finite measures on X_f and f_z–invariant finite measures on Θ_z, which we write as $\mu \mapsto \overline{\mu}_0$ (with $\overline{\mu}$ representing the normalized probability measure). Then, according to a general result of Abramov, for any ψ–invariant probability measure μ, $h_{\overline{\mu}}(f_z) = h_\mu(\psi)/\overline{\mu}_0(\Theta_z)$, where $h_\mu(\psi)$ is the entropy of the flow with respect to μ (which depends on μ, but not on the class $u \in \mathcal{A}$ represented by the cocycle $z = z_u$). Then one gets $1/h(f_z) = \inf_\mu \overline{\mu}_0(\Theta_z)/h_\mu(\psi)$. A key step in Fried's argument is to then show that for any μ the function $u \mapsto \overline{\mu}_0(\Theta_{z_u})$ extends to a *linear* continuous function $\mu_*\colon \mathcal{A} \to \mathbb{R}$. The function $W\colon u \mapsto \inf_\mu \mu_*(u)/h_\mu(\psi)$ is now given as the infimum of a family of linear continuous functions. Therefore W is concave on $\mathcal{A}$ and hence W is continuous. For every primitive integral $u \in \mathcal{A}$ we

have $W(u) = 1/h(f_{z_u})$, and the conclusion of Theorem D follows. There are several difficulties with carrying out this line of argument that we need to overcome. First, Fried's and Abramov's results are for flows, while ψ is only a semiflow. Thus we need to immediately pass, via the "natural extension" construction in dynamics, from a semiflow ψ on X_f to a related flow ψ' on another space X'_f. We then need to make sure that the above argument gets correctly translated to this new setting. Another difficulty is that the geometric "surgery" argument Fried uses to prove linearity of μ_* does not work well for X'_f. Thus we deploy a homological argument for that purpose instead of attempting to mimic the argument of Fried.

1.4 Connections to small dilatation free group automorphisms.

As a corollary of Theorem D, we see that the function $u \mapsto \log(\lambda(\varphi_u))\chi(\ker(u))$ defined on primitive integral classes u extends to a continuous function which is constant on rays. Thus from any $G = G_\varphi$ with $b_1(G) \geq 2$ and φ represented by an expanding irreducible train track map, one obtains sequences $\{\varphi_k \colon F_{N_k} \to F_{N_k}\}_{k=1}^\infty$ of automorphisms φ_k represented by expanding irreducible train track maps for which $N_k \to \infty$ and $N_k \log(\lambda(\varphi_k))$ is uniformly bounded above and below by positive constants. Moreover, by Theorem C(2), if φ is fully irreducible and hyperbolic so are all φ_k; see Corollary 8.7. This corollary mirrors a construction of pseudo-Anosov homeomorphisms with the analogous property using 3–manifolds as described by McMullen [45, Section 10]. Corollary 8.7 also complements the work of Algom-Kfir and Rafi [4] who showed that any sequence $\{\varphi_k \colon F_{N_k} \to F_{N_k}\}_{k=1}^\infty$ for which $N_k \log(\lambda(\varphi_k))$ is bounded must come from this construction (up to a certain surgery procedure). This was in turn a free-by-cyclic analogue of a result of Farb, Leininger and Margalit [24] (see also Agol [2]) complementing McMullen's construction. So in some sense, Corollary 8.7 brings this picture full circle.

1.5 Related results and constructions

As mentioned at the start of the introduction, it already follows from general facts about the BNS–invariant and related constructions (see for example Bieri, Neumann and Strebel [10] and Levitt [43]) that under the assumptions of Theorem B there exists some open cone $\mathcal{C}_G \subseteq H^1(G_\varphi)$ containing u_0, such that every primitive integral $u \in \mathcal{C}_G$ has finitely generated kernel $\ker(u)$. Additional general cohomological considerations (eg see Geoghegan, Mihalik, Sapir and Wise [36, Theorem 2.6 and Remark 2.7]) further imply that $\ker(u)$ is a free group. Consequently, u determines another splitting of G_φ as a (fg free)-by-(infinite cyclic) group. We do not use these general facts in the proof of Theorem B and instead directly establish for every primitive integral $u \in \mathcal{A}$ that $\ker(u)$ is a free group of finite rank.

The folded mapping torus X_f and the semiflow ψ are similar to Gautero's construction of a "dynamical 2–complex" [32; 33; 34; 35] (see also the discussion in [32, Section 5] for connections to other related notions) and the unpublished construction of Z Wang [63] of a "decomposed mapping torus" and various "folded complexes". Our folded mapping torus X_f turns out the be homeomorphic to the complex constructed by Brinkmann and Schleimer in [14], which treats only the case when the automorphism is induced by a homeomorphism of a surface with one boundary component. However, [14] does not construct a semiflow on the complex, but rather uses it to construct an ideal triangulation of the mapping torus of the homeomorphism.

Gautero builds a 2–complex from a topological representative $f\colon \Gamma \to \Gamma$, where Γ is a trivalent graph, and a path in outer space from Γ to $\Gamma\varphi$ using Whitehead moves (see also the variation in [35] using Stalling's folds instead of Whitehead moves, which is a modification of his original construction and is more closely related to the one here). Wang [63] on the other hand constructs from a sequence of Stallings folds from Γ to $\Gamma\varphi$ first an iterated mapping cylinder which he calls a *decomposed mapping torus*. This space is homotopy equivalent to the mapping torus and also admits a semiflow. Then for each cohomology class $u \in H^1(G_\varphi; \mathbb{R})$ satisfying certain conditions he builds what he calls a *folded complex* $Y_{f,u}$ which is a quotient of his decomposed mapping torus.

In both of these constructions, the main goal is that of finding cross sections in analogy with Fried's work [30]. In particular, these constructions are likely sufficient for proving analogues of Theorem B. However, it is unclear how to attack the remainder of Theorem A and Theorems C and D, with these constructions. It is worth noting that in Wang's construction, he uses sections of the folded complexes to find sections of the original mapping torus (though the latter will likely not satisfy Theorem B). In any event, we try to mention the relevant connections to Gautero's and Wang's construction wherever possible throughout the paper.

1.6 Alexander norm

In the 3–manifold setting, the *Thurston norm* is a seminorm

$$\mathfrak{n}_T\colon H^1(M;\mathbb{R}) = H^1(G_M;\mathbb{R}) \longrightarrow \mathbb{R}$$

defined by Thurston [60] taking integral values on primitive integral points (see also [31; 44; 47; 48; 50; 57]). He then observed that the unit ball $\boldsymbol{B}_T$ is a polyhedron. The norm of a primitive integral class is a measure of the complexity of the class, and for any class u represented by a fibration $M \to \mathbb{S}^1$, the norm is precisely $\mathfrak{n}_T(u) = -u(e)$, where e is again the Poincaré dual of the associated Euler class. In fact, in this setting $\mathscr{C}_{G_M}$ is a union of cones on the interiors of certain top-dimensional faces of

the polyhedron $\boldsymbol{B}_T$. In [46], McMullen constructed a related *Alexander norm* $\mathfrak{n}_A$ on $H^1(G;\mathbb{R})$ for any finitely presented group G, which again has a polyhedral unit ball $\boldsymbol{B}_A$ (and so as with the Thurston norm, $\mathfrak{n}_A$ is linear on cones on faces). In the case of the 3–manifold group G_M the Thurston and Alexander norms actually agree on $\mathscr{C}_{G_M}$ when $\dim(H^1(M;\mathbb{R})) > 1$; however Dunfield [23] proved that in general they do not agree on all of $H^1(G_M;\mathbb{R})$.

In the case of our free-by-cyclic group G_φ, since G_φ has a deficiency 1 presentation, the results of McMullen [46, Theorem 4.1] and Button [15, Theorem 3.1] together imply that when $\dim(H^1(G_\varphi;\mathbb{R})) > 1$, the Alexander norm calculates $\mathfrak{n}_A(u) = b_1(\ker(u)) - 1$ for any primitive integral $u \in \mathscr{C}_{G_\varphi}$, where $b_1(\ker(u)) = \dim(H^1(\ker(u);\mathbb{R}))$. As mentioned above, general cohomological considerations (eg see [36, Theorem 2.6 and Remark 2.7]) additionally imply that $\ker(u)$ is free for any primitive integral $u \in H^1(G_\varphi;\mathbb{R})$ with $\ker(u)$ finitely generated; thus we have $b_1(\ker(u)) = \mathrm{rk}(\ker(u))$ in this case. Moreover, Dunfield's work [23] implies that any primitive integral $u \in \mathscr{C}_{G_\varphi}$ must lie over the interior of a top-dimensional face of $\boldsymbol{B}_A$; therefore $\mathfrak{n}_A$ restricts to a linear function on an open cone neighborhood of any such u. Putting these results together, it follows that for every primitive integral $u_0 \in \mathscr{C}_{G_\varphi}$, the assignment $u \mapsto 1 - \mathrm{rk}(\ker(u))$ for primitive integral u extends to a linear function $u \mapsto -\mathfrak{n}_A(u)$ on some open cone containing u_0 (namely the intersection of $\mathscr{C}_{G_\varphi}$ with the cone on a top-dimensional face of $\boldsymbol{B}_A$). Thus Theorem B(2) gives a realization of this linear function as the geometrically defined E_0–class $\epsilon \in H^1(G_\varphi;\mathbb{R})$. Moreover, it shows that pairing any element of $\mathcal{A}_{X_f}$ with the E_0–class gives precisely the (negative of the) Alexander norm and thus strengthens the analogy with McMullen's result for 3–manifolds.

Remark 1.5 The results and techniques developed in this manuscript have served as the basis for a number of advances in the study of free-by-cyclic groups and $\mathrm{Out}(F_N)$ since it was made publicly available in January 2013. For example, in [19], we continue our analysis of the folded mapping torus X_f and develop an analogy with McMullen's Teichmüller polynomial [45]. Thus [19] strengthens the connection with the BNS–invariant and provides a new proof of Theorem D. However, while the results in [19] rely on the main constructions of this paper, most of the new results there are orthogonal to those of the current paper. Then in [21], we draw on the results from both of these papers to provide examples of fully irreducible automorphisms of free groups whose stretch factors illustrate a new phenomenon, complementing the results of Handel and Mosher [37].

In addition, concurrent with [19], Algom-Kfir, Hironaka, and Rafi [3] construct their own analogue of McMullen's Teichmüller polynomial. Their results rely on the construction of the folded mapping torus X_f, properties of $\mathcal{A}_{X_f}$, and in particular Theorem C.

2 Background

2.1 Graphs

A *topological graph* Γ is a locally finite 1–dimensional cell complex (where locally finite means that every 0–cell is incident to only finite many 1–cells). The 0–cells of Γ are called *vertices* and the set of all vertices of Γ is denoted by $V\Gamma$. Open 1–cells of Γ are called *topological edges* of Γ and the set of these is denoted $E_{\text{top}}\Gamma$. Every topological edge of Γ is homeomorphic to an open interval and thus admits exactly two orientations. An *oriented edge* e of Γ is a topological edge together with a choice of an orientation on it. We denote the set of all oriented edges of Γ by $E\Gamma$. If e is an oriented edge of Γ, then e^{-1} denotes the same topological edge with the opposite orientation. Thus $^{-1}\colon E\Gamma \to E\Gamma$ is a fixed-point-free involution. For an oriented edge $e \in E\Gamma$ the attaching maps for the underlying 1–cell canonically define the *initial vertex* $o(e) \in V\Gamma$ and the *terminal vertex* $t(e) \in V\Gamma$. By construction, for any $e \in E\Gamma$ we have $o(e^{-1}) = t(e)$ and $t(e^{-1}) = o(e)$. For $e \in E\Gamma$ we denote by $\overline{e}$ the closure of e in Γ.

An *orientation* on a topological graph Γ is a partition $E\Gamma = E_+\Gamma \sqcup E_-\Gamma$ such that for every $e \in E\Gamma$ we have $e \in E_+\Gamma$ if and only if $e^{-1} \in E_-\Gamma$.

Given a topological graph Γ and $v \in V\Gamma$ the *link* $\text{Lk}_\Gamma(v)$ consists of all $e \in E\Gamma$ with $o(e) = v$. A *turn* in Γ is an unordered pair e, e' of (not necessarily distinct) elements of $\text{Lk}_\Gamma(v)$. A turn e, e' is *degenerate* if $e = e'$ and *nondegenerate* otherwise. We denote the set of all turns in Γ by $\text{Turns}(\Gamma)$.

Given a finite set of points in a topological graph Γ, one can naturally construct a topological graph Γ' as the *subdivision*, in which the vertex set contains the original vertex set, together with the finite set of points.

2.2 Linear and metric structures

Given a topological graph Γ, a *linear atlas of charts* for Γ is a set

$$\Lambda = \{(e_\alpha, j_\alpha, [a_\alpha, b_\alpha])\}_{\alpha \in J},$$

where J is some index set, $e_\alpha \in E\Gamma$ for every α, and $j_\alpha \colon [a_\alpha, b_\alpha] \to \Gamma$ is a characteristic map: a continuous function for which the restriction to (a_α, b_α) is an orientation preserving homeomorphism onto the interior of e_α (in particular, $j_\alpha(a_\alpha) = o(e_\alpha)$ and $j_\alpha(b_\alpha) = t(e_\alpha)$). We further assume that (1) for every $e \in E\Gamma$, there exists at least one $\alpha \in J$ so that $e = e_\alpha$; (2) for any $\alpha, \beta \in J$ such that $e_\alpha = e_\beta$ then the restriction $j_\alpha^{-1} j_\beta \colon (a_\beta, b_\beta) \to (a_\alpha, b_\alpha)$ is the unique orientation preserving affine map; and (3) for every α there exists α' so that $e_{\alpha'} = e_\alpha^{-1}$ and $j_{\alpha'} = j_\alpha \circ \sigma$, where $\sigma \colon [a_\alpha, b_\alpha] \to [a_\alpha, b_\alpha]$ is the unique affine, orientation reversing involution. A *linear structure* on a topological graph Γ is a linear atlas $\Lambda = \{(e_\alpha, j_\alpha, [a_\alpha, b_\alpha])\}_{\alpha \in J}$ which is maximal with respect to inclusion.

A *metric atlas* on a topological graph Γ is an linear atlas $\mathcal{L} = \{(e_\alpha, j_\alpha, [a_\alpha, b_\alpha])\}_{\alpha \in J}$ on Γ such that whenever $\alpha, \beta \in J$ satisfy $e_\alpha = e_\beta$ or $e_\alpha = e_\beta^{-1}$, then $|b_\alpha - a_\alpha| = |b_\beta - a_\beta|$. Note that this implies that whenever $e_\alpha = e_\beta$ then the restriction

$$j_\alpha^{-1} j_\beta \colon (a_\beta, b_\beta) \longrightarrow (a_\alpha, b_\alpha)$$

is the unique orientation preserving isometry and whenever $e_\alpha = e_\beta^{-1}$ then the restriction

$$j_\alpha^{-1} j_\beta \colon (a_\beta, b_\beta) \longrightarrow (a_\alpha, b_\alpha)$$

is the unique orientation reversing isometry. A *metric structure* on a topological graph Γ is a metric atlas $\mathcal{L}$ on Γ which is maximal with respect to inclusion.

A linear or metric structure on Γ uniquely determines one on any subdivision of Γ. Furthermore, linear and metric structures naturally lift under covering maps.

Given a metric structure $\mathcal{L} = \{(e_\alpha, j_\alpha, [a_\alpha, b_\alpha])\}_{\alpha \in J}$, on a graph Γ, we can define $\mathcal{L}(e_\alpha) = b_\alpha - a_\alpha > 0$ to be the $\mathcal{L}$–length of e_α. From the definition of metric structure, this procedure uniquely defines a function we also denote $\mathcal{L} \colon E\Gamma \to \mathbb{R}_+$ satisfying $\mathcal{L}(e) = \mathcal{L}(e^{-1})$, for every $e \in E\Gamma$. Any such function on $E\Gamma$ is called a *length function for* Γ. The *volume* of $(\Gamma, \mathcal{L})$ is then defined to be the (possibly infinite) quantity $\mathrm{vol}_\mathcal{L}(\Gamma) := \frac{1}{2} \sum_{e \in E\Gamma} \mathcal{L}(e)$. Since Γ is locally finite, there is a unique geodesic metric $d_\mathcal{L}$ on each component of Γ defining the original weak topology on the 1–complex Γ for which the characteristic maps of that metric structure are all local isometries. Conversely, given a geodesic metric d on the components of Γ, there is a canonical metric structure $\mathcal{L}$ so that $d = d_\mathcal{L}$.

Any metric structure $\mathcal{L}$ on Γ extends to a canonical linear structure Λ, meaning that $\mathcal{L} \subset \Lambda$. Also, given a linear structure Λ on Γ and any length function $L \colon E\Gamma \to \mathbb{R}_+$, there exists a unique metric structure $\mathcal{L} \subset \Lambda$ for which $\mathcal{L}(e) = L(e)$ for all $e \in E\Gamma$. That is, when a linear structure Λ on Γ is fixed, we can uniquely specify a metric structure $\mathcal{L} \subset \Lambda$ on Γ by simply specifying the lengths of edges.

A *graph* or *linear graph* is a topological graph Γ together with a linear structure Λ on it. A *metric graph* is similarly defined as a topological graph Γ together with a metric structure $\mathcal{L}$. Via the canonical linear structure determined by a metric structure, any metric graph is also naturally a linear graph.

2.3 Paths

If Γ is a topological graph, for $n \geq 1$ a *combinatorial edge-path* of *simplicial length* n in Γ is a sequence $\gamma = e_1 e_2 \cdots e_n$ of oriented edges of Γ such that for $i = 1, \ldots, n-1$ we have $t(e_i) = o(e_{i+1})$. We denote $|\gamma| = n$ and put $o(\gamma) := o(e_1)$ and $t(\gamma) := t(e_n)$. Also, for a vertex v of Γ we view $\gamma = v$ as a combinatorial edge-path of simplicial length $|\gamma| = 0$ with $o(\gamma) = t(\gamma) = v$. For a combinatorial edge-path $\gamma = e_1 \cdots e_n$ we denote $\gamma^{-1} = e_n^{-1} \cdots e_1^{-1}$, so that γ^{-1} is a combinatorial edge-path from $t(\gamma)$ to $o(\gamma)$ with $|\gamma| = |\gamma^{-1}| = n$. For $\gamma = v$ (where $v \in V\Gamma$) we put $\gamma^{-1} = \gamma = v$.

If $\gamma = e_1 \cdots e_n$ is a combinatorial edge-path and $\boldsymbol{t}$ is a turn, we say that this turn is *contained* in γ if there exists $2 \leq i \leq n$ such that $\boldsymbol{t} = \{e_{i-1}^{-1}, e_i\}$ or $\boldsymbol{t} = \{e_i^{-1}, e_{i-1}\}$.

A *topological edge-path* of *simplicial length* n is a continuous map $g \colon [a, b] \to \Gamma$ such that there exists a combinatorial edge-path $\gamma_g = e_1 \cdots e_n$ in Γ and a subdivision $a = c_0 < c_1 < \cdots < c_n = b$ such that for each $i = 1, \ldots, n$ we have $g(c_{i-1}) = o(e_i)$, $g(c_i) = t(e_i)$, and the restriction $g|_{(c_{i-1}, c_i)}$ is an orientation preserving homeomorphism onto e_i. We say that γ_g is the combinatorial edge-path *associated to* g. Writing $o(g) := o(e_1)$ and $t(g) := t(e_n)$, then g is a path in Γ from $o(g)$ to $t(g)$. For any vertex v of Γ and $[a, b] \subset \mathbb{R}$, we also regard the constant map $g \colon [a, b] \to \Gamma$ given by $g(s) = v$ for all $s \in [a, b]$ as a topological edge-path of *simplicial length* 0 with $o(g) = t(g) = v$. If Γ is a linear graph, then a topological edge-path $g \colon [a, b] \to \Gamma$ is a *PL edge-path* if it additionally satisfies the following condition. For every edge e_i of γ_g and any chart $(e_\alpha, j_\alpha, [a_\alpha, b_\alpha])$ from the linear atlas with $e_\alpha = e_i$, the restriction of g to $[c_{i-1}, c_i]$ is equal to $j_\alpha \circ \tau_{\alpha, i}$ where $\tau_{\alpha, i}$ is the unique orientation-preserving affine homeomorphism from $[c_{i-1}, c_i]$ to $[a_\alpha, b_\alpha]$.

By dropping the requirement that $g(a)$ and $g(b)$ be vertices of Γ, the above definition is straightforwardly modified to get the notion of a *PL path* in Γ.

If we have a combinatorial path γ with $|\gamma| = n$ in a topological (respectively, linear) graph Γ, there always exists a topological (respectively, PL) edge-path $g \colon [a, b] \to \Gamma$

with $\gamma_g = \gamma$. Therefore we will often suppress the distinction between combinatorial, topological, and PL edge-paths and just call them edge-paths, unless we wish to emphasize the distinction. We call edge-paths of positive simplicial length *nondegenerate* and edge-paths of simplicial length 0 *degenerate*.

A combinatorial edge-path γ in Γ is *tight* or *reduced* if γ does not contain a subpath of the form ee^{-1} where $e \in E\Gamma$ (that is, if γ contains no degenerate turns). A topological or PL edge-path is *tight* or *reduced* if the associated combinatorial edge-path is so. More generally, a PL path $g\colon [a,b] \to \Gamma$ is *tight* if g is locally injective.

If $(\Gamma, \mathcal{L})$ is a metric graph, then the $d_{\mathcal{L}}$–length of a PL–path g in $(\Gamma, \mathcal{L})$ is denoted by $|g|_{\mathcal{L}}$. Similarly, the length of a combinatorial edge-path $\gamma = e_1 \cdots e_n$ will be denoted by $|\gamma|_{\mathcal{L}} = \sum_{i=1}^{n} \mathcal{L}(e_i)$.

2.4 Graph maps

Definition 2.1 (Graph-map) If Γ and Δ are topological graphs, a *topological graph-map* $f\colon \Gamma \to \Delta$ is a continuous map such that $f(V\Gamma) \subseteq V\Delta$ and such that for every $e \in E\Gamma$ $f|_e$ is a topological-edge-path (possibly degenerate) from $f(o(e))$ to $f(t(e))$ in Δ. More precisely, this means that for any characteristic map $j\colon [a,b] \to e$, the composition $f \circ j\colon [a,b] \to \Delta$ is a topological edge-path. In this situation we will often denote the combinatorial edge-path in Δ associated to $f \circ j$ by $f(e)$. If Γ and Δ are linear graphs and in addition we assume that for every $e \in E\Gamma$, $f|_e$ is a PL edge-path from $f(o(e))$ to $f(t(e))$, then we call f a *linear graph-map* or just a *graph-map*.

We note that if Γ and Δ are equipped with linear structures, then any topological graph-map $f\colon \Gamma \to \Delta$ is homotopic rel $f^{-1}(V\Delta) \supset V\Gamma$ to a unique (linear) graph-map $f'\colon \Gamma \to \Delta$.

Example 2.2 (Running example) Here we introduce an example that we will develop throughout the paper as a means of illustrating key ideas. Let Γ denote the graph in Figure 1. This graph has four edges, oriented as shown and labeled $\{a, b, c, d\} = E_+\Gamma$. We consider a graph-map $f\colon \Gamma \to \Gamma$ under which the edges of Γ map to the combinatorial edge-paths $f(a) = d$, $f(b) = a$, $f(c) = b^{-1}a$, and $f(d) = ba^{-1}db^{-1}ac$.

To describe this example precisely, give Γ a linear structure and let Δ be the graph obtained by subdividing Γ as pictured on the right of Figure 1. The identity then gives a graph-map $\iota\colon \Gamma \to \Delta$. Labeling the edges of Δ as indicated, there is a unique linear
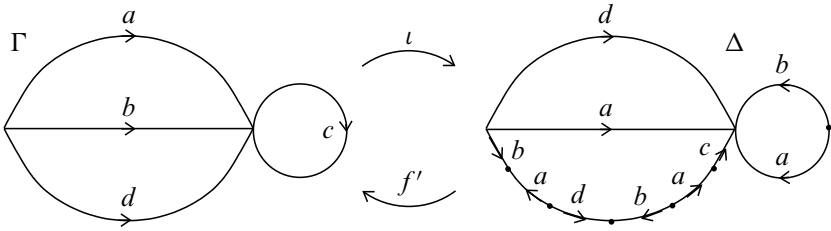
Figure 1: An example graph-map. Left: Original graph Γ. Right: Subdivided graph Δ with labels. Our example $f: \Gamma \to \Gamma$ is the composition of the "identity" $\iota: \Gamma \to \Delta$ with the map $f': \Delta \to \Gamma$ that sends edges to edges preserving labels and orientations.

graph-map $f': \Delta \to \Gamma$ for which each edge of Δ maps by an affine map to the edge of the same label in Γ. Our example is defined to be the graph-map

$$f := f' \circ \iota: \Gamma \longrightarrow \Gamma.$$

Definition 2.3 (Regular and expanding maps) A topological graph-map $f: \Gamma \to \Delta$ is *regular* if for any $e \in E\Gamma$ the combinatorial edge-path $f(e)$ is nondegenerate and tight (thus $|f(e)| > 0$). We say that a topological graph map $f: \Gamma \to \Gamma$ is *expanding* if for every edge $e \in E\Gamma$ we have $|f^n(e)| \to \infty$ as $n \to \infty$.

The next definition is convenient for carrying out the constructions in subsequent sections.

Definition 2.4 (Combinatorial maps) A *combinatorial graph map* is a regular, linear graph map $f: \Gamma \to \Gamma$ where Γ is a metric graph with all edge lengths equal to 1, and where the subdivision of any edge e along $f^{-1}(V\Gamma)$ is into equal length parts.

Remark 2.5 Given any regular topological graph map $f_0: \Gamma \to \Gamma$, then it is easy to construct a metric structure on Γ and a combinatorial graph map $f: \Gamma \to \Gamma$ that is homotopic to f_0 and for which the combinatorial edge-paths $f(e)$ and $f_0(e)$ agree for every edge e. In fact, under an appropriate expansion condition on f_0 (Definition 2.10), one can use Corollary A.9 from Appendix A to arrange for $f = f_0$ (ie no homotopy is needed).

Let $f: \Gamma \to \Delta$ be a regular topological graph-map and let $v \in V\Gamma$. Define an equivalence relation $\sim_{f,v}$ on $\mathrm{Lk}_\Gamma(v)$ by saying that for $e, e' \in \mathrm{Lk}_\Gamma(v)$ we have $e \sim_{f,v} e'$ whenever the combinatorial edge-paths $f(e)$ and $f(e')$ start with the same (oriented) edge of Δ.

Definition 2.6 (Tame maps) A topological graph-map $f\colon \Gamma \to \Delta$ is *tame* if f is regular and if for every vertex $v \in V\Gamma$ the relation $\sim_{f,v}$ partitions $\mathrm{Lk}_\Gamma(v)$ into at least two equivalence classes. Note that if $f\colon \Gamma \to \Delta$ is tame then every vertex of Γ has valence ≥ 2.

Definition 2.7 (Taken turns and Whitehead graphs) Let $f\colon \Gamma \to \Gamma$ be a regular graph-map. A turn e_1, e_2 in Γ is *taken* by f if there exist $e \in E\Gamma$ and $n \geq 1$ such that $e_1^{-1}e_2$ is a subpath of $f^n(e)$ (that is, if the turn e_1, e_2 is contained in $f^n(e)$). For any vertex $v \in V\Gamma$ we define the *Whitehead graph* of v, denoted $\mathrm{Wh}_\Gamma(f, v)$ or just $\mathrm{Wh}(f, v)$ if Γ is fixed, as follows. The graph $\mathrm{Wh}_\Gamma(f, v)$ is a simple graph whose vertex set is $\mathrm{Lk}_\Gamma(v)$. Thus the vertices of $\mathrm{Wh}_\Gamma(f, v)$ are exactly all those $e \in E\Gamma$ such that $o(e) = v$. Two distinct elements $e_1, e_2 \in \mathrm{Lk}_\Gamma(v)$ are adjacent as vertices in $\mathrm{Wh}_\Gamma(f, v)$ whenever the turn e_1, e_2 is taken by f. (Note that since $e_1 \neq e_2$, the turn e_1, e_2 is nondegenerate.)

Remark 2.8 For a regular graph-map $f\colon \Gamma \to \Gamma$, denote by $\mathrm{Turns}_n(f)$ the set of all turns contained in the paths $\{f^k(e) \mid e \in E\Gamma, 1 \leq k \leq n\}$. Then $\mathrm{Turns}_n(f) \subseteq \mathrm{Turns}_{n+1}(f)$ for all $n \geq 1$. Put $\mathrm{Turns}_\infty(f) := \bigcup_{n \geq 1} \mathrm{Turns}_n(f)$. By definition, for a vertex v of Γ the Whitehead graph $\mathrm{Wh}_\Gamma(f, v)$ records all the nondegenerate turns at v that belong to $\mathrm{Turns}_\infty(f)$.

Let Γ be a finite topological graph with at least one edge and let $f\colon \Gamma \to \Gamma$ be a topological graph-map. Choose an orientation on Γ and let $e_1, \ldots, e_k$ be an ordering of $E_+\Gamma$. The *transition matrix* $A(f) = (a_{ij})_{i,j=1}^k$ is a $k \times k$ matrix where for $1 \leq i, j \leq k$ the entry a_{ij} is the total number of occurrences of $e_j^{\pm 1}$ in the path $f(e_i)$ (note that this definition gives the transpose of that used for the transition matrix in [9]). Thus for every $n \geq 1$ we have $A(f^n) = A(f)^n$. Define $\lambda(f)$ to be the spectral radius of the transition matrix $A(f)$.

We say that $A(f)$ is *positive*, denoted $A(f) > 0$, if $a_{ij} > 0$ for all $1 \leq i, j \leq k$. We say that $A = A(f)$ is *irreducible* if for every $1 \leq i, j \leq k$ there exists $t = t(i, j) \geq 1$ such that $(A^t)_{ij} > 0$. Thus if $A(f) > 0$ then $A(f)$ is irreducible. Note that if $A(f)$ is irreducible then $\lambda(f)$ is exactly the Perron–Frobenius eigenvalue of $A(f)$. Furthermore, when $A(f)$ is irreducible, the property $\lambda(f) > 1$ is equivalent to f being expanding, which is also equivalent to $A(f)$ not being a permutation matrix.

Example 2.9 Consider again our running example graph-map $f\colon \Gamma \to \Gamma$ defined in Example 2.2. The four combinatorial edge-paths $f(a) = d$, $f(b) = a$, $f(c) = b^{-1}a$ and $f(d) = ba^{-1}db^{-1}ac$ are all reduced; therefore f is a regular graph-map. Hence

we also see that the links $\mathrm{Lk}_\Gamma(v)$ at the left and right vertices of Γ are respectively partitioned into the equivalence classes

$$\{a\}, \{b\}, \{d\} \quad \text{and} \quad \{a^{-1}\}, \{b^{-1}, c^{-1}\}, \{c\}, \{d^{-1}\}$$

by the relations $\sim_{f,v}$; thus f is tame. One may also calculate by hand the Whitehead graphs of f for the vertices of Γ, which we have illustrated in Figure 2. Namely, in the notations of Remark 2.8, Figure 2 shows all the nondegenerate turns that belong to $\mathrm{Turns}_3(f)$ is this example. A direct check shows that $\mathrm{Turns}_3(f) = \mathrm{Turns}_4(f)$ here and hence $\mathrm{Turns}_\infty(f) = \mathrm{Turns}_3(f)$. Therefore the graphs shown in Figure 2 are exactly the Whitehead graphs of the two vertices of Γ. Note that both these Whitehead graphs are connected.

Figure 2: The Whitehead graphs of the left and right vertices of Γ, respectively. We have labeled each adjacency by an edge-path that takes the corresponding turn in Γ.

Finally, the transition matrix, Perron–Frobenius eigenvalue and corresponding unit eigenvector for this example are

$$A(f) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 2 & 1 & 1 \end{pmatrix}, \quad \lambda \approx 2.4142, \quad \boldsymbol{v} \approx \begin{pmatrix} .2265 \\ .0939 \\ .1327 \\ .5469 \end{pmatrix}.$$

The characteristic polynomial of $A(f)$ is $x^4 - x^3 - 2x^2 - 3x - 1$. It is easy to check that $A(f)^3$ is positive, so $A(f)$ is irreducible. Since $\lambda(f) > 1$, we also see that f is expanding.

2.5 Entropy and expansion

The next two results are likely well known, but we could not find an explicit reference in the literature. In Appendix A, we have included proofs of these for completeness, as well as some references for closely related results.

The first result tells us that under fairly mild assumptions on a topological graph map, the topological entropy can be calculated by the transition matrix.

Proposition A.1 Let Γ be a finite, connected, topological graph with at least one edge and let $f\colon \Gamma \to \Gamma$ be regular topological graph-map. Let $h(f)$ be the topological entropy of f. Then $h(f) = \log \lambda(f)$ (recall that $\lambda(f)$ is the spectral radius of $A(f)$).

The next theorem gives a very strong rigidity result which, among other things, allows one to promote a topological graph map to a (linear) graph map, with respect to a carefully chosen linear structure, in certain situations. To state the theorem, we first make the following definition.

Definition 2.10 A topological graph-map $f\colon \Gamma \to \Gamma$ is *expanding on all scales* if for every $x \in \Gamma$, every neighborhood U of x, and every edge e of Γ, there exists a positive integer m and an open interval $W \subset U$ so that f^m maps W homeomorphically onto e.

Theorem A.3 Suppose $f\colon \Gamma \to \Gamma$ is a regular topological graph-map that is expanding on all scales. Choose an orientation on Γ and an ordering $E_+\Gamma = \{e_1, \ldots, e_n\}$.

Then the topological graph-map f is expanding, the matrix $A(f)$ is irreducible with the Perron–Frobenius eigenvalue $\lambda = \lambda(f) > 1$, and there exists a unique volume-1 metric structure $\mathcal{L}$ on Γ with the following properties:

(1) $f\colon \Gamma \to \Gamma$ is linear with respect to $\mathcal{L}$.

(2) f is a local λ–homothety, meaning that the norm of the derivative of f at every point of $\Gamma \setminus f^{-1}(V\Gamma)$ is λ.

(3)   For every $i = 1, \ldots, n$ the $d_{\mathcal{L}}$–length of e_i is v_i, where $\boldsymbol{v} = (v_1, \ldots, v_n)^T$ is the left Perron–Frobenius eigenvector (ie for the left action of $A(f)$ on column vectors) with $\sum_{i=1}^{n} v_i = 1$.

As we show in Appendix A, expanding irreducible graph maps are expanding on all scales; see Lemma A.2. Consequently, as explained Corollary A.9, given any expanding, irreducible graph map $f\colon \Gamma \to \Gamma$, there is a choice of (possibly different) metric structure on Γ making f into a combinatorial graph map (Definition 2.4).

2.6 Train-tracks

To obtain sufficient information to prove Theorems C and D we will make heavy use of a certain class of graph-maps introduced in [9] that have additional structure. This extra structure ensures that the map and all of its powers are combinatorially efficient.

Definition 2.11 (Train-track map) Let Γ be a finite connected (linear) graph with at least one edge and without valence-1 vertices. A (linear) graph-map $f\colon \Gamma \to \Gamma$ is a *train-track map* if:

(1) The map f is a homotopy equivalence (and hence f is surjective).

(2) For every edge $e \in E\Gamma$ and every $n \geq 1$ the map $f^n|_e$ is an immersion; that is, the edge-path $f^n(e)$ is tight and nondegenerate. (Note that this implies that for every $n \geq 1$ the map $f^n\colon \Gamma \to \Gamma$ is regular.)

The original definition of train-track map [9] does not allow valence 2 vertices. The next lemma guarantees that no complications occur near such vertices.

Lemma 2.12 *Suppose that $f\colon \Gamma \to \Gamma$ is a train-track map. Then f is tame. Furthermore, if v is a vertex of valence 2 in Γ then for every $n \geq 1$ the map f^n is an immersion near v; that is, for every $n \geq 1$ there exists an open neighborhood J of v in Γ such that $f^n|_J$ is injective.*

Proof By condition (2) of Definition 2.11, both conclusions of the lemma will follow if we can prove that there is a power $m_0 > 0$ so that for every vertex $v \in V\Gamma$, either $f^{m_0}(v) = v$ and f^{m_0} is injective on a neighborhood of v or else there is a turn of v taken by f^{m_0}.

Let $\mathcal{V} \subset V\Gamma$ be the set of f–*stable* vertices: that is, those vertices which are eventually permuted by f. Let $m > 0$ be such that $f^m|_{\mathcal{V}} = id_{\mathcal{V}}$, and so that $f^m(V\Gamma) = \mathcal{V}$. First observe that if $v \in V\Gamma \setminus \mathcal{V}$, then since f^m is surjective by (1), there must be a turn of v taken by f^m. It follows that any $v \in f^m(V\Gamma \setminus \mathcal{V}) \subset \mathcal{V}$ has a turn taken by f^{2m} (the image of a turn taken by f^m).

Finally, if $v \in \mathcal{V} \setminus f^m(V\Gamma \setminus \mathcal{V})$, then we claim that f^m is locally injective on some neighborhood of v or else there is a turn of v taken by f^m. To prove this claim, suppose that f^m is not locally injective on a neighborhood of v. Then the map $\mathrm{Lk}_\Gamma(v) \to \mathrm{Lk}_\Gamma(v)$ induced by f^m (defined since $f^m(v) = v$) would fail to be injective, and hence must also fail to be surjective, missing some $e \in \mathrm{Lk}_\Gamma(v)$. Surjectivity of f^m implies that there must be some edge e' so that e is an edge of $f^m(e')$. A point of $\overline{e}'$ sent to $o(e) = v$ cannot be a vertex (by assumption on v), and hence must be an interior point. But this means that v has a turn that is taken by f^m as required. Setting $m_0 = 2m$ completes the proof. $\qquad\square$

An *irreducible train track map* is a train-track map $f\colon \Gamma \to \Gamma$ such that the transition matrix $A(f)$ is irreducible. In this case, the Perron–Frobenius theory tells us that

$\lambda(f) \geq 1$ is an eigenvalue of $A(f)$, that there exists a unique (up to a scalar multiple) left eigenvector $\boldsymbol{v}$ for $A(f)$ with $A(f)\boldsymbol{v} = \lambda(f)\boldsymbol{v}$, and that all the coordinates of $\boldsymbol{v}$ are positive. We normalize $\boldsymbol{v}$ so that the sum of its coordinates is equal to 1. The coordinates of $\boldsymbol{v}$ then define a length function $E\Gamma \to \mathbb{R}_+$ which together with the linear structure induces a metric structure $\mathcal{L}$ on Γ so that $\mathrm{vol}_{\mathcal{L}}(\Gamma) = 1$ and such that for every edge $e \in E\Gamma$ we have $|f(e)|_{\mathcal{L}} = \lambda(f)\mathcal{L}(e) = \lambda(f)|e|_{\mathcal{L}}$. We call this $\mathcal{L}$ the *train-track metric structure* on Γ and the geodesic metric $d_{\mathcal{L}}$ the *train track-metric* on Γ.

Remark 2.13 An irreducible train-track map $f\colon \Gamma \to \Gamma$ typically need not be a local $\lambda(f)$–homothety with respect to the train-track metric $d_{\mathcal{L}}$. The reason is that Γ comes equipped with a linear structure Λ such that for each $e \in E\Gamma$ there exists a unique chart map $j_e\colon [0, \mathcal{L}(e)] \to \Gamma$ in Λ mapping $(0, \mathcal{L}(e))$ homeomorphically to e. If $f(e) = e_1 \cdots e_m$ then there is a unique subdivision $0 = t_0 < t_1 < \cdots < t_m = \mathcal{L}(e)$ such that $f \circ j_e$ maps (t_{i-1}, t_i) homeomorphically to e_i. Since the relative lengths of the subintervals (t_{i-1}, t_i) are specified by the linear structure, there is in general no reason to assume that we will have $\lambda(f)|t_i - t_{i-1}| = \mathcal{L}(e_i)$ for $i = 1, \ldots, m$. Thus, even though $\mathcal{L}(f(e)) = \lambda(f)\mathcal{L}(e)$, the map f may stretch the metric $d_{\mathcal{L}}$ on the subsegments $j_e([t_{i-1}, t_i])$ of e by different amounts. However, [Corollary A.7](#) shows that one may *change* the linear structure on Γ in such a way that f remains a linear graph-map and that, with respect to the new train track metric, f *is* locally a $\lambda(f)$–homothety on edges.

If $f\colon \Gamma \to \Gamma$ is a regular topological graph-map, we have a natural *derivative map* $D_f\colon E\Gamma \to E\Gamma$, where for $e \in E\Gamma$ we define $D_f(e)$ to be the initial edge of the edge-path $f(e)$. The map D_f naturally extends to $D_f\colon \mathrm{Turns}(\Gamma) \to \mathrm{Turns}(\Gamma)$ by setting $D_f(\{e, e'\}) := \{D_f(e), D_f(e')\}$, where e, e' is a turn in Γ. A turn $\boldsymbol{t} \in \mathrm{Turns}(\Gamma)$ is then called *legal* for f if $(D_f)^n(\boldsymbol{t})$ is nondegenerate for every $n \geq 0$.

Remark 2.14 Consider a finite graph Γ without valence-1 vertices and a homotopy equivalence regular (linear) graph-map $f\colon \Gamma \to \Gamma$. In this situation it is easy to see that f is a train-track map if and only if the unique nondegenerate turn at each valence-2 vertex is legal and every turn contained in an edge-path of the form $f(e)$ for $e \in E\Gamma$ is legal. Moreover, it is easy to check whether a given turn is legal because the set $\mathrm{Turns}(\Gamma)$ is finite. Therefore, it is straightforward and algorithmic to determine whether the given map f is a train track-map.

Example 2.15 The graph-map f introduced in [Example 2.2](#) is actually a train-track map. This fact can be easily checked by hand using the algorithm explained in

Remark 2.14 above, and we have also verified that f is a train-track map using Peter Brinkmann's software package xtrain.[1]

2.7 Outer automorphisms

For an integer $N \geq 2$ let $F_N = F(x_1, \ldots, x_N)$ be the free group of rank N. Recall that $\mathrm{Out}(F_N)$ is the group of all automorphisms of F_N modulo the subgroup of inner automorphisms

An element $\varphi \in \mathrm{Out}(F_N)$ is said to be *reducible* if there exists a free product decomposition $F_N = A_1 * \cdots * A_k * C$ with $k \geq 1$ and each $A_i \neq \{1\}$ (and where it is possible that $C = \{1\}$ only in the case $k \geq 2$) such that φ permutes the conjugacy classes $[A_1], \ldots, [A_k]$.

An element $\varphi \in \mathrm{Out}(F_N)$ is *irreducible* if it is not reducible. Moreover, $\varphi \in \mathrm{Out}(F_N)$ is said to be *fully irreducible* or *iwip* (for "irreducible with irreducible powers") if φ^n is irreducible for every $n \geq 1$. Thus φ is fully irreducible if and only if there do not exist $n \geq 1$ and a proper free factor A of F_N such that $\varphi^n([A]) = [A]$. Accordingly, $\Phi \in \mathrm{Aut}(F_N)$ is said to be *irreducible* (respectively, *fully irreducible*) if the outer automorphism class $\varphi \in \mathrm{Out}(F_N)$ of Φ is irreducible (respectively, fully irreducible).

Finally, $\varphi \in \mathrm{Out}(F_N)$ (or an element of $\mathrm{Aut}(F_N)$ representing φ) is called *atoroidal* or *hyperbolic* if there do not exist $n \geq 1$ and $w \in F_N, w \neq 1$, such that φ^n preserves the conjugacy class $[w]$ of w in F_N. A result of Brinkmann [13] (utilizing the Bestvina–Feighn combination theorem [7]) says that $\varphi \in \mathrm{Aut}(F_N)$ is hyperbolic if and only if the mapping torus group $G_\varphi = F_N \rtimes_\varphi \mathbb{Z}$ of φ is word-hyperbolic.

2.8 Topological representatives

Given a finite connected graph Γ and a homotopy equivalence topological graph-map $f \colon \Gamma \to \Gamma$, let $v \in V\Gamma$ be a vertex and let β be an edge-path from v to $f(v)$. We denote by $f_* = f_*^{\beta} \colon \pi_1(\Gamma, v) \to \pi_1(\Gamma, v)$ the map given by $f_*([\gamma]) = [\beta(f \circ \gamma)\beta^{-1}]$, where γ is an edge-path from v to v in Γ. We omit mention of β in the notation for f_* as we will be interested in the outer class of f_* which is independent of the choice of path β.

Definition 2.16 (Topological representative) Let $\varphi \in \mathrm{Out}(F_N)$, where $N \geq 2$. A *topological representative* of φ is a pair $(\mathfrak{m}, f)$ where $f \colon \Gamma \to \Gamma$ is a homotopy equivalence topological graph-map of a finite connected topological graph Γ to itself,

---

[1] Available at `http://gitorious.org/xtrain`

and where $\mathfrak{m}\colon F_N \to \pi_1(\Gamma, v)$ is an isomorphism such that $\mathfrak{m}^{-1} \circ f_* \circ \mathfrak{m}\colon F_N \to F_N$ is an automorphism of F_N whose outer automorphism class is φ (this condition is independent of the choice of edge-path β from v to $f(v)$). The isomorphism $\mathfrak{m}$ is called a *marking*. By convention, we will usually omit the explicit mention of marking and just say that $f\colon \Gamma \to \Gamma$ is a topological representative of φ.

If **P** is a type of topological graph-map (such as "regular" or "linear"), we say that a topological representative $f\colon \Gamma \to \Gamma$ *is* **P** if the topological graph-map f is of type **P**. Alternately we may say that $f\colon \Gamma \to \Gamma$ is a **P** *representative* of φ. Thus we may discuss regular representatives, tame representatives, irreducible train-track representatives, and so on. We emphasize that if f is a train-track representative of φ, then this automatically implies that f is a regular topological representative of φ and that $f^n\colon \Gamma \to \Gamma$ is a train-track representative of φ^n for all $n \geq 1$.

Given any $\varphi \in \mathrm{Out}(F_N)$, one can easily construct a regular topological representative for φ with Γ being a wedge of circles. Constructing representatives with more structure is in general harder and may depend on the outer automorphism φ. An important general result of Bestvina and Handel [9] says that for every irreducible $\varphi \in \mathrm{Out}(F_N)$ (where $N \geq 2$) there exists an irreducible train-track topological representative of φ. The converse is, in general, not true: it is possible for an irreducible train-track map to be a topological representative of a reducible $\varphi \in \mathrm{Out}(F_N)$. However, if $\varphi \in \mathrm{Out}(F_N)$ is fully irreducible then every train-track representative of φ has irreducible transition matrix and thus is an irreducible train-track map. We discuss the relationship between fully irreducible hyperbolic outer automorphisms and the properties of their train-track representatives in greater detail in the next subsection and in [Appendix B](#).

Example 2.17 We now return to the graph map $f\colon \Gamma \to \Gamma$ introduced in [Example 2.2](#). We have $\pi_1(\Gamma) \cong F_3$ and it is not hard to check directly that f is a homotopy equivalence; thus f is a topological representative of an element $\varphi \in \mathrm{Out}(F_3)$. Specifically, use $v_0 = o(c)$ as a base-vertex in Γ and set $\mathfrak{m}\colon F_3 = F(x_1, x_2, x_3) \to \pi_1(\Gamma, v_0)$ by $\mathfrak{m}(x_1) = b^{-1}a$, $\mathfrak{m}(x_2) = a^{-1}d$ and $\mathfrak{m}(x_3) = c$. Then the outer class φ is represented by $\varphi = \mathfrak{m}^{-1} \circ f_* \circ \mathfrak{m}$ where $f_*\colon \pi_1(\Gamma, v_0) \to \pi_1(\Gamma, v_0)$ (where β in the definition of f_* is the constant path from v_0 to itself). This automorphism of F_3 is given by $\phi(x_1) = x_2$, $\phi(x_2) = x_2^{-1}x_1^{-1}x_2x_1x_3$ and $\phi(x_3) = x_1$. We have moreover verified, with the help of Derek Holt and the computational software package kbmag[2], that the group G_φ is word-hyperbolic and therefore that the automorphism φ is hyperbolic. In [Example 2.19](#) we will see that φ is also fully irreducible.

---

[2]Available at http://www.gap-system.org/Packages/kbmag.html

2.9 Criterion for full irreducibility

We end this section with a criterion for a hyperbolic automorphism to be fully irreducible. Stating this criterion requires the notion of "clean train-tracks" from [39]:

Definition 2.18 (Clean train-track) We say that a train-track map $f\colon \Gamma \to \Gamma$ is *clean* if the Whitehead graph $\mathrm{Wh}(f, v)$ is connected for each vertex $v \in V\Gamma$, and if there exists $m > 0$ such that $A(f^m) > 0$. (This implies f is expanding, that $A(f^n) = A(f)^n$ is irreducible for all $n \geq 1$, and that $A(f^n) > 0$ for all $n \geq m$.) We say that f is *weakly clean* if f is expanding, if the Whitehead graph $\mathrm{Wh}(f, v)$ is connected for each vertex $v \in V\Gamma$ and if $A(f)$ is irreducible.

Our next theorem provides a criterion for full irreducibility in terms of the existence of weakly clean train-track representatives; this is a strengthening of the criterion proved in [39].

Theorem B.3 *Let $N \geq 3$ and let $\varphi \in \mathrm{Out}(F_N)$ be a hyperbolic element. Then the following conditions are equivalent:*

(1) *The automorphism φ is fully irreducible.*

(2) *If $f\colon \Gamma \to \Gamma$ is a train-track representative of φ then f is clean.*

(3) *There exists a clean train-track representative of φ.*

(4) *There exists a weakly clean train-track representative of φ.*

The proof of Theorem B.3 is given in Appendix B. This result seems to be of independent interest and also has some interesting consequences (eg Corollary B.4). However, as these considerations are tangential to the main results of the paper, we have relegated the proofs and discussion to an appendix.

Example 2.19 We have already seen that $f\colon \Gamma \to \Gamma$ in Example 2.2 is an expanding irreducible train-track map representing an automorphism $\varphi \in \mathrm{Out}(F_3)$. We also know that φ is hyperbolic – this fact was verified (see Example 2.17) by checking that the group G_φ is word-hyperbolic. We have additionally verified in Example 2.9 that the Whitehead graphs of both vertices of Γ with respect to f are connected. Thus f is a weakly clean train-track map. Therefore, by Theorem B.3, φ is fully irreducible.

3 Folding paths

We will need a continuous version of Stallings' folding line adapted to our setting (see [41; 59] for the classic version of Stallings folds). Our version of the folding line is more elementary and less canonical than what Francaviglia and Martino call "fast folding line" (see [28, Definition 5.6]) and what Bestvina and Feighn call "folding all illegal turns at speed 1" (see the definition immediately prior to Proposition 2.2 in [8]). While we could have used the "fast folding line" when defining the "folded mapping torus" X_f in Section 4 below, the more basic Stallings folds construction is sufficient for our purposes and simpler to analyze.

Convention 3.1 In this section let $\varphi \in \mathrm{Out}(F_N)$, let Γ be a finite connected linear graph and let $f\colon \Gamma \to \Gamma$ be a linear graph-map that is a tame topological representative of φ (see Definitions 2.1, 2.3, 2.6, and 2.16). Additionally let $\mathcal{L}$ be a metric structure on Γ compatible with the given linear graph structure.

Note that the tameness assumption implies every vertex of Γ has valence ≥ 2 and that $f(e)$ is a tight nondegenerate edge-path in Γ for every edge $e \in E\Gamma$.

Remark 3.2 Given any $\varphi \in \mathrm{Out}(F_N)$ it is easy to construct a tame topological representative of φ. This can be done, for example, via replacing an arbitrary regular topological representative $f'\colon \Gamma \to \Gamma$ of φ by an "optimal" map $f\colon \Gamma \to \Gamma$ that is freely homotopic to f, as explained in [28]. Alternatively, given $\varphi \in \mathrm{Out}(F_N)$, one can always find, via an explicit combinatorial construction, a tame representative $f\colon \Gamma \to \Gamma$ of φ with Γ being a wedge of N circles, as explained in [42]. (Also, Lemma 2.12 implies that if $f\colon \Gamma \to \Gamma$ is an expanding irreducible train-track map, then f is tame.) As indicated after Definition 2.1, such a topological representative may then be adjusted by a homotopy rel $V\Gamma$ to yield a linear graph-map (as required by Convention 3.1) with respect to any given linear structure on Γ

While our constructions of folding paths (Section 3) and associated folded mapping tori (Section 4) apply to any map f as in Convention 3.1, working in this level of generality introduces technicalities that significantly complicate the notation. To enhance readability, we therefore only describe the constructions in the following setting:

Convention 3.3 For expository purposes only, we additionally suppose that $f\colon \Gamma \to \Gamma$ is a combinatorial graph-map with respect to the given metric structure $\mathcal{L}$ (see Definition 2.4).

Remark 3.4 We emphasize that this choice to work with combinatorial graph-maps only serves to simplify the notation and that it is straightforward to adjust the constructions in Sections 3–4 to handle the general setting of Convention 3.1. In fact, the curious reader may find these details in Sections 4 and 6 of a previous version of this paper [20] where we work in the more general situation.

Moreover, this additional requirement that $f: \Gamma \to \Gamma$ be a combinatorial graph-map is not a serious constraint: Remark 2.5 explains that any regular topological graph-map is homotopic to a combinatorial graph-map (with respect to an appropriate metric structure), and Corollary A.9 shows that oftentimes the homotopy is unnecessary.

Now we form a new labeled metric graph Δ as follows. We have $\Delta = \Gamma$ as a set and as a topological space. The vertex set of Δ is $V\Delta := f^{-1}(V\Gamma) \cup V\Gamma$ and the 1–cells of Δ are the connected components of $\Delta - V\Delta$. Thus, combinatorially, Δ is obtained from Γ by subdividing each edge e of Γ into $|f(e)|$ edges. The graph Δ inherits a natural linear graph structure from Γ and we still have the graph-map $f: \Delta \to \Gamma$ which is now an affine homeomorphism on every open 1–cell. We endow Δ with a metric structure compatible with the linear structure in which every edge has length 1. Note that this metric structure is not the one that the subdivided graph Δ naturally inherits from Γ (recall that Definition 2.4 requires Γ to be a metric graph). For every edge $e \in E\Delta$ and every vertex $v \in V\Delta$ we think of $f(e) \in E\Gamma$ as the *label* of e and of $f(v) \in V\Gamma$ as the *label* of v. This labeling satisfies the property that $f(e^{-1}) = (f(e))^{-1}$ and that $o(f(e)) = f(o(e))$, $t(f(e)) = f(t(e))$ for all $e \in E\Delta$. In general, we call a (linear) graph Ψ with a labeling $\tau: V\Psi \to V\Gamma$, $\tau: E\Psi \to E\Gamma$ of vertices and edges, satisfying these properties, a Γ–*graph*. Any Γ–graph Ψ comes equipped with a canonical graph-map $\tau: \Psi \to \Gamma$, given by the labeling, which is an affine homeomorphism on open edges, and a metric compatible with the linear structure giving each edge length 1.

Since $\tau: \Psi \to \Gamma$ sends an edge to an edge, τ is a regular graph-map. A Γ–graph Ψ is *folded* if whenever $e, e' \in E\Psi$ are such that $o(e) = o(e')$ and $\tau(e) = \tau(e')$ then $e = e'$. It is easy to see that for a Γ–graph Ψ the map $\tau: \Psi \to \Gamma$ is locally injective if and only if Ψ is folded. If a Γ–graph Ψ is not folded and $e \neq e'$ are elements of $E\Psi$ such that $o(e) = o(e')$ and $\tau(e) = \tau(e')$, then identifying e and e' into a single edge with label $\tau(e) = \tau(e')$ produces a new Γ–graph Ψ'. We say that Ψ' is obtained from Ψ by a *combinatorial Stallings fold*.

We say that a Γ–graph Ψ is *tame* if the labeling graph-map $\tau: \Psi \to \Gamma$ is tame. Note that for a Γ–graph Ψ and a vertex v of Ψ, the number of equivalence classes in the partition of $\mathrm{Lk}_\Gamma(v)$ by the relation $\sim_{\tau,v}$ is equal to the number of distinct labels of edges of Ψ with origin v. Thus a Γ–graph Ψ is tame if and only if for every vertex v

of Ψ there exist some two edges $e, e' \in E\Psi$ with $o(e) = o(e') = v$ and $\tau(e) \neq \tau(e')$. Hence if Ψ is a tame Γ–graph and Ψ' is obtained from Ψ by a combinatorial Stallings fold then Ψ' is again tame.

We now return to the construction of a folding path. According to [59], we can obtain Γ from Δ by a finite sequence of combinatorial Stallings folds

$$(3\text{-}1) \qquad \Delta = \Delta_0, \quad \Delta_1, \quad \ldots, \quad \Delta_m = \Gamma,$$

where $m = \#E_{\text{top}}\Delta - \#E_{\text{top}}\Gamma$ and where each Γ–graph Δ_{i+1} is obtained from Δ_i by a single combinatorial Stallings fold identifying the edges e_i, e_i' of Δ_i. We endow each Δ_i with its natural Γ–graph metric and linear structure as described above.

Note that since by assumption, the map $f\colon \Gamma \to \Gamma$ is tame, the Γ–graph $\Delta = \Delta_0$ is tame. Since Stallings folds preserve tameness, it follows that each Δ_i is tame. In particular, this means that every vertex of Δ_i has valence ≥ 2.

We now define a 1–parameter family of metric graphs $\{\Gamma_t\}_{t \in [0,1]}$ as follows. For each $j = 0, \ldots, m$, set $\Gamma_{j/m} = \Delta_j$. For any $j = 0, \ldots, m-1$ and arbitrary $t \in [j/m, (j+1)/m]$, let Γ_t be obtained from $\Delta_j = \Gamma_{j/m}$ by isometrically identifying initial segments of e_j and e_j' of length $tm - j \in [0, 1]$. We may endow Γ_t with a metric and associated linear structure so that the quotient map $\Gamma_{j/m} \to \Gamma_t$ is a local isometry on each edge, and we do so.

Note that for each $j = 0, \ldots, m-1$, the two definitions of $\Gamma_{(j+1)/m}$ as Δ_{j+1} and as a quotient of $\Gamma_{j/m} = \Delta_j$ agree by definition of combinatorial Stallings fold. Furthermore, we note that Γ_0 is obtained from Γ by subdividing and changing the metric, but as topological spaces, $\Gamma_0 = \Gamma = \Gamma_1$.

Remark 3.5 We have noted above that each Δ_i is tame and thus has no valence-1 vertices. Therefore for every $0 \leq t \leq 1$ the graph Γ_t has no valence-1 vertices either, since the new vertices that appear in the middle of performing a single Stallings fold have valence 3.

Each Γ_t comes equipped with two natural homotopy equivalence graph maps

$$r_{t,0}\colon \Gamma_0 \longrightarrow \Gamma_t \quad \text{and} \quad r_{1,t}\colon \Gamma_t \longrightarrow \Gamma_1$$

which are locally isometric on edges and such that

$$(3\text{-}2) \qquad f = r_{1,t} \circ r_{t,0} = r_{1,0}\colon \Gamma_0 \longrightarrow \Gamma_1,$$

where we have used the identification of $\Gamma_0 = \Gamma = \Gamma_1$ as topological spaces to make sense of this equality. In fact, for any $0 \leq t \leq t' \leq 1$, there is a canonical homotopy

equivalence map $r_{t',t}\colon \Gamma_t \to \Gamma_{t'}$ which is a local isometry on edges and so that for all $0 \le t \le t' \le t'' \le 1$ we have

(3-3) $$r_{t'',t'} \circ r_{t',t} = r_{t'',t}.$$

We refer to the family $\{\Gamma_t\}_{t \in [0,1]}$, together with the maps $\{r_{t',t}\}_{0 \le t \le t' \le 1}$ defined above, as the *topological folding line* corresponding to the topological representative $f\colon \Gamma \to \Gamma$ of φ. We write $\{(\Gamma_t, r_{t',t})\}_{\{0 \le t \le t' \le 1\}}$ to denote the topological folding line.

This folding line also depends on the choice of a sequence of combinatorial Stallings folds (3-1) taking Δ to Γ, but this dependence will not affect the proofs of our main results (though the objects themselves do depend on these choices, for notational simplicity we have suppressed this dependence from the notation in the following).

Example 3.6 Let $f\colon \Gamma \to \Gamma$ be the graph-map from Example 2.2. Figure 3 illustrates the first two combinatorial Stallings folds in the folding sequence taking the Γ–graph Δ to Γ for the graph map $f\colon \Gamma \to \Gamma$.
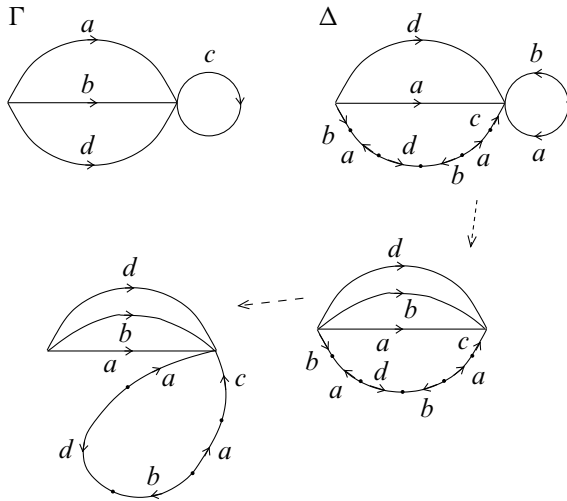

Figure 3: Two combinatorial Stallings folds used to construct a topological folding line

Remark 3.7 Technically, in the standard Outer space terminology, the "folding path" Γ_t is a family of "marked metric graphs" which connects the "marked metric graph" $(\mathfrak{m}, \Gamma_0)$ (where $\mathfrak{m}\colon F_N \to \pi_1(\Gamma_0)$ is the marking given by the fact that f is a topological representative of φ and that Γ_0 is a subdivision of Γ) to the "marked metric graph"

$(f_* \circ \mathfrak{m}, \Gamma)$. However, the markings on the graphs Γ_t are not important for the specific applications used in this paper and therefore we did not record them in our definition.

For the purposes of constructing the "folded mapping torus" of f in Section 4 below, we will mainly forget about the metric structures $\mathcal{L}_t$ on the graphs Γ_t and only look at the underlying 1–complexes Γ_t and the topological maps $\{r_{t',t}\}$ provided by the above construction. The metric structures on Γ_t will occasionally be useful, and, in particular, they will make it simpler to describe the cell structure on the "folded mapping torus" of f defined in the next section.

4 The folded mapping torus and semiflows.

Convention 4.1 Let $\varphi \in \mathrm{Out}(F_N)$ and set $G = G_\varphi$. Fix a linear graph-map $f\colon \Gamma \to \Gamma$ that is a tame topological representative φ, as in Convention 3.1, and a topological folding line $\{\Gamma_t, r_{t',t})\}_{\{0 \leq t \leq t' \leq 1\}}$. We also continue to identify $\Gamma_0 = \Gamma = \Gamma_1$ as topological spaces (though Γ_0 is a subdivision of $\Gamma = \Gamma_1$ as a topological graph). For expository purposes (see Remark 3.4), we continue to assume that f is a combinatorial graph-map as in Convention 3.3.

In this section, given the above data, we describe a cell complex we call the *folded mapping torus* of f, denoted X_f, which is a particular $K(G, 1)$ built from the data of the folding line. We also construct a semiflow on X_f that encodes the folding line, and prove a few basic properties of X_f. We end by describing the cell structure in a little more detail and defining the E_0–class $\epsilon \in H^1(X_f; \mathbb{Z})$ to be used in Theorem B.

Remark 4.2 The construction of another $K(G, 1)$ 2–complex from $\varphi \in \mathrm{Out}(F_N)$ — Gautero's "dynamical 2–complex" — and a semiflow is carried out by Gautero in [32, Proposition 3.9, Proposition 4.3] using a sequence of Whitehead moves in place of the folding path; see also the modification in [35]. Similarly, Wang's construction [63] of the decomposed mapping torus is similar to our construction via "building blocks" as explained in the proof of Proposition 4.3 below.

4.1 The folded mapping torus

Starting with the folding line $\{(\Gamma_t, r_{t',t})\}_{\{0 \leq t \leq t' \leq 1\}}$, we define an equivalence relation on $\Gamma \times [0, 1]$ by declaring (x, t) to be equivalent to (x', t') if and only if $t = t'$ and $r_{t,0}(x) = r_{t,0}(x')$. We denote the associated quotient space by

$$\Pi'\colon \Gamma \times [0, 1] \longrightarrow X'.$$

By construction, we can identify $\Pi'(\Gamma \times \{t\})$ with $r_{t,0}(\Gamma_0) = \Gamma_t$ as *sets* so that

(4-1) $$\Pi'(x, t) = r_{t,0}(x),$$

and we do so. In fact, this identification is much more than simply a setwise identification.

Proposition 4.3 *The space X' is homeomorphic to a cell-complex. The identification of the subspace $\Pi'(\Gamma \times \{t\})$ with Γ_t is a homeomorphism, where $\Pi'(\Gamma \times \{t\})$ is given the subspace topology.*

Proof Recall that for each $j = 0, \ldots, m-1$, $\Gamma_{(j+1)/m} = \Delta_{j+1}$ is obtained from $\Gamma_{j/m} = \Delta_j$ by a single combinatorial Stallings fold identifying the edges e_j, e'_j from $\Gamma_{j/m}$. We will use this to reconstruct X' as a cell complex obtained by gluing together very simple cell complexes.

For each $j = 0, \ldots, m-1$, begin with the product cell complex $\Gamma_{j/m} \times [j/m, (j+1)/m]$. The (closed) 2–cells are all thus "rectangles" obtained as products of (closed) 1–cells in $\Gamma_{j/m}$ with the interval $[j/m, (j+1)/m]$. Now we subdivide the rectangles $R_j = \bar{e}_j \times [j/m, (j+1)/m]$ and $R'_j = \bar{e}'_j \times [j/m, (j+1)/m]$ each into two triangles. More precisely, taking (orientation preserving) locally isometric characteristic maps for these 2–cells

$$\rho_j : [0, 1] \times [j/m, (j+1)/m] \longrightarrow R_j = \bar{e}_j \times [j/m, (j+1)/m],$$
$$\rho'_j : [0, 1] \times [j/m, (j+1)/m] \longrightarrow R'_j = \bar{e}'_j \times [j/m, (j+1)/m],$$

we subdivide $[0, 1] \times [j/m, (j+1)/m]$ into two triangles along the diagonal from the point $(0, j/m)$ to $(1, (j+1)/m)$. Denote the upper triangle by τ_+ and the lower triangle by τ_-. Now subdivide $\Gamma_{j/m} \times [j/m, (j+1)/m]$ by subdividing each of $\bar{e}_j \times [j/m, (j+1)/m]$ and $\bar{e}'_j \times [j/m, (j+1)/m]$ into the triangles $\rho_j(\tau_+) \cup \rho_j(\tau_-)$ and $\rho'_j(\tau_+) \cup \rho'_j(\tau_-)$, respectively; see Figure 4.
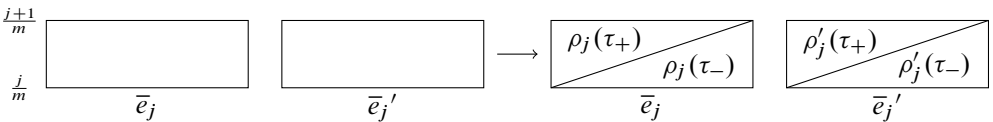

Figure 4: On the left we have the rectangles R_j, R'_j in $\Gamma_{j/m} \times [j/m, (j+1)/m]$ that are to be subdivided, and the right shows their subdivision

Now we let $\Pi'_j : \Gamma_{j/m} \times [j/m, (j+1)/m] \to X'_j$ be the quotient space obtained by identifying $\rho_j(\tau_+)$ with $\rho'_j(\tau_+)$ via the restriction to $\rho_j(\tau_+)$ of $\rho'_j \circ \rho_j^{-1}$; see Figure 5.
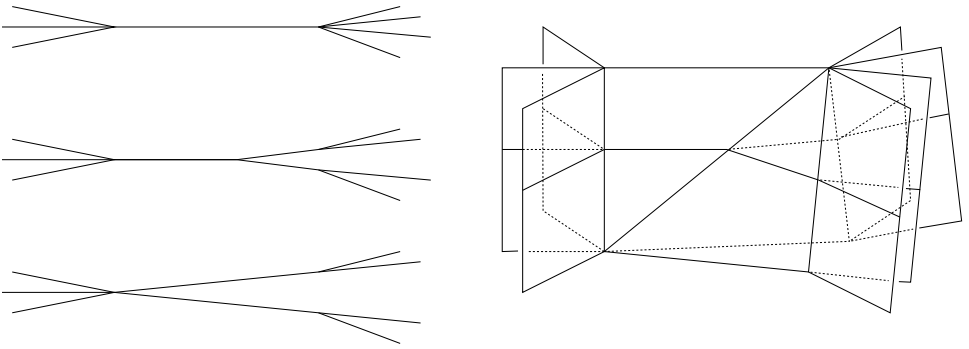
Figure 5: A local picture of X'_j near the identified triangles. On the left we have pictures of $\Gamma_{j/m}$, $\Gamma_{(2j+1)/2m}$, and $\Gamma_{(j+1)/m}$ on the bottom, middle, and top, respectively, near the folded edges. On the right we have the corresponding part of X'_j.

One minor complication is that even restricted to τ_+, the map ρ_j may not be invertible. This can only occur if e_j is a loop edge. Since $\Gamma_{j/m} \to \Gamma_{(j+1)/m}$ is a homotopy equivalence, it follows that both e_j and e'_j cannot both be loop edges (if they were, then $e'_j e_j^{-1}$ is a nonnull homotopic loop in $\Gamma_{j/m}$ sent to a null homotopic loop in $\Gamma_{(j+1)/m}$). So, without loss of generality, we assume that e_j is not a loop edge and hence we may invert ρ_j as needed; see Figure 6.
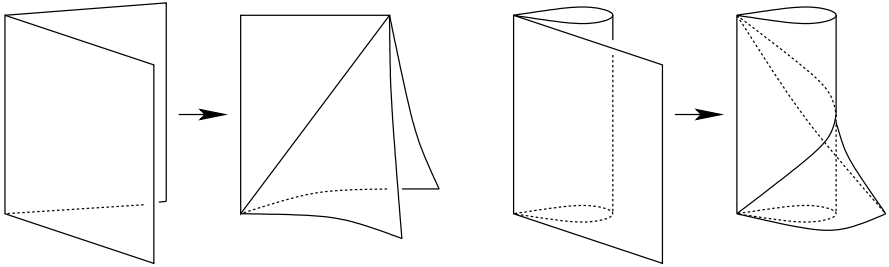

Figure 6: Two possibilities for Π'_j near $R_j \cup R'_j$ depending on whether e'_j is a loop (right) or not a loop (left)

We see that by construction, X'_j has a natural cell structure so that Π'_j is a cellular map. Furthermore, we note that for all $t \in [j/m, (j+1)/m]$, the restriction of Π'_j to $\Gamma_{j/m} \times \{t\}$ isometrically identifies the initial segments of $e_j \times \{t\}$ and $e'_j \times \{t\}$ of length $tm - j$. It is now easy to see that we may identify $\Pi'_j(\Gamma_{j/m} \times \{t\})$ with Γ_t, for all $t \in [j/m, (j+1)/m]$, *as topological spaces*. Using this identification we construct a quotient space X'' from the disjoint union $X'_0 \sqcup \cdots \sqcup X'_{m-1}$ by gluing the "top" $\Gamma_{(j+1)/m} \subset X'_j$ to the "bottom" $\Gamma_{(j+1)/m} \subset X'_{j+1}$, for each $j = 0, \ldots, m-1$.

The space X'' has a natural cell structure so that each X'_j includes into X'' by a cellular embedding.

The maps

$$\Gamma_0 \times [j/m, (j+1)/m] \xrightarrow{r_{j/m,0} \times \text{id}} \Gamma_{j/m} \times [j/m, (j+1)/m] \xrightarrow{\Pi'_j} X'_j$$

piece together to induce a map $\Pi'' \colon \Gamma \times [0, 1] \to X''$. For each $j = 0, \ldots, m-1$ and $j/m \leq t < (j+1)/m$, the restriction of Π'' to each $\Gamma \times \{t\}$ identifies (x, t) to (x', t) if and only if $r_{t,j/m}(r_{j/m,0}(x)) = r_{t,j/m}(r_{j/m,0}(x'))$. By (3-3), $r_{t,j/m} \circ r_{j/m,0} = r_{t,0}$, and it follows that the fibers of Π' and Π'' are the same. Since $\Gamma \times [0, 1]$ and X'' are compact Hausdorff spaces, X'' is homeomorphic to the quotient space X' (and identifying $X' = X''$, we have $\Pi' = \Pi''$). Furthermore, the homeomorphic identification of $\Pi'(\Gamma \times \{t\}) = \Pi''(\Gamma \times \{t\})$ with Γ_t proves the second statement. □

The spaces X'_j in the construction of $X'' = X'$ are useful in analyzing the local pictures of X', and we will refer to these as *building blocks*.

Definition 4.4 (Folded mapping torus) The *folded mapping torus* $X = X_f$ is the quotient

$$X' \longrightarrow X_f$$

obtained by gluing $\Gamma_1 \subset X'$ to $\Gamma_0 \subset X'$ using the identifications $\Gamma_0 = \Gamma = \Gamma_1$. Subdividing Γ_1, the gluing becomes cellular, and we can induce on X_f a cell structure from the one constructed in the proof of Proposition 4.3 for which the 2–cells are either *rectangles* or *triangles*. We call this the *initial cell structure* on X_f (a more useful cell structure will be defined in Section 4.4).

The folded mapping torus X_f is closely related to the mapping torus of f, which is the space

$$Z_f = \Gamma \times [0, 1]/ \sim,$$

where $(x, 1) \sim (f(x), 0)$.

Proposition 4.5 *The map $\Pi' \colon \Gamma \times [0, 1] \to X'$ descends to a continuous map*

$$\Pi \colon Z_f \longrightarrow X_f.$$

Proof Since $(x, 1)$ is identified with $(f(x), 0)$ in Z_f (and these are the only identifications), we need only verify that $\Pi'(x, 1)$ is identified with $\Pi'(f(x), 0)$ in X_f.

Using the identification $\Gamma_0 = \Gamma = \Gamma_1$ together with (3-2) and (4-1) we have

$$\Pi'(x, 1) = r_{1,0}(x) = f(x) \in \Gamma_1 = \Gamma,$$
$$\Pi'(f(x), 0) = r_{0,0}(f(x)) = f(x) \in \Gamma_0 = \Gamma.$$

Since $f(x)$ in Γ_1 is identified with $f(x)$ in Γ_0, we are done. □

In fact, it is not difficult to prove that Π is a homotopy equivalence, but this will follow naturally from the construction of certain semiflows in the next section; see Theorem 4.7.

4.2 Semiflows

Here we construct and analyze (local) semiflows — that is, (local) actions of the additive semigroup $\mathbb{R}_+$ of positive real numbers — on the spaces constructed in Section 4.1. First, let $\widehat{\psi}'$ be the local flow on $\Gamma \times [0, 1]$ defined by $\widehat{\psi}'_s(x, t) = (x, t + s)$, wherever this makes sense.

Proposition 4.6 *For all $0 \le s, t \le s + t \le 1$ we have $\Pi'(\widehat{\psi}'_s(x, t)) = r_{s+t,t}(\Pi'(x, t))$.*

Proof We simply appeal to (3-2) and (4-1) then calculate

$$\Pi'(\widehat{\psi}'_s(x, t)) = \Pi'(x, t + s) = r_{t+s,0}(x) = r_{t+s,t}(r_{t,0}(x)) = r_{t+s,t}(\Pi'(x, t)). \quad \square$$

From this it follows that the maps $r_{t',t}$ define a *local semiflow* ψ' on X' given by $\psi'_s(\Pi'(x, t)) = r_{t+s,t}(\Pi'(x, t))$, wherever this is defined. Furthermore, Π' semiconjugates $\widehat{\psi}'$ to ψ':

$$\Pi' \circ \widehat{\psi}_s = \psi'_s \circ \Pi'.$$

We note that flowlines of $\widehat{\psi}'$ only terminate at $\Gamma \times \{1\}$, and so $\widehat{\psi}'$ can be extended to a semiflow $\widehat{\psi}$ on Z_f (it is only guaranteed to be a semiflow since f is not in a homeomorphism in general). In fact, $\widehat{\psi}$ is given by

$$\widehat{\psi}_s(x, t) = \left(f^{[s+t]}(x), \ s + t - [s + t]\right)$$

for all $(x, t) \in \Gamma \times [0, 1) \subset Z_f$, where $[s + t]$ is the greatest integer less than or equal to $s + t$. Since the surjection Π' semiconjugates $\widehat{\psi}'$ to ψ', it follows that Π semiconjugates $\widehat{\psi}$ to a semiflow ψ on X_f. On the graph $\Gamma = \Gamma_0 \subset X' \to X_f$, this semiflow is given by

$$\psi_s(x) = r_{s-[s],0} \circ f^{[s]}(x) \in \Gamma_{s-[s]}.$$

In particular, ψ_1 restricted to Γ_0 is precisely the map f. Consequently, we call ψ the *suspension semiflow on X_f*.

As a first application, we prove the following theorem.

Theorem 4.7 *The map* $\Pi\colon Z_f \to X_f$ *is a homotopy equivalence.*

Proof Construct a map $F'\colon X' \to \Gamma \times [0,1]$, $F'(\Pi'(x,t)) = (r_{1,t}(\Pi'(x,t)),t)$. By (3-2) and (4-1) we have

$$F'(\Pi'(x,t)) = F'(r_{t,0}(x)) = (r_{1,t}(r_{t,0}(x)),t)) = (f(x),t).$$

It follows that F' descends to a map $F\colon X_f \to Z_f$. Moreover, this equation implies $F \circ \Pi = \widehat{\psi}_1$.

On the other hand, for $t \in [0,1)$ and $x \in \Gamma$, we have

$$\Pi(F(\Pi(x,t))) = \Pi(\widehat{\psi}_1(x,t)) = \psi_1(\Pi(x,t)).$$

Consequently, $\Pi \circ F = \psi_1$. Therefore, F is a homotopy inverse to Π with the semiflows $\widehat{\psi}$ and ψ providing the homotopies $F \circ \Pi \simeq \mathrm{id}_{Z_f}$ and $\Pi \circ F \simeq \mathrm{id}_{X_f}$, respectively. $\qquad\square$

Recall that $G = G_\varphi = \langle F_N, r \mid r^{-1}wr = \Phi(w), w \in F_N \rangle$, where $\Phi \in \mathrm{Aut}(F_N)$ is a representative of the outer automorphism class φ. The projection $\Gamma \times [0,1] \to [0,1]$ descends to a surjection $\eta'_0\colon Z_f \to \mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$. Similarly, the projection $X' \to [0,1]$, given by $\Pi'(x,t) \mapsto t$, descends to a map

$$\eta_0\colon X_f \longrightarrow \mathbb{S}^1 = \mathbb{R}/\mathbb{Z}.$$

For every $t \in \mathbb{R}/\mathbb{Z}$, the fiber $\eta_0^{-1}(t)$ is precisely Γ_t (where we take a representative $t \in [0,1]$ for the point of $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$).

An application of van Kampen's theorem implies that the isomorphism $F_N \to \pi_1(\Gamma)$ from the marking extends to an isomorphism $\Psi\colon G \to \pi_1(Z_f)$ for which

$$(\eta'_0)_*(\Psi(r)) = 1 \in \mathbb{Z} = \pi_1(\mathbb{S}^1).$$

Thus, $u_0\colon G \to \mathbb{Z}$ defined by $u_0(r^n w) = n$ for all $w \in F_N$ satisfies $u_0 = (\eta'_0)_* \circ \Psi$.

Convention 4.8 For the remainder of this paper, we use the isomorphism

$$\Pi_* \circ \Psi\colon G \longrightarrow \pi_1(X_f)$$

to identify G with $\pi_1(X_f)$. With this convention, and the fact that $\eta'_0 = \eta_0 \circ \Pi$ we see that $(\eta_0)_* = u_0$.

Since Z_f is a $K(G,1)$ for $G = G_\varphi$ (see eg [38, Theorem 1.B 11]), as a consequence of Theorem 4.7 we have:

Corollary 4.9 *The folded mapping torus X_f is a $K(G, 1)$ space.*

Example 4.10 Returning to our running example $f\colon \Gamma \to \Gamma$, Figure 7 shows the folded mapping torus X_f in this case. We have represented it as four large rectangles with a subdivided cell structure, and with identifications as indicated. See the caption of the figure for detailed description of the identifications.
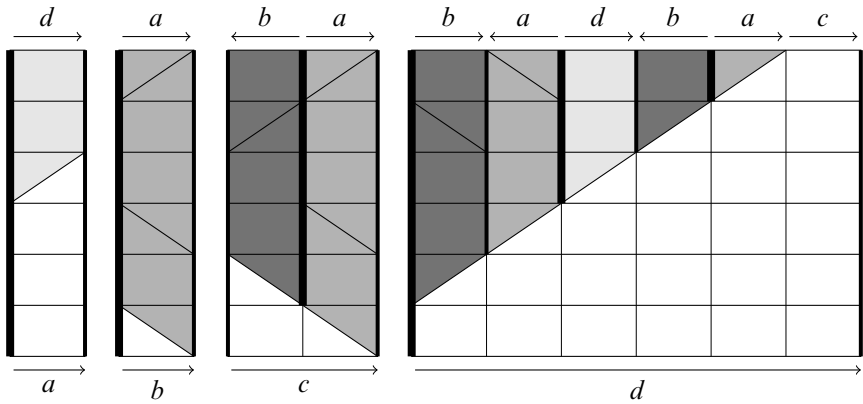

Figure 7: The folded mapping torus X_f for the example $f\colon \Gamma \to \Gamma$. The identifications on the sides of the large rectangles are indicated by thick and medium-thick lines, and these are by isometries. This identification gives the product $\Gamma \times [0, 1]$. The tops of the rectangles are identified to the bottoms as indicated by the labeled arrows, and these identifications are isometries for a and b, and local homotheties by 2 for c and by 6 for d. The result of these identifications (together with the side identifications) gives the mapping torus Z_f. Finally, the 2–cells are identified as indicated by the shading and the patterns of 2–cells (white cells are not identified to any others) to produce X_f. We have also indicated which 1–cells in the interiors of the rectangles are identified with those coming from the boundaries of the rectangles via thick and medium-thick lines. We note that performing the side identifications and the 2–cell identifications only (which can be done by isometries) gives the complex X'.

4.3 Horizontal, vertical, and edge degree

Definition 4.11 We intuitively define the *degree* $d(e)$ of a 1–cell e of the initial cell structure on X_f to be the number of 2–cells attached to e, counted with multiplicity. More precisely, we note that the attaching maps for the 2–cells are locally injective, and for every 1–cell e, we pick a point x in the interior of e, and $d(e)$ is the total number

of points in the preimage, over all attaching maps (local injectivity makes this definition independent of x). Equivalently, $d(e)$ is the minimum number of components of $U - e$, over all arbitrarily small neighborhoods U of x. For example, any point in an edge of degree 2 has a neighborhood homeomorphic to $\mathbb{R}^2$.

Definition 4.12 The *horizontal foliation* of X_f is the decomposition of X_f into fibers Γ_t, for $t \in \mathbb{S}^1$. We refer to any subset of a single fiber Γ_t as being horizontal. We refer to any arc of a flow line as a *vertical* arc. More generally, any set of points that all flow forward to a common point will be called vertical.

The complement of the closure of the 1–cells of degree different than 2 is a topological surface. We note that the restriction of the horizontal foliation to this surface is a foliation in the traditional sense. Likewise, the restriction of the semiflow to this surface is a local flow.

The 1–cells fall into three types: vertical, horizontal and skew.

Definition 4.13 The *vertical* 1–cells are 1–cells contained in semiflow lines, while the *horizontal* 1–cells are contained in fibers of η_0. The skew 1–cells are the m remaining 1–cells which occur as the (nonhorizontal/nonvertical) side of some triangle.

Note that the horizontal 1–cells all have degree 2 and the skew 1–cells all have degree 3. On the other hand, the vertical 1–cells have more general degrees.

Lemma 4.14 If e is a vertical or skew 1–cell and $v \in V\Gamma_t$ is the vertex defined by $v = \Gamma_t \cap e$, then $d(e) = \mathrm{val}(v)$, where $\mathrm{val}(v)$ is the valence of v in Γ_t. In particular, $d(e) \geq 2$ for every vertical 1–cell.

Proof First suppose e is a vertical 1–cell, and observe that $t \notin \{j/m\}_{j=0}^m$ because $e = \Pi(v_0 \times (j/m, (j+1)/m))$ for some vertex $v_0 \in V\Gamma_{j/m}$ for some j, and $v = \Gamma_t \cap e \neq \varnothing$. The fact that $d(e) = \mathrm{val}(v)$ follows from the fact that one has a basis of neighborhoods of v in X of the form $U = W \times (t - \epsilon, t + \epsilon)$ for small $\epsilon > 0$ and W a neighborhood of v in Γ_t. The minimal number of components of $W - v$ is equal to the minimal number of components of $U - e$ over all sufficiently small neighborhoods W of v in Γ_t, and this is equal to $\mathrm{val}(v)$, so $d(e) \leq \mathrm{val}(v)$. Any sufficiently small neighborhood about v in X is contained in some basis open set U, and from here one easily deduces $d(e) = \mathrm{val}(v)$.

Since the intersection of a skew 1–cell e with Γ_t corresponds to the "new" vertex of a fold, and this vertex always has $\mathrm{val}(v) = 3$ (and as we've already noted, $d(e) = 3$), the first statement for e follows.

For the second statement, we note that our folding lines never produce graphs with valence 1 vertices, by Remark 3.5. $\qquad\square$

4.4 The trapezoid structure.

To carry out the constructions in Section 6.6, it is desirable to modify the cell structure so that there are no horizontal 1–cells while maintaining a simple structure for the 2–cells. Simply erasing the horizontal 1–cells is insufficient, since the complement of the (closure of the) vertical and skew 1–cells need not be a union of disks (although it is a surface). We therefore proceed as follows.

For every edge e of $\Gamma_0 \subset X_f$, we consider the image under the semiflow for time up to 1:

$$R_e = \bigcup_{0 \leq s \leq 1} \psi_s(\overline{e}).$$

We call this set the *strip above* e. The complex X is a union of the strips above edges e of Γ_0. Some strips contain a skew edge and others do not. The example in Figure 7 has 10 strips, 9 of which contain a skew edge, and one does not.

For each strip that does not contain a skew 1–cell, we subdivide the first rectangle of the initial cell structure

$$\overline{e} \times [0, 1/m] = \bigcup_{0 \leq s \leq 1/m} \psi_s(\overline{e})$$

into two triangles by adding one of the diagonals (it does not matter which diagonal we add). The added diagonal becomes a 1–cell which we also call a skew 1–cell. The new skew 1–cells are distinguished from the previous ones by their degree: the new skew 1–cells all have degree 2, while the old ones have degree 3.

For any skew 1–cell e, consider the set of points $T_e \subset X - e$ that flow into e before meeting any other skew 1–cell. That is $y \in T_e$ if and only if there exists $r > 0$ so that $\psi_r(y) \in e$, and for all $0 \leq s < r$, $\psi_s(y)$ is not contained in any skew 1–cell. We now describe T_e in more detail.

Suppose first that $d(e) = 2$. Below e is a unique triangle 2–cell, and call it U_1. The triangle U_1 lies below e and above a horizontal 1–cell δ_1. The 1–cell δ_1 is contained in the top arc of a 2–cell σ_2 (it may be a proper subarc of the top of this 2–cell only if $\delta_1 \subset \Gamma_0 = \Gamma_1$). Let U_2 denote the subset of σ_2 directly below δ_1 (that is, the subset of this 2–cell that flows into δ_1). If σ_2 is a triangle, then the bottom of U_2 is an arc of a skew 1–cell, and hence $T_e = U_1 \cup \delta_1 \cup U_2$. If σ_2 is a rectangle, then U_2 lies above an arc δ_2 of a horizontal 1–cell e_2, and δ_2 is a subarc of a 2–cell σ_3. Let U_3 denote the subset of σ_3 below δ_2. If σ_3 is a triangle 2–cell, then

$$T_e = U_1 \cup \delta_1 \cup U_2 \cup \delta_2 \cup U_3.$$

Otherwise, σ_3 is a rectangle, and we repeat again. In this way, we eventually write T_e as

$$T_e = U_1 \cup \delta_1 \cup U_2 \cup \delta_2 \cup \cdots \cup \delta_k \cup U_k$$

for some $k \geq 1$. See Figure 8. This procedure terminates because every strip contains a skew 1–cell, hence a triangle (which explains why the new skew 1–cells were added).

If $d(e) = 3$, then below e there are exactly two triangle 2–cells. The above description can be carried out for each of the two triangles thus producing two trapezoids $T_e = T'_e \sqcup T''_e$, each of which can be described as

$$U_1 \cup \delta_1 \cup U_2 \cup \delta_2 \cup \cdots \cup \delta_k \cup U_k$$

for some $k \geq 1$.

Definition 4.15 A *trapezoid* for X is a disk T_e when $d(e) = 2$ or one of T'_e or T''_e when $d(e) = 3$.

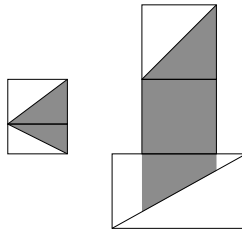Figure 8: Two possible trapezoids

Let $\mathcal{T}$ denote the set of all trapezoids. The closure of any $T \in \mathcal{T}$ is a union

$$\overline{T} = T \cup e_+(T) \cup \ell_+(T) \cup \ell_-(T) \cup e_-(T),$$

where $e_+(T) = \overline{e}$ and e is the initial skew 1–cell defining the trapezoid, $e_-(T)$ is the arc of a skew 1–cell below T, and $\ell_\pm(T)$ are the arcs of flow lines bounding T on either side (one of which may degenerate to a single point). Although T is an open disk, $\overline{T}$ need not be a closed disk (for example, it can happen that $e_-(T) \subset e_+(T)$; see Figure 9).

Because every flow line eventually intersects some skew 1–cell (again because of the addition of the new skew 1–cells), it follows that X is the union

$$X = \bigcup_{T \in \mathcal{T}} \overline{T}.$$

Since $T \cap T' = \varnothing$ for all $T \neq T' \in \mathcal{T}$ (since the first skew 1–cell a point meets under ψ is unique), we can define a new cell structure for which $\mathcal{T}$ is the set of 2–cells.

Definition 4.16 (Trapezoid cell structure) The *trapezoid cell structure* is the cell structure for which $\mathcal{T}$ is the set of 2–cells, and the vertices are precisely the union of the (three or) four corners of $\bar{T}$ over all $T \in \mathcal{T}$. We give the 1–skeleton $X_f^{(1)}$ a metric structure so that the restriction of η_0 to each edge is a local isometry to $\mathbb{S}^1$, and an orientation by insisting that these maps be orientation-preserving.

Convention 4.17 In all of what follows, we will assume that *the* cell structure on X is the trapezoid cell structure, unless otherwise stated. We refer to the initial cell structure *with the added diagonals* as the *old cell structure*. We also refer to the 1–cells of the trapezoid cell structure as either vertical or skew.

Remark 4.18 We note that every skew 1–cell for the trapezoid cell structure is contained in a skew 1–cell of the old cell structure, but that this is not true for the vertical 1–cells. However, any vertical 1–cell of the trapezoid structure eventually flows into a vertical 1–cell of the old cell structure. Since the union of the vertical 1–cells of the old structure is invariant under ψ (because $f(V\Gamma) \subset V\Gamma$), so is the union of the vertical 1–cells of the trapezoid structure.

Example 4.19 The trapezoid cell structure for our running example $f \colon \Gamma \to \Gamma$ is shown in Figure 9. There are fourteen 0–cells, twenty-seven 1–cells and thirteen 2–cells.
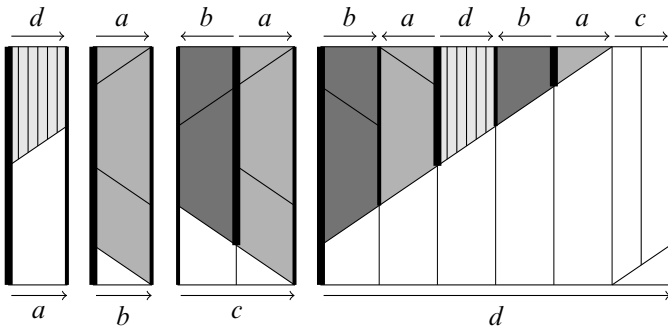

Figure 9: The trapezoid cell structure for the folded mapping torus X_f for the example $f \colon \Gamma \to \Gamma$

4.5 Trapezoidal subdivisions

In Section 6 we will construct "perturbed" fibrations $\eta_z \colon X_f \to \mathbb{S}^1$ associated to certain 1–cocycles z on X_f. To aid in the construction of η_z, which will be carried

out explicitly in Section 6.6, it will be helpful to use a subdivided cell structure on X_f that is tailored to the cocycle z. Here we describe the general type of cell structures that we will use.

For a trapezoidal 2–cell T of X_f, recall that $\ell_-(T)$, $\ell_+(T)$, $e_-(T)$, and $e_+(T)$ denote the left, right, bottom and top arcs of T, respectively. As each of these arcs is in fact a union of 1–cells of X_f, we may regard them as cellular 1–chains in X_f. The trapezoid T induces a *T–orientation* on these arcs as follows: By definition, $\ell_\pm(T)$ and $e_-(T)$ are given the positive orientation (as specified in Definition 4.16) so that the 1–cells comprising $\ell_\pm(T)$ and $e_-(T)$ must appear with positive coefficient in those 1–chains. The two sides $\ell_\pm(T)$ of T are then distinguished by the convention that $e_-(T)$ is oriented from $\ell_-(T)$ to $\ell_+(T)$. Lastly the T–orientation on the top arc $e_+(T)$ is defined so that it is also oriented from $\ell_-(T)$ to $\ell_+(T)$. Alternatively, using ψ to flow the bottom to the top defines a map $h_T\colon e_-(T) \to e_+(T)$ which is a homeomorphism on the interior of $e_-(T)$; the T–orientation on $e_+(T)$ is defined such that h_T is orientation preserving. As the T–orientation on $e_+(T)$ may not agree with the positive orientation on $X_f^{(1)}$, we see that the 1–cells comprising the 1–chain $e_+(T)$ may appear with positive or negative coefficient in this 1–chain. Since $e_+(T)$ is a sum of 1–cells all contained in a single skew 1–cell of the old cell structure, it follows that all coefficients of $e_+(T)$ have the *same* sign, and we define the *sign of T* to be $\zeta(T) \in \{\pm 1\}$ such that the cellular 1–chain $\zeta(T)e_+(T)$ has all positive coefficients. With these definitions, the boundary of T is the cellular 1–chain

$$(4\text{-}2) \qquad \partial T = e_-(T) + \ell_+(T) - e_+(T) - \ell_-(T).$$

Note that X_f may contain some trapezoids T for which one of the sides $\ell_\pm(T)$ degenerates to a single point (see the left trapezoid of Figure 8). Such trapezoids are said to be *degenerate*.

Remark 4.20 If T is a degenerate trapezoid, our naming conventions ensure that $\ell_+(T)$ is the degenerate side of T; that is, $\ell_+(T)$ is a point whereas $\ell_-(T)$ is a nondegenerate arc. Moreover, in this case the positive orientation on the skew 1–cells comprising $e_+(T)$ disagrees with the T–orientation on the arc $e_+(T)$. Thus we necessarily have $\zeta(T) = -1$ for every degenerate trapezoid.

Recall that the cell structure on X_f is constructed so that every 1–cell is either vertical or skew. The sides $\ell_\pm(T)$ of each trapezoid T consist of vertical 1–cells, and the top $e_+(T)$ consists of (possibly several) skew 1–cells. On the other hand, the bottom $e_-(T)$ of each trapezoid consists of a *single* skew 1–cell. Conversely, every skew 1–cell of X_f is equal to $e_-(T)$ for a unique trapezoid T. The subdivisions of X_f that

we construct in Section 6.3 below will again be *trapezoidal cell structures*, meaning that they have exactly this structure. Furthermore, each 2–cell T of the subdivision will be a trapezoid that is contained in a unique 2–cell T' of X_f. In fact T will be a subtrapezoid of T' having top and bottom arcs, again denoted $e_\pm(T)$, that are subarcs of $e_\pm(T')$ and with side arcs, again denoted $\ell_\pm(T)$, being arcs of flowlines running all the way from the bottom to the top of T'. Such a subdivision will be called a *trapezoidal subdivision*. As with X_f, the 1–cells are all skew or vertical and hence inherit orientations making the restriction of the fibration η_0 to each 1–cell orientation preserving. We use the same conventions for the T–orientation on the sides of a trapezoid T of the subdivision and define the sign $\zeta(T)$ just as for trapezoids of X_f. With these definitions we have $\zeta(T) = \zeta(T')$, where T' is the unique 2–cell of X_f containing T.

4.6 The semiflow on trapezoids

As described above, each trapezoid is built from a union of subrectangles of rectangles, two subtriangles of triangles, and the horizontal arcs in between:

$$T = U_1 \cup \delta_1 \cup U_2 \cup \delta_2 \cdots \cup \delta_k \cup U_k.$$

As T is a disk, ψ actually restricts to a local flow on T. This local flow has a very simple description; see Figure 10.

Proposition 4.21 *For every trapezoid $T \in \mathcal{T}$ there are real numbers $m_1, m_2, t_0 \in \mathbb{R}$ defining a (possibly degenerate) Euclidean trapezoid*

$$\widehat{T} = \{(r,t) \in \mathbb{R}^2 \mid 0 \le r \le 1, \ m_1 r \le t \le m_2 r + t_0\}$$

and a continuous map

$$\sigma_T \colon \widehat{T} \longrightarrow \overline{T}$$

such that σ_T sends horizontal lines to horizontal arcs, and for all $(r,t) \in \widehat{T}$ and $0 \le s \le m_2 r + t_0 - t$ we have

$$\sigma_T(r, t+s) = \psi_s(\sigma_T(r,t)).$$

Moreover, σ_T restricted to the interior of $\widehat{T}$ is a homeomorphism onto T.

Proof Since Π' semiconjugates $\widehat{\psi}'$ to ψ', and since ψ' generates ψ, it follows that the map $\widehat{\Pi} \colon \Gamma \times [0,1] \to X$ obtained by composing $\Pi' \colon \Gamma \times [0,1] \to X'$ with the quotient map $X' \to X$ satisfies

$$(x, t+s) \mapsto \psi_s(\widehat{\Pi}(x,t)).$$

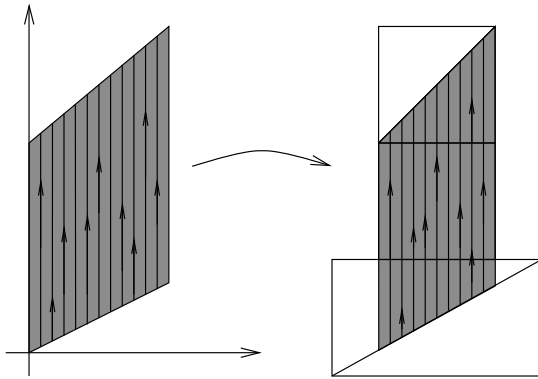Figure 10: The local flow preserving map $\sigma_T \colon \widehat{T} \to \overline{T}$

Since T is a union of images of product rectangles and triangles in $\Gamma \times [0, 1]$ under $\widehat{\Pi}$, it is straight forward to piece together the local flows $(x, t) \mapsto (x, t + s)$ on these and construct the required $\widehat{T}$ and map σ_T. □

Convention 4.22 We use the maps σ_T from Proposition 4.21 to give the interior of every trapezoid a smooth structure. Note that the local flow $\psi|_T$ on any trapezoid T is thus smooth, as is $\eta_0|_T$.

5 Local models for the semiflow

The simplest behavior of ψ is in the interior of a trapezoid where Proposition 4.21 gives a completely transparent picture of what ψ looks like, namely it is a local flow on a trapezoid by translation parallel to the direction of the vertical sides. Our next goal is to complete this description by giving "local models" for the semiflow near any point of the 1–skeleton of X_f.

Remark 5.1 The discussion in this section should be compared with [32, Section 2.3]. In [63], similar issues are addressed using work of Bestvina and Brady [6].

Definition 5.2 (Local model) A *local model* for ψ is a subset of X_f of the form

$$\mathcal{M}(K, s_0) = \bigcup_{0 \leq s \leq s_0} \psi_s(K),$$

where $s_0 > 0$ and K is a closed contractible neighborhood of a point of Γ_t, for some t. We call K the *bottom* of $\mathcal{M}$ and $\psi_{s_0}(K)$ the *top* and we call $\mathcal{M}$ minus the top and bottom the *flow-interior* of $\mathcal{M}$.

About every point of X_f we will construct a local model $\mathcal{M} = \mathcal{M}(K, s_0)$ and study the maps

$$\psi_s \colon \psi_t(K) \longrightarrow \psi_{t+s}(K)$$

induced by the flow on the horizontal pieces $\psi_t(K)$ and $\psi_{t+s}(K)$, for $0 \leq t < t+s \leq s_0$. As these maps will enjoy several useful properties, it is convenient to introduce the following terminology.

Definition 5.3 (Control rel boundary) Given two finite graphs Q_0 and Q_1 which are nondegenerate (ie have no zero-dimensional components), let ∂Q_0 and ∂Q_1 denote the set of valence 1 vertices. We will say that a continuous map $\Upsilon \colon Q_0 \to Q_1$ is *controlled rel boundary* if the following holds:

(1) For every $x \in Q_1$, $\Upsilon^{-1}(x)$ is finite.

(2) Υ restricts to a bijection from ∂Q_0 to ∂Q_1.

(3) $\Upsilon \colon (Q_0, \partial Q_0) \to (Q_1, \partial Q_1)$ is a relative homotopy equivalence. That is, there exists a continuous map $\Upsilon_0 \colon (Q_1, \partial Q_1) \to (Q_0, \partial Q_0)$ so that

$$\Upsilon_0 \circ \Upsilon \colon Q_0 \longrightarrow Q_0 \quad \text{and} \quad \Upsilon \circ \Upsilon_0 \colon Q_1 \longrightarrow Q_1$$

are homotopic rel ∂Q_0 and ∂Q_1, respectively, to the identity.

Note that $\Upsilon_0|_{\partial Q_1} = (\Upsilon|_{\partial Q_0})^{-1}$.

This notion gives rise to a useful local definition for a map between graphs (cf [32, Definition 2.11]).

Definition 5.4 (Local control) If Q_0 and Q_1 are graphs, then a map $\Upsilon \colon Q_0 \to Q_1$ is *locally controlled* if there exists subdivisions Q_0' of Q_0 and Q_1' of Q_1 with the following properties.

(1) Q_0' and Q_1' are each a finite union of nondegenerate, connected subgraphs:

$$Q_0' = Q_0^1 \cup \cdots \cup Q_0^r, \ Q_1' = Q_1^1 \cup \cdots \cup Q_1^r.$$

(2) For each $i = 0, 1$ and every $j \neq j'$, $Q_i^j \cap Q_i^{j'} \subset \partial Q_i^j \cup \partial Q_i^{j'}$.

(3) For each $j = 1, \ldots, r$, $\Upsilon(Q_0^j) = Q_1^j$ and the restriction $\Upsilon|_{Q_0^j} \colon Q_0^j \to Q_1^j$ is controlled rel boundary.

Proposition 5.5 *If Q_0, Q_1 are finite graphs and $\Upsilon \colon Q_0 \to Q_1$ is locally controlled, then Υ is a homotopy equivalence and for every $x \in Q_1$, $\Upsilon^{-1}(x)$ is finite.*

Proof Since $\Upsilon|_{Q_0^j}$ is locally controlled rel boundary for all j, it follows that

$$\Upsilon\left(\bigcup_{j=1}^{r} \partial Q_0^j\right) = \bigcup_{j=1}^{r} \partial Q_1^j$$

and the restriction of Υ to $\bigcup \partial Q_0^j$ is a bijection onto $\bigcup \partial Q_1^j$ (since this is true for the restriction $\Upsilon|_{\partial Q_0^j}$ for each j). By definition of controlled rel boundary, there exist r continuous maps

$$\Upsilon_0^j \colon Q_1^j \longrightarrow Q_0^j$$

such that $\Upsilon_0^j \circ \Upsilon|_{Q_0^j} \colon Q_0^j \to Q_0^j$ and $\Upsilon|_{Q_0^j} \circ \Upsilon_0^j \colon Q_1^j \to Q_1^j$ are homotopic relative to ∂Q_0^j and ∂Q_1^j, respectively, to the identity. Since

$$\Upsilon_0^j|_{\partial Q_1^j} = (\Upsilon|_{\partial Q_0^j})^{-1}$$

it follows that Υ_0^j and $\Upsilon_0^{j'}$ agree on $Q_1^j \cap Q_1^{j'}$ for all j, j'. Therefore, the Υ_0^j are the restrictions of a continuous map

$$\Upsilon_0 \colon Q_1 \longrightarrow Q_0.$$

The various homotopies to the identities glue together to define homotopies of $\Upsilon_0 \circ \Upsilon \colon Q_0 \to Q_0$ and $\Upsilon \circ \Upsilon_0 \colon Q_1 \to Q_1$ to the respective identities. Thus, Υ is a homotopy equivalence.

For any $x \in Q_1$, we have $x \in Q_1^j$ for some j and hence $\Upsilon^{-1}(x) = \Upsilon|_{Q_0^j}^{-1}(x)$ is finite. $\qquad\square$

Proposition 5.6 *For every open set $U \subset X_f$ and every point $x \in U$, there exists a local model neighborhood $\mathcal{M}(K, s_0) \subset U$ containing x in its interior such that*

$$\psi_s \colon \psi_t(K) \longrightarrow \psi_{t+s}(K)$$

is controlled rel boundary for all $0 \leq s, t \leq s + t \leq s_0$.

Proof Recall that $0 < 1/m < \cdots < (m-1)/m < 1$ are the times at which the Stallings folds $\Delta_j = \Gamma_{j/m}$ occur during the folding process. Let $t' \in [0, 1)$ be such that $x \in \Gamma_{t'}$ and $t' \in [(j-1)/m, j/m)$ for some $j = 1, \ldots, m$. Similarly choose $t'' \in (0, 1]$ so that $x \in \Gamma_{t''}$ and $t'' \in ((i-1)/m, i/m]$ for some $i = 1, \ldots, m$. Thus $t' = t''$ unless $x \in \Gamma_0 = \Gamma = \Gamma_1$, in which case we have $t' = 0$ and $t'' = 1$. The proposition essentially follows from the fact that for small $s > 0$ the maps $\psi_s \colon \Gamma_{t'} \to \Gamma_{t'+s}$ and $\psi_s \colon \Gamma_{t''-s} \to \Gamma_{t''}$ are just the folding maps $r_{t'+s,t'}$ and $r_{t'',t''-s}$, respectively.

Note that while $\Gamma_{t'} = \Gamma_{t''}$ as sets and topological spaces, the graph structure on $\Gamma_{t'}$ may be a combinatorial subdivision of that in $\Gamma_{t''}$ (when $t' = 0$ and $t'' = 1$). In any case, we may choose a closed contractible neighborhood K of x in $\Gamma_{t'} = \Gamma_{t''}$ so that ∂K is disjoint from $V\Gamma_{t'}$ and $V\Gamma_{t''}$ and $K \subset U$. We let $D > 0$ be any number less than the minimal distance from ∂K to $V\Gamma_{t'} \cup V\Gamma_{t''}$ in either of the metrics on $\Gamma_{t'}$ and $\Gamma_{t''}$. Fix any $0 < \delta < \min\{D/m, |j/m - t'|, |t'' - (i-1)/m|\}$. In particular

$$[t', t' + \delta] \subset [(j-1)/m, j/m) \quad \text{and} \quad [t'' - \delta, t''] \subset ((i-1)/m, i/m].$$

For any $t' \leq s \leq t' + \delta$. The map $r_{s,t'}$ identifies initial segments of a pair of edges $e_1, e_2 \subset \Gamma_{t'}$ of length $(s - t')m$ and is a homeomorphism in the complement of these segments. Set $v = o(e_1) = o(e_2)$ of $\Gamma_{t'}$ and $K_s = r_{s,t'}(K)$. For any $t' \leq s \leq s' \leq t' + \delta$ we claim that $r_{s',s}|_{K_s} \colon K_s \to K_{s'}$ is controlled rel boundary. To see this, first note that $r_{s',t'} = r_{s',s} \circ r_{s,t'}$ is a homeomorphism outside the $(s' - t')m \leq m\delta < D$ neighborhood of v. Since the distance between v and ∂K is less than D in $\Gamma_{t'}$, if $v \notin K$, then K is outside the D–neighborhood of v and $r_{s',t'}|_K$ is a homeomorphism. Consequently, $r_{s',s}|_{K_s}$ is also a homeomorphism. Otherwise the D neighborhood of v is entirely contained in K. In this case, $r_{s',s}|_{K_s}$ simply identifies initial segments of length $(s' - s)m$ of a pair of edges, and is nevertheless controlled rel boundary.

Now suppose $t'' - \delta \leq s \leq t''$ and write $K_s = (r_{t'',s})^{-1}(K) \subset \Gamma_s$. The map $r_{s,t''-\delta}$ identifies initial segments of length $(s - t'' + \delta)m \leq m\delta < D$ of a pair of edges $e_1, e_2 \subset \Gamma_{t''-\delta}$ and is a homeomorphism outside the $m\delta < D$ neighborhood of $v = o(e_1) = o(e_2)$. For each $i = 1, 2$, let $\gamma_i \subset e_i$ be the initial segment of length $m\delta$, and let $u_i \in \gamma_i$ be the endpoint different than v. Then $r_{t'',t''-\delta}(\gamma_1) = r_{t'',t''-\delta}(\gamma_2)$ is a path in $\Gamma_{t''}$ of length $m\delta < D$ and $u = r_{t'',t''-\delta}(u_1) = r_{t'',t''-\delta}(u_2) \in V\Gamma_{t''}$. Since the distance from u to ∂K in $\Gamma_{t''}$ is at most D, either $r_{t'',t''-\delta}(\gamma_1) = r_{t'',t''-\delta}(\gamma_2) \subset K$ or $r_{t'',t''-\delta}(\gamma_1) = r_{t'',t''-\delta}(\gamma_2)$ is disjoint from K. In the former case, $\gamma_1 \cup \gamma_2 \subset K_{t''-\delta}$ and $r_{t'',t''-\delta}|_{K_{t''-\delta}}$ just folds γ_1 and γ_2 together and is homeomorphism outside these segments. In the latter case, $r_{t'',t''-\delta}|_{K_{t''-\delta}}$ is a homeomorphism onto K. In either case, for any $t'' - \delta \leq s \leq s' \leq t''$ we may argue as in the situation of $t' \leq s \leq s' \leq t' + \delta$ replacing K with $K_{t''-\delta}$ to prove that the map $r_{s',s}|_{K_s}$ is controlled rel boundary.

Therefore, by choosing δ sufficiently small and setting $s_0 = 2\delta$ and $K' = K_{t''-\delta}$, we see that $\mathcal{M}(K', s_0)$ is a local model contained in U satisfying the conclusion of the proposition. $\qquad \square$

5.1 Examples of local models

Local models will play an important role in the sequel. In an effort to familiarize the reader with these tools, and to clarify the proof of Proposition 5.6, in this subsection we

will describe and illustrate some specific types of local models. In all figures throughout this subsection, we have indicated 1–cells of X_f with thick lines. A single arrow indicates a vertical 1–cell and double arrow indicates a skew 1–cell.

The simplest type of local models are homeomorphic to surfaces, in which case $K \cong [-1, 1]$ and $K \times [0, s_0] \cong \mathcal{M}(K, s_0)$ via a homeomorphism $(r, t) \mapsto \psi_t(r)$. We call these *surface local models*. These exist for any point in the interior of a trapezoid (for example, by Proposition 4.21), but they also exist near 1–cells with degree 2, and 0–cells for which all adjacent 1–cells are degree 2. In this case, note that the semiflow restricts to a local flow, and the maps

$$\psi_s \colon \psi_t(K) \longrightarrow \psi_{t+s}(K)$$

are all homeomorphisms (in particular, they are controlled rel boundary).

However, local models near the 1–cells e with $d(e) > 2$ are necessarily more complicated. It is thus useful to distinguish these, and we do so with the following definition.

Definition 5.7 Let E_0 denote the set of 1–cells e with $d(e) > 2$. Let $\mathcal{E}_0 \subset X_f^{(1)}$ denote the subgraph of the 1–skeleton with edge set E_0 and vertex set $\{o(e), t(e) \mid e \in E_0\}$. We give $\mathcal{E}_0$ the restricted metric structure and orientation from that of $X_f^{(1)}$; see Definition 4.16.

Let us next consider the local model about a point x in a 1–cell e with $d(e) > 2$. When e is a vertical 1–cell there is a local model for which the semiflow restricts to a local flow. Indeed, if $x = \Gamma_t \cap e$, then for nearby $t' < t < t''$ with $t'' - t'$ sufficiently small, we will have that $\Gamma_s \cap e$ is a vertex of Γ_s of degree $d(e)$ for all $s \in [t', t'']$. Subdividing each edge of $\Gamma_{t'}$ adjacent to $x' = \Gamma_{t'} \cap e$ by adding a point very close to x', and letting K be the star of x', we see that the flow defines a homeomorphism $K \times [t', t''] \to \mathcal{M}(K, t'' - t')$ via $(y, s) \mapsto \psi_{s-t'}(y)$. See Figure 11.

When e is a skew 1–cell with $d(e) = 3$ there is a local model neighborhood as shown at the right of Figure 11. Observe that any such x is contained in Γ_t for some $i/m < t < (i+1)/m$ and some building block X_i' from the proof of Proposition 4.3. In particular, the local models are obtained from the explicit construction of the building blocks in the proof of Proposition 4.3. Here we note that the maps $\psi_s \colon \psi_t(K) \to \psi_{t+s}(K)$ are not homeomorphisms, although they are controlled rel boundary.

Finally we consider local models near 0–cells, for which there is a variety of possible behavior. To simplify the discussion, let us assume v is a 0–cell that does not lie in the interior of an old skew 1–cell and does not have a local model homeomorphic to a surface. For such a point, we will actually illustrate two local models, one above v and one below v. By combining these, one obtains what we call a *stacked local model*.
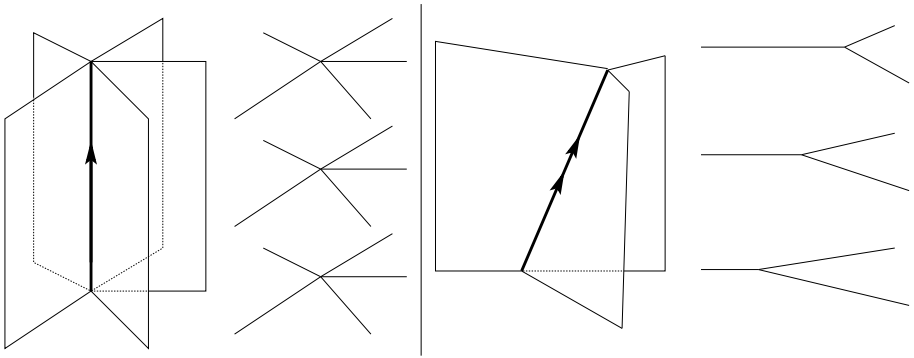
Figure 11: A local model for a vertical 1–cell e with $d(e) = 5$ (left) and a skew 1–cell (right). In both cases, $\psi_s \colon \psi_t(K) \to \psi_{t+s}(K)$ is controlled rel boundary, but only on the left is it a homeomorphism, and these are indicated to the right of each local model.

Under our assumptions, v is a 0–cell of the old cell structure, and is thus a vertex of valence at least 3 in $\Gamma_{i/m}$ for some $i = 0, \ldots, m-1$ (if v had valence 2, it would have a surface local model neighborhood). We construct a local model above v by subdividing each edge of $\Gamma_{i/m}$ adjacent to v and taking K to be the closed star of v. Taking $s_0 \in (0, 1/m)$, we obtain a local model $\mathcal{M}(K, s_0)$ for v that is in fact contained in the image of the building block X_i'.

There are essentially two possibilities for $\mathcal{M}(K, s_0)$. If $v = o(e) = o(e')$ for two distinct edges e, e' which are folded together in $\Gamma_{(i+1)/m}$, then this folding produces a skew 1–cell e_1 with $o(e_1) = v$ and $d(e_1) = 3$, and a vertical 1–cell e_2 with $o(e_2) = v$ with $d(e_2) = \mathrm{val}(v) - 1$, where $\mathrm{val}(v)$ is the valence of v in $\Gamma_{i/m}$. We have illustrated $\mathcal{M}(K, s_0)$ in Figure 12(b) (for $\mathrm{val}(v) = 5$) and (c) (for $\mathrm{val}(v) = 3$). If v is not the initial vertex of a pair of edges folded together in $\Gamma_{(i+1)/m}$, then (for s_0 sufficiently small), $\mathcal{M}(K, s_0)$ is a product. See Figure 12(a) (for $\mathrm{val}(v) = 5$).

A local model below v is described similarly. If $v \in \Gamma_0$, then $v \in \Gamma_1$ as well, so we may assume that $v \in \Gamma_{i/m}$ with $i = 1, \ldots, m$ is a vertex with $\mathrm{val}(v) \geq 3$. We again take the star K' of v in $\Gamma_{i/m}$ (after subdividing), and set $K = (\psi_{s_0})^{-1}(K')$ for a small $s_0 \in (0, 1/m)$. Then

$$\mathcal{M}(K, s_0) = \{(\psi_s)^{-1}(K') \mid 0 \leq s \leq s_0\}$$

is a local model below v. Again there are essentially two possibilities, which are illustrated in Figure 12(d)–(f) (for $\mathrm{val}(v) = 5$ or 3). Note that in all cases, the local models have the property that the flow maps $\psi_s \colon \psi_t(K) \to \psi_{t+s}(K)$ are controlled rel boundary.
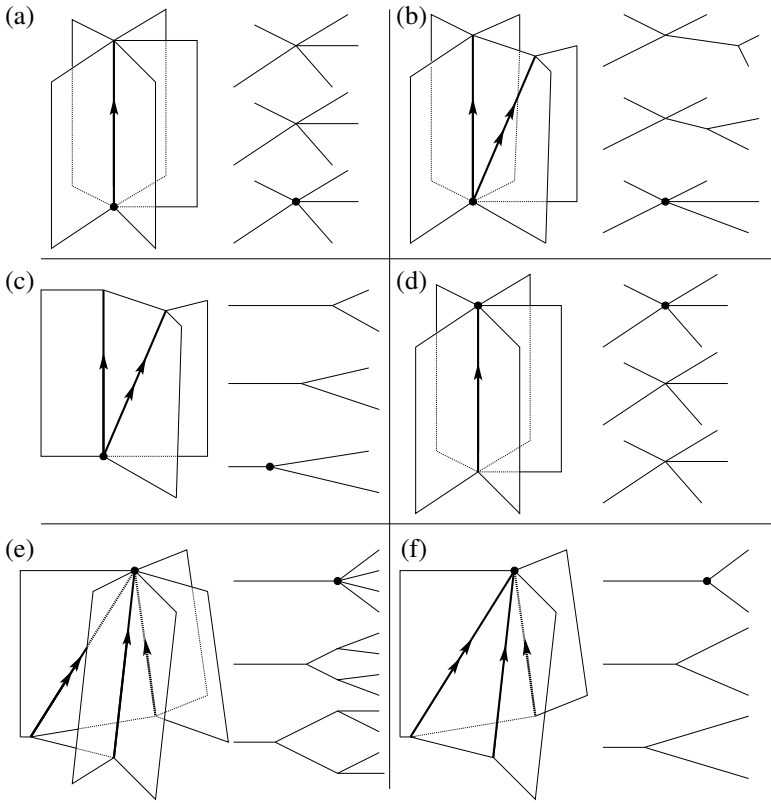
Figure 12: Pictures of local models above the 0–cell ((a)–(c)) and below the 0–cell ((d)–(e)). The maps $\psi_s \colon \psi_t(K) \to \psi_{t+s}(K)$ are indicated on the right of each local model.

5.2 The E_0–class

Recall that the 1–skeleton $X_f^{(1)}$ is oriented; see Definition 4.16. Define a cellular 1–chain $\epsilon \in C_1(X_f; \mathbb{R})$ by

$$\epsilon = \frac{1}{2} \sum_e [2 - d(e)]e,$$

where the sum is over all 1–cells e in $X_f^{(1)}$, although the coefficients are nonzero only on edges $e \in E_0$. As an application of the local models, we prove this proposition:

Proposition 5.8 *The chain ϵ is a cycle $\epsilon \in Z_1(X_f; \mathbb{R})$.*

Proof Fix a 0–cell $v \in V\mathcal{E}_0$. We need to show that the value of $\partial\epsilon$ on v is zero. To prove this fact, choose a local model neighborhood $\mathcal{M}(K, s_0)$ around v, so that

$v \in \psi_{s'}(K)$ for some $0 < s' < s_0$. Note that because $v \in V\mathcal{E}_0$, this is not a surface local model, and hence it is a stacked local model which we write as $\mathcal{M} = \mathcal{M}^+ \cup \mathcal{M}^-$ for local models $\mathcal{M}^-$ and $\mathcal{M}^+$ below and above v, respectively.

Let $e_1^-, \ldots, e_r^-$ (respectively, $e_1^+, \ldots, e_{r'}^+$) denote 1–cells in E_0 with $t(e_j^-) = v$ (respectively, $o(e_j^+) = v$). Every $e_j^\pm$ nontrivially intersects the local model $\mathcal{M}^\pm$, as shown in Figure 12. If we let $v_j^- = e_j^- \cap K = \psi_0(K)$ and $v_j^+ = e_j^+ \cap \psi_{s_0}(K)$ then for all j we have

$$\mathrm{val}(v_j^\pm) = d(e_j^\pm)$$

where $\mathrm{val}(v_j^-)$ denotes the valence of v_j^- in K, and $\mathrm{val}(v_j^+)$ denotes the valence of v_j^+ in $\psi_{s_0}(K)$. Moreover, by inspecting the local models, we can see that

$$\sum_{j=1}^r 2 - \mathrm{val}(v_j^-) = 2 - \mathrm{val}(v) = \sum_{j=1}^{r'} 2 - \mathrm{val}(v_j^+),$$

where $\mathrm{val}(v)$ denotes the valence of v in $\psi_s(K)$. Combining these two equations we obtain

$$\sum_{j=1}^r 2 - d(e_j^-) = 2 - \mathrm{val}(v) = \sum_{j=1}^{r'} 2 - d(e_j^+).$$

Finally, we observe that the coefficient of v in $\partial \epsilon$ is equal to one-half of the difference of the two sums here, and hence is zero. This completes the proof of the proposition. $\square$

Definition 5.9 We refer to the homology class of ϵ as the E_0–*class*. As an abuse of notation, we also denote the homology class by $\epsilon \in H_1(X_f; \mathbb{R})$.

Remark 5.10 We now explain the strong geometric analogy between our construction of the E_0–class and the Poincaré dual of the Euler class in the 3–manifold setting.

Suppose that $\eta_0 \colon M \to \mathbb{S}^1$ is a closed, oriented 3–manifold that fibers over $\mathbb{S}^1$ inducing some primitive integral element $(\eta_0)_* = u_0$ with pseudo-Anosov monodromy $F \colon S \to S$. Let $T\mathfrak{F} \to M$ denote the tangent plane bundle of the foliation $\mathfrak{F}$ by fibers (which is also the kernel of the 1–form $\eta_0^*(\mathrm{vol}_{\mathbb{S}^1})$). Taking any section $\sigma \colon M \to T\mathfrak{F}$ transverse to the zero section $Z \subset T\mathfrak{F}$, the Poincaré dual e of the Euler class of $T\mathfrak{F}$ is represented by the oriented 1–submanifold $\sigma^{-1}(Z)$; see [11, Section 12].

Now suppose that the stable foliation $\mathcal{F}$ is orientable, all singularities are 4–prong, and that the orientation is preserved by F. Suspending $\mathcal{F}$ defines a 2–dimensional singular foliation that intersects each fiber in the stable foliation for the first return map to that fiber. Because $\mathcal{F}$ is orientable (and the orientation is preserved by F), we can choose a section σ of $T\mathfrak{F}$ that is also tangent to $\mathcal{F}$ at every nonsingular point, and is zero

precisely at the singular points. Moreover, since the singularities are 4–prong (ie index -1), this can be done so that σ is transverse to Z. Therefore, e is represented by the suspension of the singular points of $\mathcal{F}$.

This conclusion actually holds without any assumption on orientability of $\mathcal{F}$ or on the number of prongs at the singularity, with the appropriate modifications. Namely, if $\gamma_1, \ldots, \gamma_n$ are all the closed orbits of singular points under the suspension flow, and if the number of prongs of a singular point p_i contained in γ_i is denoted $d(\gamma_i)$, then e is represented by the oriented 1–manifold

$$\frac{1}{2}\sum_{i=1}^{n}[2 - d(\gamma_i)]\gamma_i.$$

Here, $\frac{1}{2}(2 - d(\gamma_i))$ is just the index of the singularity (see eg [26, Exposé 5]). The E_0–class is the natural analogue in which the closed orbits are replaced by the 1–cells of the cell structure, and the number of prongs is replaced by degree (as the notation suggests).

6 Cocycles, fibrations, and the E_0–class

Here we begin a more detailed analysis of the map $\eta_0 \colon X_f \to \mathbb{S}^1$. As described in Section 4.1 η_0 induces the integral element $u_0 = (\eta_0)_* \in H^1(X_f; \mathbb{R}) \cong H^1(G; \mathbb{R}) \cong \mathrm{Hom}(G, \mathbb{R})$; see Theorem 4.7 and Convention 4.8. In this section we explain how to construct the neighborhood $\mathcal{A} \subset H^1(X_f; \mathbb{R})$ for which the integral elements in this neighborhood are similarly induced by maps $\eta \colon X_f \to \mathbb{S}^1$. We also begin to explain how much of the structure of η_0 is also present in η. We complete this section with the proof of Theorem B.

6.1 Cellular cocycle representatives

The first task is to provide an explicit cellular cocycle $z_0 \in Z^1(X_f; \mathbb{Z})$ representing u_0. This is done as follows. Recall that unless otherwise stated, the cell structure is always the trapezoid cell structure.

The restriction of η_0 to any cell $\sigma \subset X_f$ can be lifted to the universal cover:

$$\begin{array}{ccc} & & \mathbb{R} \\ & \tilde{\eta}_0|_\sigma \nearrow & \downarrow \\ \sigma & \xrightarrow{\ \eta_0|_\sigma\ } & \mathbb{S}^1 \end{array}$$

Moreover, this lift is unique up to composing with a covering transformation $x \mapsto x+n$, $n \in \mathbb{Z}$.

For any 1–cell e the difference in the values at the end points of the lift $\tilde{\eta}_0|_e$ is independent of the lift, and we denote this difference by $z_0(e)$. This assignment defines a cellular cochain $z_0 \in C^1(X_f; \mathbb{R})$.

Lemma 6.1 *The cochain z_0 is a cocycle representing the class u_0.*

Proof For any based loop $\gamma \colon [0,1] \to X_f$ constructed as a concatenation of oriented 1–cells $e_1^{\delta_1} e_2^{\delta_2} \cdots e_k^{\delta_k}$, where $\delta_j \in \{\pm 1\}$, we obtain lifts $\tilde{\eta}_0|_{e_j}$ to $\mathbb{R}$ by restricting a lift

$$\widetilde{\eta_0 \circ \gamma} \colon [0,1] \longrightarrow \mathbb{R}$$

of $\eta_0 \circ \gamma$. Then

$$(6\text{-}1) \qquad u_0(\gamma) = (\eta_0)_*(\gamma) = \widetilde{\eta_0 \circ \gamma}(1) - \widetilde{\eta_0 \circ \gamma}(0)$$

$$= \sum_{j=1}^{k} \delta_j z_0(e_j) = z_0 \left(\sum_{j=1}^{k} \delta_j e_j \right) = z_0(\gamma).$$

In particular, the value of z_0 around the boundary of any 2–cell is the value of u_0 around the cell which is zero. Thus z_0 is a cocycle, and it agrees with u_0 on every loop. $\qquad \square$

6.2 Positive cocycles

Definition 6.2 Call a cocycle $z \in Z^1(X_f; \mathbb{R})$ *positive* if

$$z(e) > 0$$

for every 1–cell e.

Remark 6.3 This definition is analogous to the notion of *positivity* in Gautero's dynamical 2–complex (see [32, Definition 2.1]), and will serve a similar purpose. See also Wang [63].

Since the restriction of η_0 to any 1–cell e is an orientation preserving diffeomorphism, the discussion in the previous section together with Lemma 6.1 shows that z_0 is a positive cocycle.

Proposition 6.4 *The set of cohomology classes represented by positive 1–cocycles is an open, convex cone $\mathcal{A} \subset H^1(X_f; \mathbb{R})$.*

Proof Choose a basis for $H^1(X_f; \mathbb{R})$ represented by cocycles $\zeta_1, \ldots, \zeta_b$. Suppose z is positive, so that $z(e) > 0$ for every $e \in E_0$. Then one can find $s > 0$ so that for all $s_1, \ldots, s_b$ with $|s_j| < s$ for every j, we also have

$$z(e) + \sum_{j=1}^{b} s_j \zeta_j(e) > 0$$

for every 1–cell e. The set

$$\{[t(z + s_j \zeta_j)] \in H^1(X_f; \mathbb{R}) \mid t > 0, \ |s_j| < s\}$$

is a neighborhood of $[z] \in H^1(X_f; \mathbb{R})$, consisting entirely of cohomology classes represented by positive cocycles. Therefore, $\mathcal{A}$ is open.

Suppose z_1, z_2 are positive and $s, t > 0$. For any 1–cell e, we have $z_1(e) > 0$ and $z_2(e) > 0$ and thus

$$(tz_1 + sz_2)(e) = tz_1(e) + sz_2(e) > 0.$$

Therefore, $\mathcal{A}$ is a convex cone. $\qquad\square$

Definition 6.5 (Positive cone) We denote by $\mathcal{A}_{X_f}$ the set of all $u \in H^1(G_\varphi; \mathbb{R}) = H^1(X_f; \mathbb{R})$ such that u can be represented by a positive 1–cocycle. We call $\mathcal{A}_{X_f}$ the *positive cone* corresponding to X_f.

For each positive 1–cocycle z representing an integral class in $u \in \mathcal{A}$, we now seek to construct a "perturbed" fibration $\eta_z \colon X_f \to \mathbb{S}^1$ that, among other things, induces the homomorphism $u \colon G \to \mathbb{Z}$ on the level of fundamental group. Our construction will make use of the general structure of trapezoidal subdivisions of X_f introduced in Section 4.5, and we refer the reader to that section for our conventions regarding T–orientations and the boundary arcs $\ell_\pm(T)$ and $e_\pm(T)$ of a trapezoid T.

6.3 The subdivision procedure

Consider a trapezoidal subdivision Y of X_f (see Section 4.5). Suppose that Y contains a trapezoid T_0 whose top arc $e_+(T_0)$ consists of a single skew 1–cell. This 1–cell is then necessarily the bottom of another trapezoid T_1 which may again have a single 1–cell along its top arc $e_+(T_1)$. Continuing in this manner we may be able to find a sequence of trapezoids $T_0, T_1, \ldots$ whose top arcs each consist of a single skew 1–cell $\zeta(T_i)e_+(T_i) = e_-(T_{i+1})$. A maximal such sequence (which is completely determined by the initial trapezoid T_0) must either terminate in a trapezoid whose top arc contains multiple 1–cells, or, since there are only finitely many such trapezoids in Y, must

eventually become periodic. In the latter case, the trapezoids $T_i, T_{i+1}, \ldots, T_{i+k} = T_i$ comprising the periodic part of the sequence glue together to give an annulus or Möbius strip in Y which is by construction invariant under the flow ψ. Let us call such an object an *invariant band*.

Lemma 6.6 *Every invariant band B of a trapezoidal subdivision Y of X_f contains a closed orbit $\mathcal{O}$ of ψ in its interior that intersects each trapezoid of the band in a single vertical arc.*

Proof Let us write the invariant band as $B = T_0 \cup \cdots \cup T_k$, where $T_k = T_0$. Since B is invariant under ψ and every flowline of X_f intersects the graph $\Gamma \subset X_f$ infinitely often, it must be the case that some trapezoid of B, say T_0, intersects Γ nontrivially. Moreover, since T_0 is a subtrapezoid of a unique trapezoid T' of X_f and since Γ is by construction disjoint from the interior of $e_{\pm}(T')$ and can intersect T' in at most one connected component (see Section 4.4), we see that $\Gamma \cap T_0$ must consist of a single arc c running from $\ell_-(T_0)$ to $\ell_+(T_0)$. As the intersection of Γ with any trapezoid of X_f is either empty or contained in a single edge of Γ, we furthermore see that $c \subset \Gamma$ is a subarc of a single edge e of Γ.

Let $f' : c \to c$ denote the first return map of the restricted flow $\psi|_B$ to $c \subset B$. Since the first return map of ψ to Γ is the original map f, it follows that $f' = f^m|_c$ for some $m \geq 1$. Using the linear structure on Γ to identify c with a closed interval $[0, 1]$, we now see that f' is an affine homeomorphism from c to itself. Therefore, f' necessarily has a fixed point x in the interior of c. Taking the flowline through x then gives a closed orbit of ψ contained in the interior of B that intersects each trapezoid T_i in a single vertical arc. $\qquad\square$

Remark 6.7 When $f : \Gamma \to \Gamma$ is an expanding irreducible train track map, Lemma A.2 shows that f is expanding on all scales and so there cannot exist a subarc $c \subset \Gamma$ of an edge of Γ and a power $m \geq 1$ for which $f^m|_c$ is a self-homeomorphism. Therefore the proof of Lemma 6.6 implies that a trapezoidal subdivision of X_f cannot contain any invariant bands when $f : \Gamma \to \Gamma$ is an expanding irreducible train track map.

We now define the subdivision procedure that will be used in Section 6.6 below.

Definition 6.8 (Standard subdivision) Let Y be a trapezoidal subdivision of X_f. The *standard subdivision* $\widehat{Y}$ of Y is constructed as follows. Firstly, for each invariant band $B = T_0 \cup \cdots \cup T_k$ (with $T_k = T_0$) of Y, Lemma 6.6 provides a closed orbit $\mathcal{O}$ of ψ in the interior of B. The top and bottom arcs of each trapezoid T_i consist of single skew 1–cells of Y, and the maps $h_{T_i} : e_-(T_i) \to e_+(T_i)$ give us homeomorphisms

$e_-(T_i) \to e_-(T_{i+1})$ between these 1–cells. We subdivide each of these 1–cell into two 1–cells by adding a vertex at the point $v_i = \mathcal{O} \cap e_-(T_i)$ for $0 \leq i < k$. We furthermore subdivide each trapezoid T_i of B into two trapezoids by adding a vertical 1–cell $\beta_i = \mathcal{O} \cap T_i$ from v_i to v_{i+1} (notice that these 1–cells are by construction arcs of flowlines). Let Y^* denote the auxiliary trapezoidal subdivision of Y obtained by simultaneously performing this procedure to each invariant band of Y. In light of Remark 6.7, we see that this has no effect (that is $Y^* = Y$) in the case that $f \colon \Gamma \to \Gamma$ is an expanding irreducible train track map.

We now further subdivide Y^* as follows. First, the vertical 1–cells are not subdivided. Each skew 1–cell of Y^* is $e_-(T)$ for some trapezoid T of Y^*, and we may write the top arc of T as a concatenation of skew 1–cells of Y^*. Let $v_0, v_1, \ldots, v_k$ be the vertices of $e_+(T)$, traversed in that order along $e_+(T)$ (with the T–orientation), so that v_0 is the top endpoint of $\ell_-(T)$ and v_k is the top endpoint of $\ell_+(T)$. We let $u_i = h_T^{-1}(v_i)$, where $h_T \colon e_-(T) \to e_+(T)$ is the map defined by ψ, and we subdivide $e_-(T)$ at the points $u_1, \ldots, u_{k-1}$ yielding new 1–cells $\alpha_i = (u_{i-1}, u_i)$ for $1 \leq i \leq k$ (so if $k = 1$, $e_-(T)$ is not subdivided). Lastly, we subdivide the trapezoid T of Y^* into k trapezoids $T_1, \ldots, T_k$ by adding vertical 1–cells β_i (ie arcs of flowlines) connecting u_i to v_i for each $1 \leq i < k$; see Figure 13. It is straightforward to check that simultaneously applying these subdivisions (of $e_-(T)$ and T) for each trapezoid T of Y^* yields a trapezoidal cell structure $\widehat{Y}$ that subdivides Y.

If Y' is a subdivision of a trapezoidal cell structure Y, we say that a 1–cell σ of Y *becomes subdivided* in Y' if there exists a 1–cell σ' of Y' such that $\sigma' \subsetneq \sigma$. Equivalently, this is the case if and only if σ is not itself a 1–cell of Y'. The same terminology applies to 2–cells of Y, and we note that some 2–cells of Y may not become subdivided in the standard subdivision $\widehat{Y}$ (namely those trapezoids T that are not contained in an invariant band and for which $e_+(T)$ consists of a single skew 1–cell).

Lemma 6.9 *Let Y_0 be a trapezoidal subdivision of X_f, and let $Y_0, Y_1, \ldots$ be the sequence of standard subdivisions defined recursively by $Y_{n+1} = \widehat{Y}_n$. Then every skew 1–cell of Y_n eventually becomes subdivided in some further subdivision Y_m with $m > n$.*

Proof Suppose that a skew 1–cell σ of Y_n remains unsubdivided in Y_m for all $m > n$. Then by nature of the standard subdivision procedure, σ must equal $e_-(T_0)$ for a trapezoid T_0 of Y_n whose top arc $\zeta(T_0)e_+(T_0)$ is a *single* skew 1–cell of Y_n that also remains unsubdivided in Y_m for all $m > n$. In particular, we must have $\zeta(T_0)e_+(T_0) = e_-(T_1)$ for another trapezoid satisfying the same property. Continuing

in this way, we construct an infinite sequence of trapezoids $T_0, T_1, T_2, \ldots$ whose top arcs each consist of a single skew 1–cell $\zeta(T_i)e_+(T_i) = e_-(T_{i+1})$ that remains unsubdivided in Y_m for all $m > n$. Such an infinite sequence must eventually become periodic yielding an invariant band in Y_n. By construction, every skew 1–cell in this invariant band subdivided in $(Y_n)^*$ and therefore also in Y_{n+1}. But this contradicts the fact that each skew 1–cell $e_-(T_i)$ remains unsubdivided in Y_m for all $m > n$. $\square$

6.4 Cocycle refinement

If Y is any trapezoidal cell structure on X_f and Y' is a further trapezoidal subdivision of Y, then any 1–cell σ of Y may be viewed as a cellular 1–chain of Y'. With this in mind, we say that a cocycle $z' \in Z^1(Y'; \mathbb{R})$ is a *refinement* of $z \in Z^1(Y; \mathbb{R})$ if $z'(\sigma) = z(\sigma)$ for every 1–cell σ of Y.

The standard subdivision procedure from Definition 6.8 has an associated method of refinement for positive 1–cocycles. Before describing this refinement, we first introduce some notation.

Definition 6.10 (Skew ratio) Let Y be a trapezoidal subdivision of X_f and let $z \in Z^1(Y; \mathbb{R})$ be a positive cocycle on Y. For a trapezoid T of Y whose top arc is a 1–chain $e_+(T) = \zeta(T)(\sigma_1 + \cdots + \sigma_k)$ we define the *skew ratios* of T to be the numbers

$$0 < \rho_i(T) = \frac{z(\sigma_i)}{\zeta(T)z(e_+(T))} \leq 1.$$

As $z(\sigma_1) + \cdots + z(\sigma_k) = \zeta(T)z(e_+(T))$, we observe that $\rho_1(T) + \cdots + \rho_k(T) = 1$. The *maximal ratio* for T is then defined to be $\rho(T) = \max_{1 \leq i \leq k} \rho_i(T)$, with the convention that $\rho(T) = 0$ in the case that $k = 1$. We thus have $0 \leq \rho(T) < 1$ for every trapezoid T of Y. Finally, we set

$$\rho(Y) = \max_T \rho(T) < 1,$$

where the maximum is over all trapezoids T of Y. To emphasize the dependence on the cocycle z, these quantities will sometimes be notated more elaborately as $\rho_i(T, z)$, $\rho(T, z)$ and $\rho(Y, z)$.

Definition 6.11 (Standard refinement) Let $z \in Z^1(Y; \mathbb{R})$ be a positive cocycle on a trapezoidal subdivision Y of X_f. The *standard refinement* of z is the positive cocycle $\hat{z} \in Z^1(\hat{Y}; \mathbb{R})$ defined as follows: First consider a trapezoid T contained in an invariant band of Y. Then in the auxiliary subdivision Y^*, the bottom arc $e_-(T)$ is subdivided into two skew 1–cells σ_1, σ_2 and the trapezoid T is subdivided into two trapezoids

by the addition of a new vertical 1–cell β. Correspondingly, we define an auxiliary refinement $z^* \in Z^1(Y^*; \mathbb{R})$ of z by declaring

$$(6\text{-}2) \qquad z^*(\sigma_1) = z^*(\sigma_2) = \frac{z(e_-(T))}{2} \quad \text{and} \quad z^*(\beta) = \frac{z(\ell_-(T)) + z(\ell_+(T))}{2}$$

for the new 1–cells arising in any such trapezoid and leaving z unchanged on the remaining 1–cells of Y. It is straightforward to check that z^* then defines a positive refinement of z.

Next consider any trapezoid T of Y^* with top arc a 1–chain $e_+(T) = \zeta(T)(\sigma_1 + \cdots + \sigma_k)$ in Y^*. As in Definition 6.8 we let $\alpha_1, \ldots, \alpha_k$ be the new skew 1–cells for the subdivision of $e_-(T)$ in $\widehat{Y}$, and $\beta_1, \ldots, \beta_{k-1}$ the new vertical 1–cells of $\widehat{Y}$ inside T. These are indexed so that β_i connects the terminal vertex of α_i to the terminal vertex of $h_T(\alpha_i) = \zeta(T)\sigma_i$. We then construct the refinement $\widehat{z}$ of z^* by leaving the value unchanged on the vertical 1–cells of $\ell_\pm(T)$ (which are again vertical 1–cells of $\widehat{Y}$), and for each $1 \le i \le k$ declaring

$$(6\text{-}3) \qquad\qquad \widehat{z}(\alpha_i) = \rho_i(T) z^*(e_-(T))$$

and

$$(6\text{-}4) \quad \widehat{z}(\beta_i) = z^*(\ell_+(T))(\rho_1(T) + \cdots + \rho_i(T)) + z^*(\ell_-(T))(\rho_{i+1}(T) + \cdots + \rho_k(T)).$$

It is straightforward to see that this procedure defines a positive cocycle that refines z^* and thus also z. See Figure 13 for an illustration of this refinement procedure.

Note that skew ratios are essentially preserved under the standard refinement procedure:

Lemma 6.12 Let $z \in Z^1(Y; \mathbb{R})$ be a positive cocycle on a trapezoidal subdivision Y of X_f. Then for the standard subdivision $\widehat{Y}$ and corresponding refinement $\widehat{z} \in Z^1(\widehat{Y}; \mathbb{R})$ we have

$$\rho(\widehat{Y}, \widehat{z}) \le \max\{\rho(Y, z), \tfrac{1}{2}\}.$$

Proof First consider a trapezoid T^* of Y^*. If T^* is a new trapezoid that was created while constructing the subdivision Y^*, then its top arc $e_+(T^*)$ necessarily consists of a single skew 1–cell of Y^* and so we have $\rho(T^*, z^*) = 0$ in this case. Otherwise $T^* = T_0$ for some trapezoid T_0 of Y and we have $\rho(T^*, z^*) = \rho(T_0, z)$ except in the case that the 1–chain $e_+(T_0) = \zeta(T_0)(\sigma_1 + \cdots \sigma_k)$ in Y contains a 1–cell σ_i that becomes subdivided into two 1–cells σ_i' and σ_i'' in Y^*. In that case by (6-2) and Definition 6.10 we have

$$z^*(\sigma_i') = z^*(\sigma_i'') = \frac{z(\sigma_i)}{2} = \tfrac{1}{2}\rho_i(T_0, z)\zeta(T_0)z(e_+(T_0))$$

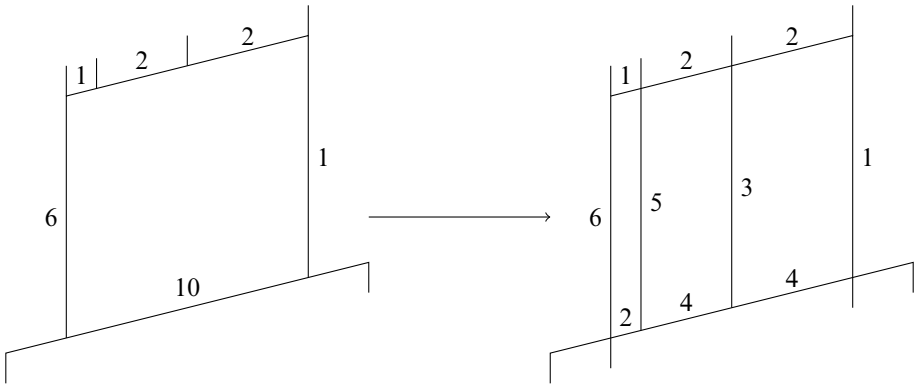$$= \tfrac{1}{2}\rho_i(T_0, z)\zeta(T^*)z^*(e_+(T^*)).$$

Figure 13: A local picture of the subdivision and cocycle refinement procedure for a trapezoid T when $e_+(T)$ consists of more than one skew 1–cell. The numbers along the 1–cells on the left represent the values of some positive cocycle z and the numbers along the 1–cells on the right represent the values of the refinement $\hat{z}$. In this example, we have assumed for simplicity that the trapezoids above T are not subdivided. In general the values along the top will also be "distributed" among the 1–cells in the respective subdivisions.

If $k = 1$, so that $\rho_i(T_0, z) = 1$, then $\rho(T^*, z^*) = \frac{1}{2}$, and otherwise it follows that $\rho(T^*, z^*) \le \rho(T_0, z)$. Whence we have $\rho(T^*, z^*) \le \max\{\rho(Y, z), \frac{1}{2}\}$ for every trapezoid T^* of Y^*.

Next suppose that T is a trapezoid of $\widehat{Y}$ and that we have $e_+(T) = \zeta(T)(\alpha_1 + \cdots + \alpha_k)$ with $k \ge 2$. Notice that, by construction, the top arc $e_+(T)$ of T consists of a *single* 1–cell α of Y^*. Let T^* be the trapezoid of Y^* with $e_-(T^*) = \alpha$. Since the 1–cell α evidently becomes subdivided in $\widehat{Y}$, by definition of the standard subdivision we must have $e_+(T^*) = \zeta(T^*)(\sigma_1 + \cdots + \sigma_k)$, as a 1–chain in Y^*, where h_{T^*} sends α_i to σ_i. From (6-3) we have

$$\hat{z}(\alpha_i) = \rho_i(T^*, z^*)z^*(e_-(T^*)).$$

By the refinement property, we also have $z^*(e_-(T^*)) = \hat{z}(\alpha) = \zeta(T)\hat{z}(e_+(T))$. Therefore

$$\rho_i(T, \hat{z}) = \frac{\hat{z}(\alpha_i)}{\zeta(T)\hat{z}(e_+(T))} = \rho_i(T^*, z^*) \le \rho(T^*, z^*) \le \max\{\rho(Y, z), \frac{1}{2}\}$$

as required. $\qquad\square$

Lemma 6.12, Lemma 6.9 and the definition of standard refinement now give the following:

Lemma 6.13 Let Y_0 be a trapezoidal subdivision of X_f and $z_0 \in Z^1(Y_0; \mathbb{R})$ a positive cocycle. Consider the sequence of standard subdivisions $Y_0, Y_1, \dots$ with corresponding cocycle refinements $z_0, z_1, \dots$ defined recursively by $Y_{n+1} = \widehat{Y}_n$ and $z_{n+1} = \widehat{z}_n$. Then the numbers

$$S_n = \max\{z_n(\sigma) \mid \sigma \text{ is a skew } 1\text{–cell of } Y_n\}$$

tend to zero as $n \to \infty$.

Proof Let $\delta = \max\{\rho(Y_0, z_0), \frac{1}{2}\} < 1$. Applying Lemma 6.12 inductively, we see that $\rho(T, z_n) \leq \delta$ for all trapezoids T of Y_n and for all $n \geq 0$. According to (6-2) and (6-3), whenever a skew 1–cell $\sigma = e_-(T)$ of Y_n becomes subdivided in Y_{n+1}, each resulting subcell $\sigma' \subsetneq \sigma$ satisfies

$$z_{n+1}(\sigma') \leq \max\{\rho(T), \tfrac{1}{2}\} z_n(e_-(T)) \leq \delta z_n(\sigma).$$

By Lemma 6.9, for each n there exists $m > n$ so that every 1–cell of Y_n has become subdivided in Y_m. The above inequality implies $S_m \leq \delta S_n$ and the result follows. $\square$

6.5 An unconstrained refinement

Our construction of the perturbed fibration $\eta_z \colon X_f \to \mathbb{S}^1$ associated to a primitive integral element $[z] \in \mathcal{A}$ will rely on finding a subdivision of X_f and a positive refinement of z in which every trapezoid enjoys the following property.

Definition 6.14 (Unconstrained trapezoids) Let Y be a trapezoidal subdivision of X_f and let $z \in Z^1(Y; \mathbb{R})$ be a 1–cocycle. A trapezoid T of Y is said to be *unconstrained by z* if

$$\max\{0, \ z(e_-(T))\} \leq \min\{z(\ell_-(T)), \ z(\ell_-(T)) + z(e_+(T))\}.$$

We note that when z is positive, the above definition reduces to $z(e_-(T)) \leq z(\ell_-(T))$ since the cocycle condition $z(\partial T) = 0$ (see (4-2)) immediately yields

$$0 < z(e_-(T)) < z(\ell_+(T)) + z(e_-(T)) = z(\ell_-(T)) + z(e_+(T)).$$

Therefore, in this case a trapezoid is automatically unconstrained whenever $z(e_+(T)) < 0$. In particular, we see that trapezoids with $\zeta(T) = -1$ are unconstrained by *every* positive 1–cocycle.

We now employ the constructions from Sections 6.3–6.4 to find such a refinement. For the proof, we recall from Section 4.5 that X_f may contain degenerate trapezoids and that the sign of every degenerate trapezoid T is $\zeta(T) = -1$ (Remark 4.20).

Proposition 6.15 Let Y_0 denote the trapezoidal cell structure X_f and let $z_0 = z \in Z^1(X_f; \mathbb{R})$ be a positive cocycle. Consider the sequence $Y_0, Y_1, \ldots$ of standard subdivisions defined recursively by $Y_{n+1} = \widehat{Y}_n$ with corresponding cocycle refinements $z_{n+1} = \widehat{z}_n$. Then there exist some $n \geq 0$ so that every trapezoid T of Y_n is unconstrained by $z_n \in Z^1(Y_n; \mathbb{R})$.

Proof Let D denote the set of degenerate trapezoids of X_f. Any $T' \in D$ has $\zeta(T') = -1$, and furthermore any trapezoid T of a subdivision Y_n which is contained in some $T' \in D$ also has $\zeta(T) = -1$ (see Remark 4.20 and the last paragraph of Section 4.5). Therefore, any such T is unconstrained by z_n as the calculation following Definition 6.14 demonstrates. So we need only consider trapezoids T not contained in any trapezoid in D.

Let $M > 0$ be the minimum of $z(\beta)$ over all vertical 1–cells of X_f. We claim that for $n \geq 0$ every vertical 1–cell β of Y_n that is not contained in the interior of some $T' \in D$ also has $z_n(\beta) \geq M$. This is true for $n = 0$ by definition of M; arguing by induction we assume it is true for some $n \geq 0$ and prove it for $n + 1$. The condition holds for all vertical 1–cells of Y_{n+1} that are also vertical 1–cells of Y_n by induction hypothesis. Thus it suffices to consider a new vertical 1–cell of Y_{n+1} that is not contained in any trapezoid in D and is not a 1–cell of Y_n. By construction, such a 1–cell can only arise as a vertical arc β_i connecting $e_-(T)$ to $e_+(T)$ for some trapezoid T of Y_n that is not contained in any $T' \in D$. We then have $z_n(\ell_{\pm}(T)) \geq M$ since the 1–cells comprising these chains are not contained in the interior of any $T' \in D$. As $\rho_1(T) + \cdots + \rho_k(T) = 1$, equations (6-2) and (6-4) show that $z_{n+1}(\beta_i)$ is a convex combination of $z_n(\ell_+(T)) \geq M$ and $z_n(\ell_-(T)) \geq M$, and hence $z_{n+1}(\beta_i) \geq M$, as required.

By Lemma 6.13 we may now choose n so that every skew 1–cell σ of Y_n satisfies $z_n(\sigma) < M$. For any trapezoid T of Y_n that is not contained in a trapezoid of D, we then have $z_n(\ell_{\pm}(T)) \geq M$ by the preceding paragraph. Therefore

$$z_n(e_-(T)) < M \leq z_n(\ell_-(T)),$$

showing that T is unconstrained by z_n and completing the proof. $\quad\square$

6.6 Perturbed fibrations

We now utilize the trapezoidal subdivisions and cocycle refinements developed in Sections 4.5–6.5 to construct perturbed fibrations $\eta_z \colon X_f \to \mathbb{S}^1$ for positive 1–cocycles representing primitive integral classes in $\mathcal{A}$. The proof will also utilize the local model

neighborhoods introduced in Section 5. For any local model $\mathcal{M} \subset X_f$, we denote a lift to the universal covering $\pi\colon \mathbb{R} \to \mathbb{S}^1$ of the restriction of η_0 to $\mathcal{M}$ by

$$\eta_0^{\mathcal{M}}\colon \mathcal{M} \longrightarrow \mathbb{R}$$

and write $I_{\mathcal{M}} = \eta_0^{\mathcal{M}}(\mathcal{M})$.

Lemma 6.16 *Let $z \in Z^1(X_f; \mathbb{R})$ be a positive cocycle representing an integral class $u \in \mathcal{A}$. Then there exists a map $\eta_z\colon X_f \to \mathbb{S}^1$ and a finite covering $\mathbb{M}$ of $X_f^{(1)}$ by local model neighborhoods with the following properties.*

(1) $(\eta_z)_* = u\colon \pi_1(X_f) \to \pi_1(\mathbb{S}^1) = \mathbb{Z}$ *in* $H^1(X_f; \mathbb{R})$.

(2) *For each local model $\mathcal{M} \in \mathbb{M}$, there is an orientation preserving injective local diffeomorphism $h_{\mathcal{M}}\colon I_{\mathcal{M}} \to \mathbb{R}$ so that the restriction $\eta_z|_{\mathcal{M}}$ satisfies*

$$\eta_z|_{\mathcal{M}} = \pi \circ h_{\mathcal{M}} \circ \eta_0^{\mathcal{M}}.$$

(3) *The restriction of η_z to any semiflow line is a local diffeomorphism and the restriction to the interior of each trapezoid is a submersion.*

(4) *For some $x_0 \in \mathbb{S}^1$, $\eta_z^{-1}(x_0) \cap X_f^{(1)} \subset (X_f^{(1)} \setminus X_f^{(0)})$ consists of a finite set of points which flow into periodic orbits of ψ. Moreover, if f is an expanding irreducible train track map, then all points of $\eta_z^{-1}(x_0) \cap X_f^{(1)}$ eventually flow into vertical 1–cells of X_f.*

Part (2) of the lemma basically says that in each local model $\mathcal{M}$, the map η_z factors through the restriction to $\mathcal{M}$ of η_0 (but to make this correct, we need to use $\eta_0^{\mathcal{M}}$ instead of the restriction of η_0 to $\mathcal{M}$ exactly). As a consequence of (2), we see that the fibers of η_z are horizontal in $\mathcal{M}$.

Proof The map $\eta_z\colon X_f \to \mathbb{S}^1$ will be defined so that for any 1–cell e, if we lift the restriction to the universal cover $\tilde{\eta}_z|_e\colon e \to \mathbb{R}$, then the difference in the values at the endpoints (which is independent of the lift) is exactly $z(e)$. A calculation as in Equation (6-1) shows that any such map will satisfy conclusion (1) of the lemma. The plan is to first construct a map on the 1–skeleton $\eta_z'\colon X_f^{(1)} \to \mathbb{S}^1$ satisfying this property, so that the restriction to each 1–cell e is a local diffeomorphism onto its image, and so that (4) is satisfied. Assuming that every trapezoid of X_f is unconstrained by z, we may then extend η_z' over the 2–cells so that (3) is satisfied. Finally, we modify η_z' in a neighborhood of $X_f^{(1)}$ and construct the covering $\mathbb{M}$ so that (2) is satisfied.

By Proposition 6.15, after passing to a trapezoidal subdivision Y of X_f and a corresponding positive refinement z' of z, we may assume that every trapezoid of Y is

unconstrained by z'. Suppose we can prove the lemma for the cell structure Y and corresponding refinement z'. The first assertion of (4) will then follow from the fact that $X_f^{(1)} \subset Y^{(1)}$. Furthermore, in the case that f is an expanding irreducible train track map, all vertical 1–cells of Y eventually flow into the vertical 1–cells of X_f (see Definition 6.8), so the second assertion of (4) will follow as well. Since any finite cover of the 1–skeleton of Y also covers the 1–skeleton of X_f, part (2) will follow, and (3) will follow similarly by using local models near the new vertical 1–cells. Finally, (1) does not make reference to the cell structure and so will follow automatically. Thus it suffices to prove the lemma for Y and z'; by an abuse of notation we therefore continue to refer to this subdivision and refined cocycle simply as $Y = X_f$ and $z' = z$, respectively.

Choose a maximal tree $Q \subset X_f^{(1)}$ and fix some vertex $v_0 \in Q$ to serve as a basepoint. Define a map $\hat{\eta}_z \colon Q \to \mathbb{R}$ so that $\hat{\eta}_z(v_0) = 0$ and so that the restriction to every 1–cell e in Q is an injective local diffeomorphism and so that

$$\hat{\eta}_z(t(e)) - \hat{\eta}_z(o(e)) = z(e).$$

Because z represents an integral cohomology class, we see that for any 1–cell $e \in E_0$ which is not in Q, although this equation can fail, we still have

$$\hat{\eta}_z(t(e)) - \hat{\eta}_z(o(e)) - z(e) \in \mathbb{Z}.$$

Therefore, composing $\hat{\eta}_z$ with π, we obtain a map $\eta'_z \colon Q \to \mathbb{S}^1$ which can we can extend to all of $X_f^{(1)}$ with the property that for any 1–cell e, any lift $\tilde{\eta}'_z|_e \colon e \to \mathbb{R}$ is an injective local diffeomorphism and

$$\tilde{\eta}'_z|_e(t(e)) - \tilde{\eta}'_z|_e(o(e)) = z(e).$$

We next modify $\eta'_z \colon X_f^{(1)} \to \mathbb{S}^1$ by a homotopy rel $X_f^{(0)}$, maintaining the property that η'_z is a local diffeomorphism on each edge, so that on a neighborhood U of $X_f^{(0)}$ our map η'_z factors through the restriction of η_0 to U. More precisely, we arrange that each connected component U_0 of U contains a single vertex of $X_f^{(0)}$ and that for any lift $\tilde{\eta}_0 \colon U_0 \to \mathbb{R}$ of the restriction of η_0 to U_0 there exists a local diffeomorphism $h \colon \mathbb{R} \to \mathbb{R}$ so that

$$\eta'_z|_{U_0 \cap X_f^{(1)}} = \pi \circ h \circ \tilde{\eta}_0|_{U_0 \cap X_f^{(1)}}.$$

This modification ensures that two points of $U_0 \cap X_f^{(1)}$ are identified by η'_z if and only if they are also identified by $\tilde{\eta}_0$.

After shrinking U if necessary, we may then choose a point $x_0 \in \mathbb{S}^1 \setminus \eta'_z(U)$. Note that when $f \colon \Gamma \to \Gamma$ is an expanding irreducible train track map, the set $\bigcup_{k \geq 1} f^{-k}(V\Gamma)$ is dense in Γ (see the proof of Theorem A.3). Therefore, the set of points in X_f

that eventually flow into vertical 1–cells by ψ is dense and these eventually flow into periodic orbits of ψ (see Remark 4.18). More generally, if $\mathcal{P} \subset \Gamma$ is the set of periodic points for f, then $\bigcup_{k \geq 1} f^{-k}(\mathcal{P})$ is also dense (as can be seen from the linear structure of f), and hence the set of points of X_f that eventually flow into periodic orbits is always dense. Therefore, we may modify $\eta'_z: X_f^{(1)} \to \mathbb{S}^1$ by a homotopy rel $U \supset X_f^{(0)}$ so that $(\eta'_z)^{-1}(x_0)$ (which is a finite subset of $X_f^{(1)} \setminus U$) consists of points that eventually flow into periodic orbits of ψ (and that furthermore flow into vertical 1–cells of X_f in the case that f is an expanding irreducible train track map). The extension $\eta_z: X_f \to \mathbb{S}^1$ of η'_z that we construct below will agree with η'_z on $X_f^{(1)}$ and will therefore satisfy condition (4).

We then extend η'_z over the 2–cells of X_f as follows. Consider a trapezoid T of X_f, and let $v \in X_f^{(0)}$ denote the common initial vertex of $e_-(T)$ and $\ell_-(T)$ (recall that $e_-(T)$ is oriented from $\ell_-(T)$ to $\ell_+(T)$). Lifting $\eta'_z|_{\partial T}$ to the universal cover gives us a map $\widetilde{\eta}'_z|_{\partial T}: \partial T \to \mathbb{R}$ which, after composing with a translation, we may assume sends v to 0. By construction, $\widetilde{\eta}'_z|_{\partial T}$ then maps $e_-(T)$ and $e_+(T)$ locally diffeomorphically onto the intervals

$$[0, z(e_-(T))] \quad \text{and} \quad [z(\ell_-(T)), z(\ell_-(T)) + z(e_+(T))],$$

respectively (in the case $\zeta(T) = -1$, note that the endpoints of this latter interval have been written out of order). Since T is unconstrained by z, these intervals have disjoint interiors. Therefore the map $\widetilde{\eta}'_z|_{\partial T}$ does not identify any interior point of $e_-(T)$ with any interior point of $e_+(T)$. Rather, $\widetilde{\eta}'_z|_{\partial T}$ identifies each point of $e_-(T)$ with a point of $\ell_-(T)$, and each point of $e_+(T)$ with a point of either $\ell_+(T)$ (in the case $\zeta(T) = 1$) or $\ell_-(T)$ (in the case $\zeta(T) = -1$). It follows that we may extend $\pi \circ \widetilde{\eta}'_z|_{\partial T}$ to a map $\eta'_z: \overline{T} \to \mathbb{S}^1$ whose fibers are transverse to the flow. Explicitly, identifying $\overline{T}$ with the Euclidean trapezoid $\widehat{T}$ via the map $\sigma_T: \widehat{T} \to \overline{T}$ provided by Proposition 4.21, we extend $\widetilde{\eta}'_z|_{\partial T}$ over $\overline{T}$ so that the fibers are straight (Euclidean) lines in $\widehat{T}$ and define η'_z on $\overline{T}$ as $\pi \circ \widetilde{\eta}'_z|_{\overline{T}}$. The unconstrained condition then ensures that each fiber has nonvertical slope and is therefore transverse to the foliation of $\widehat{T}$ by (vertical) flowlines. Extending η'_z over every trapezoidal 2–cell of X_f in this manner, we obtain the map $\eta'_z: X_f \to \mathbb{S}^1$.

Notice that the fibers of η'_z are by construction transverse to $X_f^{(1)}$. We may therefore modify η'_z by a homotopy rel the 1–skeleton to obtain a map $\eta_z: X_f \to \mathbb{S}^1$ whose fibers remain transverse to the flow and are furthermore horizontal in a neighborhood W of $X_f^{(1)}$. This map satisfies conclusions (1) and (3) of the lemma.

Finally to prove part (2), we construct the cover $\mathbb{M}$ of $X_f^{(1)}$ by local models. For each 0–cell v of X_f we use Proposition 5.6 to find a local model neighborhood $\mathcal{M}_v$ of v

contained in $U \cap W$. The local models $\mathcal{M}_v$ and $\mathcal{M}_{v'}$ are then disjoint for distinct 0–cells (since each component of U contains a single 0–cell). We next cover each edge by finitely many 1–cell local models that are each contained in W and let $\mathbb{M}$ denote the resulting cover. Since each $\mathcal{M} \in \mathbb{M}$ is contained in W, we see that the fibers of $\eta_z|_\mathcal{M}$ are necessarily horizontal. This implies that two points of $\mathcal{M}$ are identified by a lift $\widetilde{\eta}_z|_\mathcal{M}$ if and only if they are also identified by $\eta_0^\mathcal{M}$. Therefore the maps $\widetilde{\eta}_z|_\mathcal{M}$ and $(\eta_0^\mathcal{M})^{-1}$ may be composed to yield a well defined map

$$h_\mathcal{M} = \widetilde{\eta}_z|_\mathcal{M} \circ (\eta_0^\mathcal{M})^{-1} \colon I_\mathcal{M} \longrightarrow \mathbb{R}.$$

This map $h_\mathcal{M}$ is a local diffeomorphism (as can be seen by restricting to a flowline in $\mathcal{M}$) and thus satisfies conclusion (2) of the lemma. $\qquad\square$

Remark 6.17 Alternatively, one could work with *continuous* 1–*forms* on X as described in [25]. We do not need any of the technical machinery developed in that setting here, and our situation requires other special considerations, so we decided to avoid adding another layer of terminology.

Definition 6.18 (Associated fibrations and the graph Θ_z) Let $u \in H^1(X_f; \mathbb{R})$ be an integral class represented by a positive 1–cocycle $z \in Z^1(X_f; \mathbb{R})$. A map $\eta_z \colon X_f \to \mathbb{S}^1$ satisfying the conclusions of Lemma 6.16 will be called an *associated fibration* for z. For any point $x \in \mathbb{S}^1$, let $\Theta_{x,z} = \eta_z^{-1}(x)$ be the full preimage of x. By the second and third condition of Lemma 6.16, the fiber $\Theta_{x,z}$ is a finite topological graph with vertex set equal to the intersection with $X_f^{(1)}$. The *associated graph* Θ_z for this fibration is defined by choosing a point $x_0 \in \mathbb{S}^1$ as in part (4) and letting $\Theta_z := \Theta_{x_0,z}$ as a topological space; as topological graph we equip Θ_z with the vertex set

$$V\Theta_z = \left(\bigcup_{s \geq 0} \psi_s(V\Theta_{x_0,z}) \right) \cap \Theta_{x_0,z}.$$

According to part (4), $V\Theta_{x_0,z}$ flows into periodic orbits of ψ, and hence Θ_z is indeed a finite subdivision of $\Theta_{x_0,z}$. Note that by construction, every edge of Θ_z is contained in a trapezoidal 2–cell of X_f (since this is already true for $\Theta_{x_0,z}$).

Let u, z, η_z and Θ_z be as in Definition 6.18 above. To understand the graph Θ_z, we need to study some general properties of the graphs $\Theta_{x,z}$ (where $x \in \mathbb{S}^1$ is arbitrary) and how they transform under ψ. By the third condition of Lemma 6.16, we can reparameterize the semiflow ψ_s so that the composition $\mathbb{R}_+ \to \mathbb{S}^1$ of any semiflow line with the projection to the circle has unit speed. We denote this reparameterization by ψ_s^z. With this reparameterization we have

$$\psi_s^z(\Theta_{x,z}) = \Theta_{x+s,z},$$

where $s \in \mathbb{R}$ and $x + s \in \mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$. By Lemma 6.16(2), given a local model $\mathcal{M} = \mathcal{M}(K, s_0) \in \mathbb{M}$, there exists s_0^z so that

$$\mathcal{M} = \bigcup_{0 \le s \le s_0^z} \psi_s^z(K).$$

Furthermore, there is a homeomorphism $h_{\mathcal{M},z} \colon [0, s_0] \to [0, s_0^z]$ so that

$$\psi_s(K) = \psi_{h_{\mathcal{M},z}(s)}^z(K) \quad \text{for all } s \in [0, s_0]$$

and consequently Proposition 5.6 is valid for ψ^z in place of ψ.

Lemma 6.19 *For any $x \in \mathbb{S}^1$ there exists an $s_x > 0$ such that for all $-\frac{1}{2}s_x \le s \le s' \le \frac{1}{2}s_x$ the map*

$$\psi_{s'-s}^z \colon \Theta_{x+s,z} \longrightarrow \Theta_{x+s',z}$$

is a homotopy equivalence.

Proof For any $x' \in \mathbb{S}^1$ and $0 < s' < 1$ let

$$N(x', s') = \bigcup_{0 \le s \le s'} \psi_s^z(\Theta_{x',z}) = \eta_z^{-1}([x', x' + s']).$$

Now cover the intersections of $\Theta_{x,z}$ with $X_f^{(1)}$ by the flow-interiors of local models. By replacing the local models $\mathcal{M} \in \mathbb{M}$ provided by Lemma 6.16 with smaller local models if necessary, we can assume that $\Theta_{x,z}$ does not intersect the top or bottom of any $\mathcal{M}$ and that for all $\mathcal{M} \ne \mathcal{M}' \in \mathbb{M}$, we have

$$\Theta_{x,z} \cap \mathcal{M} \cap \mathcal{M}' = \varnothing.$$

Then we can choose $0 < s_x < 1$ so that for any local model $\mathcal{M} = \mathcal{M}(K, s_0)$ there exist $0 < s_1^{\mathcal{M}} < s_2^{\mathcal{M}} < \cdots < s_r^{\mathcal{M}} < s_0^z$ (with $r = r(\mathcal{M})$) with

$$N(x - \tfrac{1}{2}s_x, s_x) \cap \mathcal{M} = \bigcup_{j=1}^{r} \left(\bigcup_{0 \le s \le s_x} \psi_{s_j^{\mathcal{M}}+s}^z(K) \right).$$

Furthermore, by choosing s_x sufficiently small we can also guarantee that

$$N(x - \tfrac{s_x}{2}, s_x) \cap \mathcal{M} \cap \mathcal{M}' = \varnothing$$

for all $\mathcal{M} \ne \mathcal{M}' \in \mathbb{M}$.

Applying Proposition 5.6 it follows that for all $-\frac{1}{2}s_x \le s \le s' \le \frac{1}{2}s_x$ and all $\mathcal{M} \in \mathbb{M}$, we have

$$\psi_{s'-s}^z \colon \Theta_{x+s,z} \cap \mathcal{M} \longrightarrow \Theta_{x+s',z} \cap \mathcal{M}$$

is controlled rel boundary. Any component of the complement

$$U \subset N(x - \tfrac{1}{2}s_x, s_x) \setminus \bigcup_{\mathcal{M} \in \mathbb{M}} \mathcal{M}$$

is entirely contained in some trapezoid T. In T the semiflow is a flow, and so the restriction

$$\psi^z_{s'-s} \colon \Theta_{x+s,z} \cap U \longrightarrow \Theta_{x+s',z} \cap U$$

is a homeomorphism.

It follows that $\psi^z_{s'-s} \colon \Theta_{x+s,z} \to \Theta_{x+s',z}$ is locally controlled, and so by Proposition 5.5 it is a homotopy equivalence. □

Corollary 6.20   *For any $x \in \mathbb{S}^1$ and $s \in \mathbb{R}$ we have that*

$$\psi^z_s \colon \Theta_{x,z} \longrightarrow \Theta_{x+s,z}$$

is a homotopy equivalence. In particular, the first return map

$$f_z = \psi^z_1 \colon \Theta_z \longrightarrow \Theta_z$$

is a homotopy equivalence. Furthermore, $f_z(V\Theta_z) \subset V\Theta_z$.

Proof   We can write $\psi^z_s = \psi^z_{s_1} \circ \psi^z_{s_2} \circ \cdots \circ \psi^z_{s_k}$, where each restriction

$$\psi^z_{s_j} \colon \Theta_{x_{j-1},z} \longrightarrow \Theta_{x_j,z} \quad \text{for } j = 1, \ldots, k,$$

is a homotopy equivalence given by Lemma 6.19 for some $x_0, \ldots, x_k$ with $x_0 = x$ and $x_k = x + s$. This proves the first (and hence also the second) statement. The last statement follows from the definition of $V\Theta_z$ (see Definition 6.18) since f_z is the first return map of ψ. □

As a consequence we have the following Corollary (cf [32, Theorem 3.1] and [63, Lemma 4.1.3]).

Corollary 6.21   *For any primitive integral element $u \in \mathcal{A}$ represented by a positive 1–cocycle z with associated fibration η_z, the fiber Θ_z is connected and the induced map $\pi_1(\Theta_z) \to \pi_1(X_f) = G$ is an isomorphism onto the kernel of $u = (\eta_z)_* \colon G \to \mathbb{Z}$.*

Proof   The flow $\psi^z_s \colon \Theta_{x,z} \to \Theta_{x+s,z}$ is a homotopy equivalence for all x, s, so the number of components n of $\Theta_{x,z}$ is constant, independent of x. We note that $X = \bigcup_{x \in \mathbb{S}^1} \Theta_{x,z}$ and use this to construct a quotient $F \colon X \to X/\!\sim$, where $\sim$ is the equivalence relation whose equivalence classes are precisely the connected components of $\Theta_{x,z}$ for $x \in \mathbb{S}^1$. There is a larger equivalence relation $\sim'$ whose equivalence classes

are the sets $\Theta_{x,z}$ for $x \in \mathbb{S}^1$, and this quotient $X/\sim'$ is just $\mathbb{S}^1$ itself. The quotient $X \to \mathbb{S}^1$ factors as $X \to X/\sim \to \mathbb{S}^1$, and the $X/\sim \to \mathbb{S}^1$ is everywhere n-to-1. In fact, this map is an n–fold covering map, $X/\sim = \mathbb{S}^1 \to \mathbb{S}^1$. Therefore, the homomorphism $u\colon G \to \mathbb{Z}$ factors through the induced map of this cover $F_* \colon G \to \pi_1(X/\sim)$. But the induced map from $X/\sim = \mathbb{S}^1 \to \mathbb{S}^1$ is multiplication by $\pm n$, and hence $u = \pm n F_*$. Since u is primitive, it follows that $n = 1$, so $\Theta_{x,z}$ is connected for all x.

For the second statement, apply van Kampen's theorem to calculate $\pi_1(X_f)$ as a semidirect product $\pi_1(\Theta_z) \rtimes \mathbb{Z}$, and observe that the projection onto $\mathbb{Z}$ is $(\eta_z)_*$. $\square$

6.7 E_0–class and the proof of Theorem B

In this section we will establish Theorem B from Section 1. Recall (cf Convention 4.1) that for any $\varphi \in \mathrm{Out}(F_N)$, where $N \geq 2$, given a tame (linear) graph-map $f\colon \Gamma \to \Gamma$ that is a topological representative for φ along with some folding sequence as in (3-1), we have constructed the corresponding folded mapping torus X_f for φ. As noted in Remark 3.2, for any $\varphi \in \mathrm{Out}(F_N)$ there exists a tame representative f as above.

Recall also that in Definition 5.7 we defined a subgraph $\mathcal{E}_0$ of $X_f^{(1)}$, that in Definition 5.9 we defined the E_0–class $\epsilon \in H_1(X_f; \mathbb{R})$, and that sections of the semiflow ψ were defined in Section 1.

We now prove:

Theorem B *Let $G_\varphi = F_N \rtimes_\varphi \mathbb{Z}$ be the mapping torus group of $\varphi \in \mathrm{Out}(F_N)$ (where $N \geq 2$), let $f\colon \Gamma \to \Gamma$ be a tame graph-map which is a topological representative of φ, and let X_f be a folded mapping torus for f with the associated semiflow ψ, positive cone $\mathcal{A}_{X_f} \subset H^1(X_f; \mathbb{R})$, and E_0–class $\epsilon \in H_1(X_f; \mathbb{R})$. Given a positive cellular 1–cocycle $z \in Z^1(X_f; \mathbb{R})$ representing a primitive integral element $u \in \mathcal{A}_{X_f} \subset H^1(X_f; \mathbb{R}) = H^1(G_\varphi; \mathbb{R})$, there exists a map $\eta_z\colon X_f \to \mathbb{S}^1$ and a fiber $\Theta_z = \eta_z^{-1}(x_0)$ for some $x_0 \in \mathbb{S}^1$ so that:*

(1) *Θ_z is a finite connected topological graph such that the inclusion $\Theta_z \subset X_f$ is π_1–injective and such that $\pi_1(\Theta_z) = \ker(u) \leq \pi_1(X_f) = G_\varphi$.*

(2) *$u(\epsilon) = \chi(\Theta_z) = 1 - \mathrm{rk}(\ker(u))$.*

(3) *Θ_z is a section of ψ.*

(4) *The first return map $f_z\colon \Theta_z \to \Theta_z$ of ψ is a homotopy equivalence with $f(V\Theta_z) \subset V\Theta_z$ and $(f_z)_*$ represents $\varphi_u \in \mathrm{Out}(\ker(u))$.*

Proof The existence of η_z comes from Lemma 6.16, and property (1) is a consequence of Corollary 6.21. Property (3) is immediate and (4) follows from Corollary 6.20.

All that remains is to prove (2). Since $\chi(\Theta_z) = 1 - \mathrm{rk}(\ker(u))$, to establish (2) it suffices to prove $u(\epsilon) = \chi(\Theta_z)$.

Write $\epsilon = \sum_{e \in E_0} c_e e \in Z_1(X_f; \mathbb{R})$, where $c_e = (2 - d(e))/2$. Let x be any point not in the η_z–image of the zero-skeleton $X_f^{(0)}$. Since $(\eta_z)_* = u$ and z is a cocycle that represents u, we can calculate $z(\epsilon)$ by pushing ϵ forward by η_z, and pairing this 1–cocycle on $\mathbb{S}^1$ with the Poincaré dual of the fundamental class. This class is represented in de Rham cohomology, $H^1(\mathbb{S}^1; \mathbb{R}) = H^1_{\mathrm{dR}}(\mathbb{S}^1; \mathbb{R})$ by any 1–form ω on $\mathbb{S}^1$ with total volume 1. Pick such a 1–form ω with support contained in an interval about x disjoint from $\eta_z(X_f^{(0)})$.

Next, recall that the restriction of η_z to any $e \in E_0$ is an orientation preserving local diffeomorphism. In particular, we can consider the restriction of η_z to e as a smooth, singular 1–chain in $\mathbb{S}^1$, denoted e_z, which is something that we can integrate ω over. By our choice of ω and since η_z is orientation preserving on e, we have

$$\int_{e_z} \omega = \left| \eta_z^{-1}(x) \cap e \right|.$$

Thus, from the previous paragraph we have

$$z(\epsilon) = \sum_{e \in E_0} c_e \int_{e_z} \omega = \frac{1}{2} \sum_{e \in E_0} (2 - d(e)) \left| \eta_z^{-1}(x) \cap e \right|.$$

Now observe that $\eta_z^{-1}(x) \cap \mathcal{E}_0$ is precisely the set of vertices of $\Theta_{x,z}$ of valence greater than 2. Indeed, if $y \in \eta_z^{-1}(x) \cap e$, then the valence $\mathrm{val}(y)$ in $\Theta_{x,z}$ is precisely $d(e)$. Therefore, from this fact and the previous equation we have

$$z(\epsilon) = \frac{1}{2} \sum_{y \in V\Theta_{x,z}} (2 - \mathrm{val}(y)) = \chi(\Theta_{x,z}).$$

By definition, Θ_z is a subdivision of $\Theta_{x_0,z}$ for some x_0 not in the image of $X_f^{(0)}$. Hence $z(\epsilon) = \chi(\Theta_{x_0,x}) = \chi(\Theta_z)$, as required. $\qquad \square$

7 Train track maps and full irreducibility

Recall that if $\varphi \in \mathrm{Out}(F_N)$ is represented by an expanding irreducible train-track map $f: \Gamma \to \Gamma$, then f is tame by Lemma 2.12 and so automatically satisfies Convention 3.1. Thus given any folding sequence (3-1) for f and compatible metric structure on Γ

(cf Convention 4.1), we can construct the corresponding folded mapping torus X_f for φ.

In this section we prove Theorem C from Section 1.

Theorem C *Let $\varphi \in \mathrm{Out}(F_N)$ (where $N \geq 2$) be an outer automorphism represented by an expanding irreducible train track map $f: \Gamma \to \Gamma$, let X_f be a folded mapping torus, and ψ the associated semiflow. Let $u \in \mathcal{A}_{X_f} \subset H^1(G_\varphi; \mathbb{R}) = H^1(X_f; \mathbb{R})$ be a primitive integral element which is represented by a positive 1–cocycle $z \in Z^1(X_f; \mathbb{R})$, and let $\Theta_z \subset X_f$ be the topological graph from Theorem B so that the first return map $f_z: \Theta_z \to \Theta_z$ is a homotopy equivalence representing φ_u. Then:*

(1) *The graph Θ_z can be chosen so that the first return map $f_z: \Theta_z \to \Theta_z$ is an expanding irreducible train-track map.*

(2) *If we additionally assume that φ is hyperbolic and fully irreducible, then $\varphi_u \in \mathrm{Out}(\ker(u))$ is also hyperbolic and fully irreducible.*

The proof of Theorem C, will be given after establishing a sequence of lemmas. First we make the following convention.

Convention 7.1 For the remainder of the section, for the purposes of proving part (1) of Theorem C, we assume that $f: \Gamma \to \Gamma$ is an expanding irreducible train track map representing $\varphi \in \mathrm{Out}(F_N)$. We do not assume φ is hyperbolic and fully irreducible until specifically stated; see Convention 7.6.

Note that by Lemma 2.12 the map f is tame. We equip Γ with the linear and metric structures making f into a combinatorial graph map (see Lemma A.2 and Corollary A.9), and then construct the corresponding folding line and the folded mapping torus X_f. We assume z is a positive cocycle representing a primitive integral class $u \in \mathcal{A} = \mathcal{A}_{X_f}$ and that η_z is an associated fibration defining the connected topological graph Θ_z and $\Theta_{x,z}$ for all $x \in \mathbb{S}^1$; see Definition 6.18. Recall that the time 1 map ψ_1^z of the reparameterized flow is the first return map for Θ_z and is denoted

$$f_z: \Theta_z \longrightarrow \Theta_z.$$

From Corollary 6.20, f_z is a homotopy equivalence and $f_z(V\Theta_z) = V\Theta_z$.

By Corollary 6.21, the inclusion induces an isomorphism $\pi_1(\Theta_z) \cong \ker(u)$, which we use to identify $\ker(u) = \pi_1(\Theta_z)$. Then f_z induces the monodromy $(f_z)_* = \varphi_u \in \mathrm{Out}(\ker(u))$.

Thus f_z satisfies all the requirements of being a topological representative of φ_u (in the sense of Definition 2.16) except that we don't know yet that f_z is a topological

graph-map. We will show that f_z is a topological graph-map, and that, with respect to an appropriate linear structure on Θ_z, the map f_z is an expanding irreducible train track map. When φ is fully irreducible, we will find that f_z is moreover a weakly clean train track map. The fact that φ_u is fully irreducible will then follow from Theorem B.3.

We begin our analysis of f_z with the following technical lemma.

Lemma 7.2 *Suppose* $\alpha\colon (-\delta, \delta) \to \sigma$ *is an arc in the interior of a* 2–*cell* σ *of* X_f *transverse to flow* ψ_s. *Then the map*

$$\Psi_\alpha\colon (-\delta, \delta) \times \mathbb{R}_{\geq 0} \longrightarrow X_f$$

defined by $\Psi_\alpha(t, s) = \psi_s(\alpha(t))$ *is locally injective.*

Proof The 2–cell σ is a trapezoid whose bottom edge is an arc in a skew 1–cell of the old cell structure, which in turn lives in the strip $\{\psi_s(e_0) : 0 \leq s \leq 1\}$ over an edge e_0 of Γ_0. Since Γ_0 is a subdivision of Γ, we see that there is an edge $e \in E\Gamma$ so that $\alpha(-\delta, \delta)$ is contained in $\{\psi_s(e) : s \geq 0\}$. If we take a characteristic map $j\colon [0, 1] \to X_f$ for e, it therefore suffices to show that

$$\Psi_j\colon (0, 1) \times \mathbb{R}_{\geq 0} \longrightarrow X_f$$

is locally injective. For this, observe that by composing with η_0, it follows that if $\Psi_j(s, t) = \Psi_j(s', t')$ then $t - t'$ is an integer. Given a point (s_0, t_0), choose a neighborhood of the form $W = U \times (t_0 - \frac{1}{2}, t_0 + \frac{1}{2})$. If $\Psi_j(s, t) = \Psi_j(s', t')$ for $(s, t), (s', t') \in W$, then $t = t'$.

To find U, we proceed as follows. Let $T > t_0 + \frac{1}{2}$ be an integer. Since f is a train track map, there is a neighborhood U of s_0 so that $f^T|_{e(U)}$ is injective. Suppose $s, s' \in U$ and $|t - t_0| < \frac{1}{2}$ is such that $\Psi_j(s, t) = \Psi_j(s', t)$. Then $\psi_t(e(s)) = \psi_t(e(s'))$. Since $t < T$, we have $T - t > 0$ and the semiflow property implies

$$f^T(e(s)) = \psi_T(e(s)) = \psi_{T-t}(\psi_t(e(s)))$$
$$= \psi_{T-t}(\psi_t(e(s'))) = \psi_T(e(s')) = f^T(e(s')).$$

Therefore, $s = s'$ and Ψ_j is injective on W. Since (s_0, t_0) was arbitrary, Ψ_j is locally injective. □

Now suppose ψ_s^z is the reparameterizations of the semiflow ψ_s coming from a positive perturbation z of z_0. We can define a map $\Psi_\alpha^z(t, s) = \psi_s^z(\alpha(t))$ as in the lemma. Because ψ_s^z is a reparameterization of ψ_s, it follows that $\Psi_\alpha^z = \Psi_\alpha \circ H$ for some homeomorphism

$$H\colon (-\delta, \delta) \times \mathbb{R}_{\geq 0} \longrightarrow (-\delta, \delta) \times \mathbb{R}_{\geq 0}$$

(determined by the reparameterization). Thus, the conclusion of the lemma remains true for Ψ_z^z in place of Ψ_α. In particular, since edges of Θ_z are arcs in the interiors of 2–cells of X_f, this implies the following.

Lemma 7.3 *For each $n \geq 1$ the map $f_z^n \colon \Theta_z \to \Theta_z$ is locally injective on the interior of each edge of Θ_z. Moreover, f_z is a regular topological graph map.*

Proof The first statement follows immediately from the above discussion of the map Ψ_z^z. To see the second statement, first recall that by Corollary 6.20, we have $f_z(V\Theta_z) \subseteq V\Theta_z$. Let e be an edge of Θ_z and let $j \colon [0,1] \to \Theta_z$ be such that $j(0) = o(e)$, $j(1) = t(e)$ and $j|_{(0,1)}$ is a homeomorphism onto e. To show that f_z is a topological graph map, we must check that $f_z \circ j$ is a topological edge-path in Θ_z. The regularity of f_z will then follow immediately from local injectivity on edges.

The set $(f_z \circ j)^{-1}(V\Theta_z)$ of points that map to vertices of Θ_z is a closed, and hence compact, subset of $[0,1]$. Therefore, by the local injectivity of $f_z \circ j$ and finiteness of $V\Theta_z$, the set is finite and we may write its points as $0 = c_0 < \cdots < c_n = 1$. Now consider a subinterval $I = (c_{i-1}, c_i)$. Since $(f_z \circ j)(I)$ is disjoint from $V\Theta_z$, this image must be contained in a single edge e' of Θ_z. Moreover, the intermediate value theorem implies that $(f_z \circ j)|_I$ maps onto e'. Therefore, by the local injectivity, $(f_z \circ j)|_I$ is a homeomorphism onto e', as required. $\qquad \square$

We next desire to construct a linear structure on Θ_z and to establish the necessary combinatorial properties of f_z. To this end we introduce the following notation.

Each edge e of Θ_z lives in trapezoid of X_f whose bottom is an arc of a skew 1–cell of the old cell structure. Therefore, there exists an arc λ_e contained in an edge of Γ so that a reparameterization of the flow maps λ_e homeomorphically onto e. More precisely, there exists a continuous function $\tau_e \colon \lambda_e \to \mathbb{R}_{\geq 0}$ so that the map $\rho_e \colon \lambda_e \to X_f$ defined by $\rho_e(y) = \psi_{\tau_e(y)}(y)$ sends λ_e homeomorphically onto e. If we then choose an integer $m_e \geq \max\{\tau_e(\lambda_e)\}$ and define $s_e \colon e \to [0, m_e]$ by $s_e(x) = m_e - \tau_e(\rho_e^{-1}(x))$, we find that the map $\rho_e' \colon e \to X_f$ defined by $\rho_e'(x) = \psi_{s_e(x)}(x)$ maps e into $\Gamma_0 = \Gamma$ in such a way that

$$\lambda_e \xrightarrow{\;\rho_e\;} e \xrightarrow{\;\rho_e'\;} \Gamma$$

is just $f^{m_e}|_{\lambda_e}$.

Recall that by assumption $f \colon \Gamma \to \Gamma$ is an expanding irreducible train-track map. Thus its transition matrix $A(f)$ is irreducible and has $\lambda(f) > 1$.

Lemma 7.4 *The topological graph map $f_z \colon \Theta_z \to \Theta_z$ is expanding on all scales.*

Proof Consider any neighborhood $U \subset \Theta_z$ of a point $x \in \Theta_z$, and an edge $e \in E\Theta_z$. After restricting to a smaller open set we may assume that U is an open interval contained in an edge e_0 of Θ_z. As above, there are arcs λ_{e_0} and λ_e that get mapped homeomorphically onto e_0 and e by the flow. In particular λ_{e_0} contains an open subinterval I so that the restriction $\rho_{e_0}|_I : I \to U$ is a homeomorphism.

Let $\hat{e}$ denote the edge of Γ containing λ_e. By Lemma A.2, $f : \Gamma \to \Gamma$ is expanding on all scales. Therefore there is a subinterval $I' \subset I$ and a power $n \geq 1$ so that f^n maps I' homeomorphically onto $\lambda_e \subset \hat{e}$. In fact, by restricting to smaller neighborhoods, we can arrange for n to satisfy $n \geq m_{e_0}$. Let $U' = \rho_{e_0}(I') \subset U$. It follows that $\rho_e \circ f^{n-m_e} \circ \rho'_{e_0}$ restricts to a homeomorphism from $U' = \rho_{e_0}(I')$ to

$$\rho_e \circ f^{n-m_e} \circ \rho'_{e_0}(\rho_{e_0}(I')) = \rho_e \circ f^n(I') = \rho_e(\lambda_e) = e.$$

However, this map is just $f_z^k|_{U'}$ for some integer $k \geq 1$. $\square$

Lemmas 7.3 and 7.4 show that f_z satisfies the hypotheses of Theorem A.3. Therefore by Theorem A.3 and Theorem B(4) we conclude the following:

Corollary 7.5 *There is a linear structure Λ on Θ_z with respect to which $f_z \colon \Theta_z \to \Theta_z$ is a linear graph map with irreducible transition matrix $A(f_z)$ such that $\lambda(f_z) > 1$. Consequently, f_z is an expanding irreducible train-track map representing φ_u for $u = [z]$.* $\square$

We have thus established the conclusion of Theorem C(1). From this point on, we fix a linear structure Λ on Θ_z provided by Corollary 7.5.

Convention 7.6 From now on and until the end of this section, unless stated otherwise we further assume that the element $\varphi \in \mathrm{Out}(F_N)$ is hyperbolic and fully irreducible.

By Theorem B.3 we now know that f is a clean train track map. Thus the Whitehead graph $\mathrm{Wh}(f, v)$ associated to each vertex v of Γ is connected and $A(f)^m$ is positive for some $m \geq 1$. We now show that all Whitehead graphs of $f_z \colon \Theta_z \to \Theta_z$ are connected as well.

Recall that $\Gamma_t = \eta_0^{-1}(t)$ denotes the horizontal fiber of X_f at height $t \in [0, 1]$, and that $\Gamma_0 = \Gamma$.

Definition 7.7 (z–modified Whitehead graph) For each point $x \in \Gamma_t$ we define a corresponding z–*modified Whitehead graph* $\mathrm{Wh}^z(\Gamma_t, x)$ as follows. If x lies in the interior of an edge of Γ_t, then subdivide that edge at x. The vertices of $\mathrm{Wh}^z(\Gamma_t, x)$

are as usual the set of "directions" at x, that is the set of edges in the (subdivided) graph that originate at x. Two such edges e_1, e_2 are adjacent in $\mathrm{Wh}^z(\Gamma_t, x)$ if there exists an edge e of Θ_z and a power $n \geq 0$ so that the edge-path $\psi_t \circ f^n \circ \rho'_e(e)$ takes the corresponding turn e_1, e_2; in this case we call e_1, e_2 a z–*taken turn*.

Lemma 7.8 *The z–modified Whitehead graph $\mathrm{Wh}^z(\Gamma_t, x)$ is connected for all $t \in [0, 1]$ and $x \in \Gamma_t$.*

Proof We first claim that $\mathrm{Wh}^z(\Gamma_0, x)$ is connected for each $x \in \Gamma_0$. Indeed, for every edge e of Θ_z there exists an $n \geq 1$ so that the path $f^n \circ \rho'_e(e)$ crosses every edge of $\Gamma_0 = \Gamma$. This follows from the positivity of $A(f)^n$ for large n together with the facts that f is expanding on all scales and $\rho'_e(e)$ contains an open neighborhood in Γ. This shows that $\mathrm{Wh}^z(\Gamma_0, x)$ is connected whenever x lies in the interior of an edge of Γ_0. On the other hand, when x is a vertex of Γ_0 it is easy to see that $\mathrm{Wh}^z(\Gamma_0, x) = \mathrm{Wh}(f, x)$, which we already know is connected. Indeed, since $f^n \circ \rho'_e(e)$ crosses every edge of Γ_0, whenever $k \geq 1$ and $c \in E\Gamma_0$ are such that $f^k(c)$ takes a turn at x, then $f^{n+k} \circ \rho'_e(e)$ takes the same turn at x.

Now let $t > 0$ and let x be any point of Γ_t. Let $P = (\psi_t)^{-1}(x) \subset \Gamma_0$ be the preimage. Notice that ψ_t maps each z–taken turn at $p \in P$ to a z–taken turn at x; that is, ψ_t induces a map of z–modified Whitehead graphs $\psi_t^*: \mathrm{Wh}^z(\Gamma_0, p) \to \mathrm{Wh}^z(\Gamma_t, x)$ for each $p \in P$. Indeed, each edge e_1 originating at p is mapped to an edge $\psi_t(e_1)$ originating at x; and if $e \subset \Theta_z$ is an edge such that $f^n \circ \rho'_e(e)$ takes a turn e_1, e_2 at $p \in P$, then $\psi_t \circ f^n \circ \rho'_e(e)$ takes the turn $f_t(e_1), f_t(e_2)$ at x.

Let $T \subset X_f$ denote the union of the flow lines $\psi_{[0,t]}(p)$ for the points $p \in P$ (ie T is the set $\{\psi_s(p) \mid s \in [0, t], p \in P\}$). Then T is an oriented trivalent tree with $|P|$ leaves on the bottom and one point x at the top. Each point $y \in T$ lives in some Γ_s and has a corresponding z–modified Whitehead graph $\mathrm{Wh}^z(\Gamma_s, y)$. The fact that $\mathrm{Wh}^z(\Gamma_t, x)$ is connected will follow from the following claim.

Claim $\mathrm{Wh}^z(\Gamma_s, w)$ *is connected for each $w \in T$.*

We have already demonstrated this for the $|P|$ points at the base of the tree. If $w \in \Gamma_s$ lies in the interior of an edge of T, and $y \in \Gamma_r$, with $r < s$, lies in the same edge, then the flow ψ_{s-r} induces a quotient map $\mathrm{Wh}^z(\Gamma_r, y) \to \mathrm{Wh}^z(\Gamma_s, w)$ (the fact that y and w lie in the same edge of T implies that y does not get identified with any other points in the interval $[r, s]$). In particular, $\mathrm{Wh}^z(\Gamma_s, w)$ is connected provided $\mathrm{Wh}^z(\Gamma_r, y)$ is.

Now suppose that 2 upward-oriented edges b_1, b_2 of T meet at a point $w \in \Gamma_s$. Restricting to a sufficiently small time interval $[r, s]$, we may assume that there is

a constant z–modified Whitehead graph G_i along each b_i. By induction on the combinatorial distance to P within T we may suppose that each G_i is connected. Let $e_1, e_2 \in E\Gamma_r$ denote the two edges of Γ_r whose initial segments are folded together by $r_{s,r} \colon \Gamma_r \to \Gamma_s$. The closures of these edges necessarily contain the points $y_1, y_2 \in \Gamma_r \cap T$ for which $\psi_{s-r}(y_1) = \psi_{s-r}(y_2) = w$. The initial segment α_i of e_i (oriented from y to $o(e_i)$) then corresponds to a vertex in $G_i = \mathrm{Wh}^z(\Gamma_r, y_i)$. Since α_1 and α_2 are folded together by ψ_{s-r}, we see that the induced map ψ_{s-r}^* on z–modified Whitehead graphs identifies these two vertices. Therefore $\mathrm{Wh}^z(\Gamma_s, w)$ is the wedge of the connected graphs G_1 and G_1 and in particular is connected. □

Corollary 7.9 *The Whitehead graphs of all vertices of Θ_z with respect to $f_z \colon \Theta_z \to \Theta_z$ are connected.*

Proof Recall that $V\Theta_z$ is defined to be the set of intersection points of Θ_z with $X_f^{(1)}$, together with some additional degree-2 vertices obtained by subdividing at points of intersection of the flow lines through this first set of points; see Definition 6.18.

Let v be a vertex of Θ_z. Let us first suppose that v is not contained in a skew 1–cell of X_f. One may then find a neighborhood $U \subset X_f$ of v and a time t such that the flow ψ maps $\Gamma_t \cap U$ homeomorphically onto $\Theta_z \cap U$. Let x be the point of $\Gamma_t \cap U$ that flows onto v. From the definition of z–modified Whitehead graphs, we see that the flow induces a graph isomorphism $\mathrm{Wh}^z(\Gamma_t, x) \to \mathrm{Wh}_{\Theta_z}(f_z, v)$. Therefore $\mathrm{Wh}_{\Theta_z}(f_z, v)$ is connected by Lemma 7.8.

Now suppose that v lies on a skew 1–cell. In this case we can find a neighborhood $U \subset X_f$ of v so that $U \cap \Theta_z$ is a tree with one vertex v and 3 edges, and a time t so that $U \cap \Gamma_t$ consist of 2 disjoint edges. Moreover, each component of $U \cap \Gamma_t$ contains a point x_i that flows onto v. By Lemma 7.8, we know that each graph $\mathrm{Wh}^z(\Gamma_t, x_i)$ is connected (in this case just a single edge with two vertices). The flow now induces a quotient map

$$\mathrm{Wh}^z(\Gamma_t, x_1) \sqcup \mathrm{Wh}^z(\Gamma_t, x_2) \longrightarrow \mathrm{Wh}_{\Theta_z}(f_z, v)$$

that takes a vertex from each $\mathrm{Wh}^z(\Gamma_t, x_i)$ and identifies these to a single vertex in $\mathrm{Wh}_{\Theta_z}(f_z, v)$. Therefore $\mathrm{Wh}_{\Theta_z}(f_z, v)$ is the wedge of 2 line segments and is in particular connected. □

Corollary 7.10 *The map $f_z \colon \Theta_z \to \Theta_z$ is a weakly clean train-track map.*

Proof We already know, by Corollary 7.5, that $f_z \colon \Theta_z \to \Theta_z$ is an expanding train track map with an irreducible transition matrix $A(f_z)$. Corollary 7.9 implies that the Whitehead graphs $\mathrm{Wh}_{\Theta_z}(f_z, v)$ are connected for all $v \in V\Theta_z$. Thus f_z is a weakly clean train-track map. □

Proof of Theorem C Suppose first that $f\colon \Gamma \to \Gamma$ is a combinatorial expanding irreducible train-track map representing φ, without assuming that φ is hyperbolic and fully irreducible. Recall that $u \in \mathcal{A}$ is a primitive integral element, that $z \in Z^1(X_f; \mathbb{R})$ is a positive 1–cocycle representing the cohomology class u and that $f_z\colon \Theta_z \to \Theta_z$ induces the monodromy $(f_z)_* = \varphi_u \in \mathrm{Out}(\ker(u)) = \mathrm{Out}(\pi_1(\Theta_z))$. Lemma 7.3 shows that f_z is a topological graph-map and thus f_z is a topological representative of the automorphism φ_u of $\ker(u) = \pi_1(\Theta_z)$. Corollary 7.5 implies that there exists a linear structure Λ on Θ_z so that $f_z\colon \Theta_z \to \Theta_z$ is an expanding irreducible train-track map representing φ_u. This verifies part (1) of Theorem C.

Suppose now that φ is additionally assumed to be hyperbolic and fully irreducible, in which case G_φ is word-hyperbolic and f is a clean train-track representing φ. Since $G_{\varphi_u} \cong G_\varphi$ is word-hyperbolic, we also know that φ_u is a hyperbolic automorphism. By Corollary 7.10 we additionally have f_z is a weakly clean train track map. Hence φ_u is fully irreducible by Theorem B.3. □

Remark 7.11 The full irreducibility conclusion for φ_u of part (2) in Theorem C does not necessarily hold under the weaker assumption in part (2) that φ only be fully irreducible (ie without the assumption that φ be hyperbolic). Recall that, as proved in [9], a fully irreducible $\varphi \in \mathrm{Out}(F_N)$ is nonhyperbolic if and only if φ comes from a pseudo-Anosov homeomorphism f of a once-punctured connected surface S without boundary with $F_N = \pi_1(S)$. In this case the mapping torus M_f of f is a hyperbolic 3–manifold fibering over the circle with fiber S and $G_\varphi = \pi_1(M)$. If M admits a different fibration over $\mathbb{S}^1$ with fiber Y, then M_f also decomposes as the mapping torus of a homeomorphism f' of Y and this f' is again pseudo-Anosov. However, in general, there is no reason to expect that the new fiber Y will have exactly one puncture, even for a fibration of M coming from a class $u \in H^1(G_\varphi, \mathbb{R})$ which is projectively close to the class u_0 associated with the original fibration. If such Y has at least two punctures, then the element of $\mathrm{Out}(\pi_1(Y))$ corresponding to f' is not fully irreducible, since the punctures of Y are permuted by f' and the peripheral curves around the punctures are primitive elements of $\pi_1(Y)$.

8 Dynamics and entropy

In this section, we establish Theorem D from Section 1.

Theorem D Let $\varphi \in \mathrm{Out}(F_N)$ (where $N \geq 2$) be an element represented by an expanding irreducible train track map $f\colon \Gamma \to \Gamma$. Let X_f be a folded mapping torus, ψ the associated semiflow, and $\mathcal{A}_{X_f} \subset H^1(X_f; \mathbb{R})$ the positive cone of X_f. Then

there exists a continuous, homogeneous of degree -1 function $\mathfrak{H}\colon \mathcal{A}_{X_f} \to \mathbb{R}$ such that for every primitive integral $u \in \mathcal{A}_{X_f}$

$$\log(\lambda(\varphi_u)) = \mathfrak{H}(u).$$

Moreover, $1/\mathfrak{H}$ is concave, and hence $\mathfrak{H}$ is convex.

Our proof closely follows Fried's proof of the analogous statement for mapping tori of pseudo-Anosov homeomorphisms (which applies in a more general setting); see [29]. To carry out the arguments here, we work with the natural extension, as we now explain.

8.1   The natural extension of the first return map

Fix any integral class $u \in \mathcal{A}$ represented by a positive cocycle z and let $\eta_z\colon X = X_f \to \mathbb{S}^1$ be an associated fibration. Recall that Θ_z is the preimage of a point under η_z which we assume (as we may) is $0 \in \mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$, so that $\Theta_z = \eta_z^{-1}(0)$. Additionally, $f_z\colon \Theta_z \to \Theta_z$ denotes the first return map of ψ. The failure of $f_z\colon \Theta_z \to \Theta_z$ to be a homeomorphism can be rectified in a certain sense by the construction of the *natural extension*. This is the space

$$\Theta_z' = \{\underline{x} = (\dots, x_{-1}, x_0, x_1, \dots) \mid x_i \in \Theta_z \text{ such that } x_i = f_z(x_{i-1}) \text{ for all } i\}.$$

equipped with the product topology, together with the *shift map* $f_z'\colon \Theta_z' \to \Theta_z'$ defined by

$$(f_z'(\underline{x}))_i = x_{i+1}$$

for all $i \in \mathbb{Z}$. That is, the ith coordinate of $f_z'(\underline{x})$ is the $(i+1)^{\text{st}}$ coordinate of $\underline{x}$. There is a continuous surjective semiconjugacy $\pi_z\colon \Theta_z' \to \Theta_z$ given by

$$\pi_z(\dots, x_{-1}, x_0, x_1, \dots) = x_0.$$

That is, $\pi_z \circ f_z' = f_z \circ \pi_z$ (π_z is a semiconjugacy of f_z' onto f_z). This fact is immediate from the definition of Θ_z' and f_z'. Since π_z is surjective, the topological entropies satisfies the inequality

$$h(f_z') \geq h(f_z).$$

However, the system $f_z'\colon \Theta_z' \to \Theta_z'$ is the "smallest" homeomorphism which admits a semiconjugacy to $f_z\colon \Theta_z \to \Theta_z$, in the sense that any other homeomorphism which can be semiconjugated onto $f_z\colon \Theta_z \to \Theta_z$ factors through π_z via a surjection to Θ_z'. The key fact for us is that topological entropy is actually preserved in this case; see [22, Section 6.8]:

Lemma 8.1 *The topological entropy of the extension satisfies*

$$h(f_z') = h(f_z).$$

8.2 A natural extension of the semiflow

We can make a similar construction for the semiflow $\psi_s \colon X \to X$. Say that a map $\gamma \colon \mathbb{R} \to X$ is a *biinfinite flow line* for ψ if

$$\psi_s(\gamma(t)) = \gamma(t + s)$$

for all $t \in \mathbb{R}$ and $s \geq 0$. Let X' denote the space of all biinfinite flow lines, equipped with the topology of uniform convergence on compact sets:

$$X' = \{\gamma \mid \gamma \text{ is a biinfinite flow line}\}.$$

This comes equipped with a flow Ψ defined by $(\Psi_s(\gamma))(t) = \gamma(t + s)$ for all $t, s \in \mathbb{R}$. There is a map $\Pi \colon X' \to X$ given by $\Pi(\gamma) = \gamma(0)$ which is a continuous surjection (since there is a biinfinite flow line through every point). Again from the definitions, we see that for all $s \geq 0$ we have

$$\psi_s \circ \Pi = \Pi \circ \Psi_s$$

which is to say, Π semiconjugates Ψ to ψ. This is also clearly the minimal flow with this property.

By composing η_z with Π we obtain a map

$$\eta_z' = \eta_z \circ \Pi \colon X' \longrightarrow \mathbb{S}^1.$$

We put $\Theta_z'' := (\eta_z')^{-1}(0)$ and denote the first return map of Ψ to this section by $f_z'' \colon \Theta_z'' \to \Theta_z''$. As the next lemma shows, f_z'' is nothing but the natural extension of $f_z \colon \Theta_z \to \Theta_z$ already constructed above.

Lemma 8.2 *There is a homeomorphism $F \colon \Theta_z'' \to \Theta_z'$ that conjugates f_z'' to f_z'.*

Proof First observe that $\gamma \in X'$ is in Θ_z'' if and only if $\gamma(0) \in \Theta_z$. Such a biinfinite flow line $\gamma \colon \mathbb{R} \to X$ meets Θ_z at infinitely many points in both forward and backward time. More precisely, there is an increasing, biinfinite, proper sequence of times $\{t_j\}_{j \in \mathbb{Z}}$ with $t_0 = 0$ so that

$$(8\text{-}1) \qquad\qquad \gamma(t) \in \Theta_z \quad \Leftrightarrow \quad t = t_j \text{ for some } j.$$

Since f_z is the first return map of ψ to Θ_z, it follows that $f_z(\gamma(t_j)) = \gamma(t_{j+1})$ for all j. Therefore, for any $\gamma \in \Theta_z''$, we may define

$$F(\gamma) = (\ldots, \gamma(t_{-1}), \gamma(t_0), \gamma(t_1), \ldots)$$

(that is, the ith coordinate of $F(\gamma)$ is $\gamma(t_i)$), which is an element of Θ_z'.

To see that F is injective, we claim that the sequence

$$(\ldots, \gamma(t_{-1}), \gamma(t_0), \gamma(t_1), \ldots)$$

determines γ. This is because, for any $t > s$, $\gamma(s)$ determines $\gamma(t)$ by the formula $\gamma(t) = \psi_{t-s}(\gamma(s))$. As $j \to -\infty$, $t_j \to -\infty$, and so for any $t \in \mathbb{R}$, $\gamma(t)$ is determined by some $\gamma(t_j)$ with $t_j < t$. Therefore, F is injective. Similarly, given any sequence $\underline{x} \in \Theta_z'$, we can use this reasoning to construct $\gamma \in \Theta_z''$ with $F(\gamma) = \underline{x}$.

Suppose now that $\gamma_n \to \gamma$ is a sequence of flow lines converging uniformly on compact sets with $\gamma(0), \gamma_n(0) \in \Theta_z$ for all n and $\gamma_n(0) \to \gamma(0)$. Let $\{t_j\}_{j \in \mathbb{Z}}$ denote the sequence for γ as above, and similarly, $\{t_{n,j}\}_{j \in \mathbb{Z}}$ the sequence for γ_n. The uniform convergence on compact sets implies that the points of intersection of γ_n with Θ_z are converging uniformly on compact sets to the intersection points of γ with Θ_z, which is to say, $\gamma_n(t_{n,j}) \to \gamma(t_j)$ for all j as $n \to \infty$. Therefore F is continuous. Because X is compact, one can show that X' is also compact, hence so is Θ_z''. Since Θ_z' is Hausdorff, the continuous map $F \colon \Theta_z'' \to \Theta_z'$ is a homeomorphism.

For any $\gamma \in \Theta_z''$ and any j, we have $f_z(\gamma(t_j)) = \gamma(t_{j+1})$, and therefore F conjugates f_z'' to the shift map f_z'. $\qquad\square$

Because of this lemma, we will ignore the distinction between $f_z'' \colon \Theta_z'' \to \Theta_z''$ and $f_z' \colon \Theta_z' \to \Theta_z'$, and just refer to Θ_z' as the section of Ψ and $f_z' \colon \Theta_z' \to \Theta_z'$ the first return map of Ψ.

8.3 The variational principal and Abramov's theorem

The topological entropy of f_z' is approximated from below by metric entropies $h_\nu(f_z')$ for invariant probability measures ν on Θ_z', according to the *variational principle*; see [22, Section 6.8].

Theorem 8.3 *For any integral class $u \in \mathcal{A}$ represented by a positive cocycle z we have*

$$h(f_z') = \sup_\nu h_\nu(f_z'),$$

where ν ranges over all f_z'–invariant Borel probability measures on Θ_z'.

The key to proving continuity (and the other properties) of the entropy is a result of Abramov which relates entropy of f_z' to that of the flow Ψ. To make sense of this, we have to relate f_z–invariant measures on Θ_z' to Ψ–invariant measures on X'. In fact, there is a bijection between these two spaces of measures defined as follows.

Given a Ψ–invariant probability measure μ on X', we can construct a measure $\bar{\mu}_0$ on any section Θ_z' by the formula

$$(8\text{-}2) \qquad \bar{\mu}_0(A) = \frac{1}{t}\mu(\{\psi_s(x) \mid x \in A, 0 \leq s \leq t\})$$

for any Borel measurable set $A \subset \Theta_z'$. This may not be a probability measure, and indeed the total measure $\bar{\mu}_0(\Theta_z')$ is an important quantity. The f_z'–invariant probability measure on Θ_z is the normalization

$$\bar{\mu} = \frac{1}{\bar{\mu}_0(\Theta_z')}\bar{\mu}_0.$$

There is a flow-equivariant local homeomorphism

$$\Theta_z' \times \mathbb{R} \longrightarrow X'$$

given by $(\underline{x}, s) \mapsto \Psi_s(\underline{x})$ (where the flow on the domain is just the natural translation action on the second coordinate). The product measure $\bar{\mu}_0 \times m$, where m is Lebesgue measure on $\mathbb{R}$, is flow-invariant and pushes forward to a flow invariant measure. From the construction of $\bar{\mu}_0$, it follows that μ is the push forward of $\bar{\mu}_0 \times m$. As an abuse of notation, we will just write $\mu = \bar{\mu}_0 \times m$. This can be used to describe the inverse construction.

Abramov's theorem relating the entropy of a flow to the first return map of a section (see [1; 18]) takes the following form in our situation (see also [29]).

Theorem 8.4 *For any integral class $u \in \mathcal{A}$ represented by a positive cocycle z and any Ψ–invariant probability measure μ on X, let $\bar{\mu}_0$ and $\bar{\mu}$ be defined as above. Then*

$$h_{\bar{\mu}}(f_z') = \frac{h_\mu(\Psi)}{\bar{\mu}_0(\Theta_z')}.$$

As an immediate corollary of these two results, we have the following useful expression for the reciprocal of the topological entropy.

Corollary 8.5 *For any integral class $u \in \mathcal{A}$ represented by a positive cocycle z, we have*

$$\frac{1}{h(f_z')} = \inf_\mu \frac{\bar{\mu}_0(\Theta_z')}{h_\mu(\Psi)},$$

where the infimum is over all Ψ–invariant Borel probability measures μ on X', and $\bar\mu_0$ is constructed from μ as above.

8.4 The measure $\bar\mu_0$ and linearity

Given a Ψ–invariant probability measure μ on X', we now have a function defined on the integral classes $u \in \mathcal{A}$ which we denote

$$\mu_*(u) = \bar\mu_0(\Theta'_{z_u}),$$

where z_u is a positive cocycle representing u. If z'_u is another positive cocycle representing u with fiber $\Theta'_{z'_u}$, then we can flow Θ'_{z_u} homeomorphically onto $\Theta'_{z'_u}$, and hence

$$\bar\mu_0(\Theta'_{z'_u}) = \bar\mu_0(\Theta'_{z_u})$$

is independent of the choice of z_u (and fibration η_{z_u}).

Our goal is now to prove the following.

Lemma 8.6 *If μ is a Ψ–invariant probability measure on X', then μ_* is the restriction of a linear function (of the same name):*

$$\mu_* \colon H^1(X;\mathbb{R}) \longrightarrow \mathbb{R}.$$

Proof Given an integral class $u \in \mathcal{A}$ represented by a positive cocycle $z = z_u$, we can reparameterize the flow as Ψ^z, so that for the projection of the flow lines to $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ by η'_z are local isometries. The bijections

$$\{\Psi\text{–invariant measures on } X'\} \leftrightarrow \{f'_z\text{–invariant measures on } \Theta'_z\}$$
$$\leftrightarrow \{\Psi^z\text{–invariant measures on } X'\}$$

will allow us to turn the Ψ–invariant measure μ on X' into a Ψ^z–invariant measure μ^z on X', given by

$$\mu_z = \bar\mu_0 \times m_z,$$

where m_z is the Lebesgue measure on $\mathbb{R}$, and the subscript z is only there to remind us that it is induced by the reparameterized flow Ψ^z.

The measure μ_z is *not* a probability measure. Indeed, the parameterization is chosen so that the map

$$\Theta'_z \times [0,1) \longrightarrow X', \quad (\underline{x}, s) \mapsto \Psi^z_s(\underline{x})$$

is 1–1 and onto. It follows that

$$\mu_z(X') = \bar\mu_0(\Theta'_z).$$

What is important about this formula is that the left-hand side does not involve the section Θ'_z. Moreover, we could have used a different section Θ'_w, and as long as we use the flow Ψ^z, we would produce the same measure μ_z.

Now suppose that $u, v \in \mathcal{A}$ are two integral classes represented by the positive cocycles z and w, respectively. Then $u + v$ is represented by $z + w$, which is also positive. A useful observation is that because η_z and η_w are maps to the *additive group* $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$, we can add them to produce a map $\eta_z + \eta_w$. Each of η_z and η_w satisfy Lemma 6.16, and without loss of generality, the finite cover $\mathbb{M}$ by local models in the statement is the same for both η_z and η_w. Indeed, we may always pass to a refinement of the cover without affecting the validity of the lemma, and we will freely do so without further mention. In addition, by adding a constant in $\mathbb{S}^1$ to each of η_z and η_w, we may assume that the point $x_0 \in \mathbb{S}^1$ is the same in both, and is the identity $0 \in \mathbb{S}^1$.

Setting $\eta_{z+w} = \eta_z + \eta_w$, it is straightforward to see that this map satisfies properties (1), (2) and (3) of Lemma 6.16 for the cocycle $z + w$ representing the class $u + v$. It may not satisfy property (4), however. To address this issue, we proceed as follows. Note that $\Xi = \eta_{z+w}^{-1}(0) \cap X_f^{(1)}$ is the finite set of points $\xi \in X_f^{(1)}$ with the property that $\eta_z(\xi) = -\eta_w(\xi)$. Then $\Xi = \Xi_0 \cup \Xi_1$ where $\eta_z(\xi) = 0 = \eta_w(\xi)$ for $\xi \in \Xi_0$ and $\eta_z(\xi) \neq 0 \neq \eta_w(\xi)$ for $\xi \in \Xi_1$. Consequently, $\Xi_0 \subset \Theta_z \cap \Theta_w$ and $\Xi_1 \cap \Theta_z = \varnothing$. After modifying η_z by an arbitrarily small homotopy if necessary (which is the identity on Θ_z), we may assume that $\Xi_1 \subset X_f^{(1)} \setminus X_f^{(0)}$ and that this set flows into the union of vertical 1–cells. Since Ξ_0 already has this property (being contained in $\Theta_z \cap \Theta_w$), it follows that η_{z+w} satisfies all properties of Lemma 6.16.

The map $\eta_{z+w} = \eta_z + \eta_w$ determines a corresponding map

$$\eta'_{z+w} = \eta'_z + \eta'_w \colon X' \longrightarrow \mathbb{S}^1.$$

The measure μ_z and μ_w are defined by taking the product $\bar{\mu}_0$ with the Lebesgue measures m_z and m_w, induced by the reparameterized flows Ψ^z and Ψ^w, respectively. Since these flows are parameterized so that η'_z and η'_w restricted to flow lines are local isometries, it follows that μ_{z+w} is defined by taking the product $\bar{\mu}_0$ with $m_z + m_w$. Consequently, we have

$$\mu_{z+w} = \mu_z + \mu_w.$$

Hence

$$\mu_*(u + v) = \mu_{z+w}(X') = \mu_z(X') + \mu_w(X') = \mu_*(u) + \mu_*(v)$$

and μ_* is linear on the integral points of $\mathcal{A}$. It follows that μ_* uniquely extends to a linear function on $H^1(X; \mathbb{R})$. $\qquad \square$

8.5 Completing the proof

Proof of Theorem D If $u \in \mathcal{A}$ is a primitive integral element and z_u is a positive 1–cocycle, then, by Corollary 6.20 the map $f_{z_u} \colon \Theta_{z_u} \to \Theta_{z_u}$ is a homotopy equivalence taking vertices to vertices. By Corollary 6.21 we also know that Θ_{z_u} is connected and furthermore, it has no valence-1 vertices (see the proof of Theorem B). By Lemma 7.3 the map $f_{z_u} \colon \Theta_{z_u} \to \Theta_{z_u}$ is a regular topological graph-map. Corollary 6.21 now implies that f_{z_u} is a regular topological representative of the monodromy $\varphi_u \in \mathrm{Out}(\ker(u))$. By Proposition A.1 we have $h(f_{z_u}) = \log \lambda(f_{z_u})$ where $\lambda(f_{z_u})$ is the spectral radius of the transition matrix $A(f_{z_u})$. We have also seen in the proof of Theorem C that for an appropriate choice of a linear structure on Θ_{z_u} the map $f_{z_u} \colon \Theta_{z_u} \to \Theta_{z_u}$ is an expanding irreducible train-track map representing φ_u. Therefore $\lambda(f_{z_u}) = \lambda(\varphi_u)$ is the growth rate of φ_u and thus $h(f_{z_u}) = \log \lambda(f_{z_u}) = \log \lambda(\varphi_u) = h(\varphi_u)$, the algebraic entropy of φ_u. Since φ_u is fully irreducible, we have $h(f_{z_u}) = h(\varphi_u) > 0$.

Define a function $W \colon \mathcal{A} \to \mathbb{R}$ by

$$W(u) := \inf_{\mu} \frac{\mu_*(z_u)}{h_\mu(\Psi)},$$

where $u \in \mathcal{A}$, where z_u is a positive cocycle representing u, and the infimum is over all Borel Ψ–invariant probability measures μ on X'. By Lemma 8.1 we have $W(u) = 1/h(f'_{z_u}) = 1/h(f_{z_u})$ for any primitive integral class $u \in \mathcal{A}$. The linearity of μ_* implies that the function W is homogeneous of degree 1. By Corollary 8.5, we see that W is the infimum of a collection of linear functions, and is therefore concave and hence continuous. For any primitive integral $u \in A$ we have $0 < h(f_{z_u}) < \infty$ and hence $0 < W(u) = 1/h(f_{z_u}) < \infty$. Since the primitive integral classes are projectively dense in $\mathcal{A}$, concavity and homogeneity of W imply that $0 < W(u) < \infty$ for all $u \in \mathcal{A}$. Therefore setting $\mathfrak{H}(u) := 1/W(u)$ for all $u \in \mathcal{A}$ gives us a continuous homogeneous of degree -1 function $\mathfrak{H} \colon \mathcal{A} \to (0, \infty)$. By construction, for every primitive integral $u \in \mathcal{A}$ we have $\mathfrak{H}(u) = \log \lambda(f_{z_u}) = \log \lambda(\varphi_u) = h(\varphi_u)$, as required. □

As mentioned in the introduction, we have the following analogue from the 3–manifold setting of McMullen's [45, Section 10] which complements the results of Algom-Kfir and Rafi [4].

Corollary 8.7 Let $\varphi \in \mathrm{Out}(F_N)$ (where $N \geq 2$) be an element represented by an expanding irreducible train track map $f \colon \Gamma \to \Gamma$, and let X_f, $\mathcal{A}_{X_f}$, and $\mathfrak{H}$ be as in Theorem D. The function

$$D \colon \mathcal{A} \longrightarrow \mathbb{R}_+$$

given by $D(u) = \mathfrak{H}(u)u(\epsilon)$ is continuous and constant on rays. In particular, for any compact subset $K \subset \mathcal{A}$, D is bounded on the cone over K,

$$C_K = \{c\tau \mid c \in \mathbb{R}_+, \, \tau \in K\}.$$

Consequently, if $\dim(H^1(G; \mathbb{R})) \geq 2$, then $G \cong F_{N_k} \rtimes_{\varphi_k} \mathbb{Z}$ for some sequence of automorphisms $\varphi_k \in \mathrm{Out}(F_{N_k})$, with $N_k \to \infty$, represented by expanding irreducible train track maps, and there are constants $0 < c_1 < c_2 < \infty$ so that

$$c_1 \leq N_k \log \lambda(\varphi_k) \leq c_2.$$

If φ is fully irreducible and hyperbolic so are all φ_k.

Proof Continuity follows from linearity of $u \mapsto u(\epsilon)$ and continuity of $\mathfrak{H}$. The fact that D is constant on rays comes from the canceling homogeneity degrees: $+1$ for ϵ and -1 for $\mathfrak{H}$. The second statement follows from the fact that $D(K) = D(C_K)$, and K is compact. The existence of φ_k and constants $0 < c_1 < c_2 < \infty$ follows from the first statement and Theorems B and C by taking any sequence $\{u_k\} \subset \mathcal{A}$ of primitive integral elements for which the corresponding rays $\mathbb{R}_+ u_k$ converge to a point in $\mathcal{A}$. The last statement is a consequence of Theorem C, part (3). $\square$

Remark 8.8 It is possible to avoid the use of Theorem A.3 in the proof of Theorem D and in Theorem C(2).

Namely, suppose $f \colon \Gamma \to \Gamma$ is a regular topological graph-map of a finite connected topological graph Γ with an irreducible transition matrix $A(f)$. It is straightforward to find a metric structure $\mathcal{L}$ on Γ and a regular linear graph map $f' \colon (\Gamma, \mathcal{L}) \to (\Gamma, \mathcal{L})$ such that f' is isotopic to f and has "the same combinatorics" as f and such that f' is a local $\lambda(f)$–homothety on edges with respect to $d_{\mathcal{L}}$. By "the same combinatorics" we mean that for any $e \in E\Gamma$ and $n \geq 1$ we have $f^n(e) = (f')^n(e)$ as combinatorial edge-paths. Thus $A(f) = A(f')$, $\lambda(f) = \lambda(f')$, and a turn in Γ is taken by f if and only if it is taken by f'. Hence $\mathrm{Wh}_\Gamma(f, v) = \mathrm{Wh}_\Gamma(f', v)$ for any $v \in V\Gamma$. We have $h(f) = h(f') = \log \lambda(f)$ by Proposition A.1. Also, if f^n is locally injective on edges for all $n \geq 1$ then $(f')^n$ is locally injective on edges for all $n \geq 1$; that is, f' is a train-track map.

In the proof of Theorem C, we can first verify that $f_z^n \colon \Theta_z \to \Theta_z$ is locally injective on edges for all $n \geq 1$ and that $A(f_z)$ is irreducible with $\lambda(f_z) > 1$, and that all Whitehead graphs of all the vertices of Θ_z with respect to f_z are connected. We can then pass to the modified isotopic linear graph-map $f_z' \colon \Theta_z \to \Theta_z$ constructed as above. Then it will follow that f_z' is a (weakly) clean train-track map representing the same outer automorphism φ_u of $\pi_1(\Theta_z) = \ker(u)$ as f_z. Hence φ_u is fully irreducible by

Theorem B.3, yielding Theorem C. Moreover, since $h(f) = h(f') = \log \lambda(f) = h(\varphi_u)$, the proof of Theorem D would go through without appealing to Theorem A.3 as well. We included the proof of Theorem A.3 since it appears to be of independent interest and may be useful in the further analysis of the properties of the monodromies φ_u. Moreover, Theorem C(1), which does require the use of Theorem A.3, provides a stronger analogy with the motivating results of Fried for 3–manifolds.

Appendix A: Entropy and expansion

A.1: Entropy

The following statement about computing topological entropy of topological graph-maps appears to be a known "folklore" fact for the setting of linear graph-maps with an irreducible transition matrix, although we were unable to find an explicit reference for it in the literature. In that context, this fact is a straightforward corollary of Theorem 7.1 in the paper [12] of Bowen (using the "canonical eigenmetric" on the graph provided by Corollary A.7 below) which allows one to compute the topological entropy by counting n–periodic points of f as $n \to \infty$. The version we provide below is more general and applies to topological graph-maps that may be the identity map on some small nondegenerate subintervals of some edges (thus f could have continuumly many fixed points, in which cases Bowen's result is not applicable) and also which do not necessarily have irreducible transition matrix (so that a "uniformly expanding" metric on the graph as required by Bowen's theorem is not necessarily available even in the linear-graph context).

Proposition A.1 *Let Γ be a finite connected topological graph with at least one edge and let $f: \Gamma \to \Gamma$ be regular topological graph-map. Let $h(f)$ be the topological entropy of f. Then $h(f) = \log \lambda(f)$ (recall that $\lambda(f)$ is the spectral radius of $A(f)$).*

Proof This statement follows, for example, by a direct application of [5, Lemma 3.2]. In the notations of [5], for any $x \in \Gamma$ the set $f^{-1}(x)$ is finite and we have $K_f(x) = \{x\}$. With $A = V\Gamma$, the triple (Γ, A, f) satisfies the condition from [5] of being a "monotone representative of an action". In the notations of [5], the partition $\mathcal{A}_f$ of Γ consists of all the vertices and open 1–cells of Γ and the partition $\mathcal{A}_f^n$ of Γ consists of all the f^n–preimages of vertices and of all the connected components of the f^n–preimages of open 1–cells. Thus, up to an additive constant independent of n, the cardinality $\#(\mathcal{A}_f^n)$ of the partition $\mathcal{A}_f^n$ is equal to $\sum_{e \in E_+\Gamma} |f^n(e)|$ (where $E_+\Gamma$ comes from any choice an orientation on Γ). In other words, $\#(\mathcal{A}_f^n)$ is equal, up to a finite additive error independent of n, to the sum of all the entries of the matrix $(A(f))^n$.

By [5, Lemma 3.2] we have

$$(*) \qquad\qquad h(f) = \lim_{n\to\infty} \frac{1}{n} \log \#(A_f^n).$$

The assumptions on f imply that $\lambda(f) \geq 1$. If $\lambda(f) = 1$ then the sum of all the entries of $(A(f))^n$ grows polynomially in n and equation $(*)$ yields $h(f) = 0$. Thus we have $h(f) = 0 = \log 1 = \log \lambda(f)$, as required.

If $\lambda > 1$, then there is a constant $c \geq 1$ independent of n such that for every $n \geq 1$ we have $\frac{1}{c}\lambda(f)^n \leq \#(A_f^n) \leq c\lambda(f)^n$. Then by $(*)$ we again have $h(f) = \log \lambda(f)$, as required. $\qquad\qquad\square$

A.2: Expansion and promotion of topological graph-maps

Given a topological graph-map $f\colon \Gamma \to \Gamma$, we say that f is *expanding on all scales* if for every $x \in \Gamma$, every neighborhood U of x, and every edge e of Γ, there exists a positive integer m and an open interval $W \subset U$ so that f^m maps W homeomorphically onto e. Observe that as an immediate consequence, $f(e)$ is always a nondegenerate combinatorial edge-path (it cannot be constant on any open set).

Let us first observe that for graph maps, expanding on all scales is a rather mild condition that can be deduced from the combinatorial data.

Lemma A.2 *Let $f\colon \Gamma \to \Gamma$ be a graph map of a finite (linear) graph Γ. Suppose that f is expanding and that $A(f)$ is irreducible. Then f is expanding on all scales.*

Proof The graph Γ comes equipped with a linear structure Λ. For each edge e and any chart $(e_\alpha, j_\alpha, [a_\alpha, b_\alpha])$ with $e_\alpha = e$, the composition $f \circ j_\alpha\colon [a_\alpha, b_\alpha] \to \Gamma$ is a PL edge-path. Thus if $e_1 \cdots e_n$ is the combinatorial edge-path $f(e)$ then there is a partition $a_\alpha = c_0 < c_1 \cdots < c_n = b_\alpha$ of $[a_\alpha, b_\alpha]$ so that $f \circ j_\alpha$ restricts to a homeomorphism of (c_{i-1}, c_i) onto e_i for each $i = 1, \ldots, n$. We then define the *subdivision ratios* of the edge e to be the numbers $\rho_i^e = (c_i - c_{i-1})/(b_\alpha - a_\alpha)$ for $i = 1, \ldots, n$. These numbers all satisfy $0 < \rho_i^e < 1$ unless $n = 1$. Notice that since the transition functions between charts are affine, these numbers do not depend on the choice of the chart $j_\alpha\colon [a_\alpha, b_\alpha] \to \Gamma$. Since Γ has finitely many edges, there are only finitely many subdivision ratios, and we let ρ denote the maximal ratio that is strictly less than 1.

For an edge $e_0 \in E\Gamma$ we will now analyze what happens to e_0 under iteration by f. By the maximality condition for linear structures, Λ contains a chart of the form $(e_0, j, [0, 1])$. For $k \geq 0$ the set $j^{-1} \circ f^{-k}(V\Gamma)$ partitions $[0, 1]$ into subintervals, each of which is mapped homeomorphically onto an edge of Γ by f^k. Let $\mathcal{S}_k$ denote

the set of intervals in this subdivision. Suppose that an interval $I \in \mathcal{S}_k$ is mapped homeomorphically onto $e \in E\Gamma$ by f^k. Then I becomes further subdivided in $\mathcal{S}_{k+1}$ into subintervals of lengths $\text{len}(I)\rho_1^e, \ldots, \text{len}(I)\rho_{|f(e)|}^e$. In particular, if $|f(e)| \geq 2$, then each of these subintervals has length at most $\text{len}(I)\rho$.

Let l_k denote the maximal length of any subinterval in $\mathcal{S}_k$. Notice that, since intervals are only ever subdivided (ie $f^{-k}(V\Gamma) \subset f^{-k-1}(V\Gamma)$), l_k is nonincreasing. We claim that l_k in fact tends to 0 as $k \to \infty$. Indeed, since f is expanding, there exits $M \geq 1$ so that $|f^M(e)| \geq 2$ for every $e \in E\Gamma$ (notice that $|f^n(e)|$ is nondecreasing function of n). It follows that every interval in $\mathcal{S}_k$ becomes subdivided in $\mathcal{S}_{k+M}$. Therefore $l_{k+M} \leq l_k\rho$ for all $k \geq 0$ and we conclude $l_k \to 0$.

We now prove that f is expanding on all scales. Given $x \in \Gamma$ and a neighborhood U of x, we may find an open interval $U' \subset U$ that is contained in some edge e_0 of Γ. The above discussion shows that for some sufficiently large k, U' contains an interval $W' \subset U'$ that is mapped homeomorphically over some edge e_1. Let $e \in E\Gamma$ be any edge. Since $A(f)$ is irreducible, there exists a power n so that the combinatorial edge-path $f^n(e_1)$ crosses e or e^{-1}. Therefore, there is a further subinterval $W \subset W'$ so that $f^{k+n}(W)$ is mapped homeomorphically onto e. $\qquad\square$

While a topological graph map is always homotopic to a graph map with respect to a linear structure, the next theorem says that if f is expanding on all scales, then f is already linear with respect to some linear structure.

Recall that if Γ is a topological graph and $e \in E\Gamma$ then $\bar{e}$ denotes the closure of e in Γ.

Theorem A.3 *Suppose $f\colon \Gamma \to \Gamma$ is a regular topological graph-map that is expanding on all scales. Choose an orientation on Γ and an ordering $E_+\Gamma = \{e_1, \ldots, e_n\}$.*

Then the topological graph-map f is expanding, the matrix $A(f)$ is irreducible with the Perron–Frobenius eigenvalue $\lambda = \lambda(f) > 1$, and there exists a unique volume-1 metric structure $\mathcal{L}$ on Γ with the following properties:

(1) *$f\colon \Gamma \to \Gamma$ is linear with respect to $\mathcal{L}$.*

(2) *f is a local λ–homothety, meaning that the norm of the derivative of f at every point of $\Gamma \setminus f^{-1}(V\Gamma)$ is λ.*

(3) *For every $i = 1, \ldots, n$ the $d_\mathcal{L}$–length of e_i is v_i, where $\boldsymbol{v} = (v_1, \ldots, v_n)^T$ is the left Perron–Frobenius eigenvector (ie for the left action of $A(f)$ on column vectors) with $\sum_{i=1}^n v_i = 1$.*

Proof We assume for simplicity that Γ has no loop-edges. It will be clear from the proof how to adjust the argument to cover the case where some loop-edges are present.

We will first establish the existence of a metric $d_{\mathcal{L}}$ satisfying conditions (1) and (2) of the theorem. The assumptions on f imply that $A(f)$ is irreducible. Indeed, given any two edges $e_i, e_j \in E_+\Gamma$, the edge e_j is a neighborhood of any point $x \in e_j$, and as such there exists m and an open interval W in e_j so that f^m maps W homeomorphically onto e_i. Therefore, there is at least one occurrence of e_i or e_i^{-1} in $f^m(e_j)$, and hence $(A(f)^m)_{ij} \neq 0$.

Clearly, $A(f)$ is not a permutation matrix and thus the topological graph-map f is expanding.

Now let $\Gamma_0 = \Gamma$ and for $k \geq 1$ define Γ_k to be the graph obtained by subdividing Γ_{k-1} along $f^{-1}(V\Gamma_{k-1})$. By construction $V\Gamma_{k-1} \subset V\Gamma_k$ for all $k \geq 0$. Furthermore, since $f(V\Gamma) \subset V\Gamma$, it follows that $V\Gamma_k = (f^k)^{-1}(V\Gamma)$. Given any edge $e_{k,0} \in E\Gamma_k$ for any $k \geq 0$, the subdivision makes $e_{k,0}$ into a combinatorial edge-path $\gamma_{e_{k,0}} = e_{k+1,1}e_{k+1,2}\cdots e_{k+1,r}$ in Γ_{k+1}. Also, for every $k \geq 1$, we have $f(e_{k,0})$ is an edge $e_{k-1,0} \in E\Gamma_{k-1}$.

Claim A.4 *The set*

$$\mathcal{V} = \bigcup_{k=0}^{\infty} V\Gamma_k$$

is dense in Γ.

Proof of claim Given any $x \in \Gamma$ and any neighborhood U, we need only prove that $V\Gamma_k \cap U \neq \varnothing$ for k sufficiently large. Let e be any edge of Γ and let m and $W \subset U$ be as in the definition of f being expanding on all scales. Without loss of generality, we can assume that $\overline{W} \subset U$ (by replacing U with a strictly smaller neighborhood before finding m and W). Then $f^m(\overline{W}) = \overline{e}$, and hence $(f^m)^{-1}(o(e) \cup t(e)) \subset U$. Therefore $(f^m)^{-1}(V\Gamma) = V\Gamma_m$ nontrivially intersects U, and we set $k = m$. □

Claim A.5 *There exists a sequence of length functions $\{L_k \colon E\Gamma_k \to \mathbb{R}_+\}_{k=0}^{\infty}$ with the following two properties.*

(1) *For every edge $e_{k,0} \in E\Gamma_k$, with $k \geq 0$, if $\gamma_{e_{k,0}} = e_{k+1,1}e_{k+1,2}\cdots e_{k+1,r}$ is the combinatorial edge-path representing $e_{k,0}$ in $E\Gamma_{k+1}$, then*

$$L_k(e_{k,0}) = \sum_{\ell=1}^{r} L_{k+1}(e_{k+1,\ell}).$$

(2) For every edge $e_{k,0} \in E\Gamma_k$, with $k \geq 1$, if we write $f(e_{k,0}) = e_{k-1,0}$, then

$$\lambda L_k(e_{k,0}) = L_{k-1}(e_{k-1,0}).$$

Proof of claim   To construct this sequence of length functions, we define L_0 first to be given by $L(e_i) = L(e_i^{-1}) = v_i$, where $\boldsymbol{v} = (v_1, \ldots, v_n)$ is the Perron–Frobenius eigenvector for λ. Then for $k \geq 1$ define L_k in terms of L_{k-1} inductively by the second condition. Equivalently, we note that for every $k \geq 1$ and every edge $e_{k,0} \in E\Gamma_k$, we have that f^k maps $e_{k,0}$ homeomorphically onto some edge $e_i \in E\Gamma$ and so its L_k–length is $\lambda^{-k} v_i$.

It remains to verify that the first property holds. Fix $e_{k,0}$ and we let i be as above so that f^k maps $e_{k,0}$ homeomorphically onto e_i. The next iteration of f maps e_i over each edge e_j exactly a_{ij} times and we have

(A-1) $$\lambda v_i = \sum_{i=1}^{n} a_{ij} v_j.$$

On the other hand, writing

$$e_{k,0} = e_{k+1,1} e_{k+1,2} \cdots e_{k+1,r}$$

and letting $j(\ell)$ be such that f^{k+1} maps $e_{k+1,\ell}$ homeomorphically onto $e_{j(\ell)}$, it follows that

$$f(e_i) = e_{j(1)} e_{j(2)} \cdots e_{j(r)}.$$

and so comparing with (A-1) we have

$$\lambda v_i = \sum_{\ell=1}^{r} v_{j(\ell)}.$$

The L_{k+1}–length of $e_{k+1,\ell}$ is $\lambda^{-(k+1)} v_{j(\ell)}$ and therefore we have

$$L_k(e_{k,0}) = \lambda^{-k} v_i = \lambda^{-k} \left(\lambda^{-1} \sum_{\ell=1}^{r} v_{j(\ell)} \right)$$

$$= \lambda^{-(k+1)} \sum_{\ell=1}^{r} v_{j(\ell)} = \sum_{\ell=1}^{r} L_{k+1}(e_{k+1,\ell}). \qquad \square$$

We now prove that there exists a metric structure with the required properties. As a consequence of the second condition, we note that the L_k–length of any edge of $E\Gamma_k$ is at most λ^{-k} times the maximum L_0–length of the edges in $E\Gamma_0 = E\Gamma$. In particular, the maximum L_k–length of any edge in $E\Gamma_k$ tends to zero as $k \to \infty$.

Now suppose that $e \in E_+\Gamma$ and that $x, y \in \mathcal{V} \cap \bar{e}$ are any two points. We define a "distance in e" $d_{\mathcal{V},e}(x, y)$ between x and y to be the sum of the L_k–length of the edges between x and y inside e, provided $x, y \in V\Gamma_k$ (which is valid for all sufficiently large k, depending on x and y). By the first condition, this definition of $d_{\mathcal{V},e}(x, y)$ is independent of k and hence we get a metric $d_{\mathcal{V},e}$ on $\mathcal{V} \cap \bar{e}$. Moreover, by construction, with respect to the ordering on $\mathcal{V} \cap \bar{e}$ induced by the orientation on e, if $x < y < z$ are three points in $\mathcal{V} \cap \bar{e}$, then

$$d_{\mathcal{V},e}(x, z) = d_{\mathcal{V},e}(x, y) + d_{\mathcal{V},e}(y, z).$$

Because the L_k–lengths of edges in $E\Gamma_k$ tend to zero and because $\mathcal{V} \cap e$ is dense in $\bar{e}$, we can extend $d_{\mathcal{V},e}$ to a geodesic metric d_e on $\bar{e}$ (inducing the given topology): for an arbitrary pair of points x', y' in $\bar{e}$, we can approximate $d_e(x', y')$ from above (respectively, below) by $d_{\mathcal{V},e}$–distances between points in $\mathcal{V}$ limiting to x', y' from outside (respectively, inside) the arc between x' and y'. For $x, y \in \mathcal{V} \cap \bar{e}$ we will have $d_e(x, y) = d_{\mathcal{V},e}(x, y)$.

The collection of metrics d_e on edges of Γ determines a geodesic metric d on Γ consistent with the original topology on Γ. By construction of d_e and using local injectivity of f on edges, it follows that for any edge $e \in E\Gamma$ and any sufficiently close points $x, y \in \mathcal{V}$ we have $d(f(x), f(y)) = \lambda d(x, y)$. Since $\mathcal{V}$ is dense in e, it follows that f is locally a λ–homothety on edges with respect to d, as required. This completes the proof of the existence of a metric structure $\mathcal{L}$ satisfying conditions (1) and (2) of the theorem.

To see that such $\mathcal{L}$ is unique note that condition (2) of the theorem uniquely determines the $d_{\mathcal{L}}$–lengths of the edges of Γ. Proceeding inductively on k, we see that conditions (1) and (2) then uniquely determines the $d_{\mathcal{L}}$–lengths of all edges of Γ_k for all $k \geq 0$. For $k = 0$ we have $\Gamma_0 = \Gamma$ and the statement holds. If $k \geq 1$ and $e \in E\Gamma_k$ then $e' = f(e)$ is an edge of Γ_{k-1}, so that, by the inductive hypothesis, the $d_{\mathcal{L}}$ length of e' is already determined. Then by condition (1) it follows that the $d_{\mathcal{L}}$–length of e is equal to $1/\lambda$ times the $d_{\mathcal{L}}$ length of e', which completes the inductive step.

It thus follows that conditions (1) and (2) of the theorem uniquely define the $d_{\mathcal{L}}$ length along an edge of Γ between any two points of $\mathcal{V}$ belonging to that edge. Since the set $\mathcal{V}$ is dense in Γ, the uniqueness of $\mathcal{L}$ follows. $\qquad\square$

Remark A.6 This proof follows closely the strategy for constructing the stable foliation for a pseudo-Anosov in [26, Exposé 9.5].

Lemma A.2 and Theorem A.3 immediately imply:

Corollary A.7 Let Γ be a finite (linear) graph and let $f\colon \Gamma \to \Gamma$ be a regular graph-map such that f is expanding and $A(f)$ is irreducible.

Then there exists a (possibly different) linear structure Λ and compatible volume-1 metric structure $\mathcal{L} \subset \Lambda$ on the underlying topological graph Γ such that, with respect to $\mathcal{L}$, the map $f\colon \Gamma \to \Gamma$ is still linear and the restriction of f to any edge is affine with constant derivative $\lambda(f)$. (Thus locally on every edge, f is a $\lambda(f)$–homothety with respect to $d_{\mathcal{L}}$).

We call the objects Λ, $\mathcal{L}$ and $d_{\mathcal{L}}$ provided by Corollary A.7 (or Theorem A.3) the *canonical linear structure*, *canonical metric structure* and *canonical eigenmetric* associated to the given map $f\colon \Gamma \to \Gamma$, respectively. These canonical structures lead to the following rigidity statement. First we say that two regular topological graph maps $f_1, f_2\colon \Gamma \to \Gamma$ are *combinatorially equivalent* if for every $e \in E\Gamma$, we have $f_1(e) = f_2(e)$ as combinatorial edge-paths.

Theorem A.8 Suppose Λ_1, Λ_2 are linear structures on a graph Γ and suppose that for each $i = 1, 2$, $f_i\colon \Gamma \to \Gamma$ is an expanding Λ_i–linear graph map. If f_1 and f_2 are combinatorially equivalent and $A(f_1) = A(f_2)$ is irreducible, then f_1 and f_2 are conjugate by a homeomorphism $h\colon \Gamma \to \Gamma$ homotopic to the identity:

$$h f_1 h^{-1} = f_2.$$

Proof sketch Let d_i be the canonical eigenmetric associated to f_i, for $i = 1, 2$. If we write d_i^e for the induced path metric on the interior of $e \in E\Gamma$, then the homeomorphism $h\colon \Gamma \to \Gamma$ is the unique map that fixes $V\Gamma$ and restricts to an isometry $(e, d_1^e) \to (e, d_2^e)$ for each edge $e \in E\Gamma$. $\square$

Corollary A.9 Let $f\colon \Gamma \to \Gamma$ be a regular topological graph-map that is expanding on all scales. Then there exits a metric structure $\mathcal{L}$ on Γ with respect to which f is a combinatorial graph-map.

Proof First equip Γ with the linear structure Λ_1 provided by Theorem A.3, so that $f_1 = f\colon \Gamma \to \Gamma$ is an expanding irreducible (linear) graph-map. Next choose any metric structure $\mathcal{L}_2$ (with associated linear structure Λ_2) for which each edge has length 1. As indicated in Remark 2.5, we may then build a combinatorially equivalent (linear) graph-map $f_2\colon \Gamma \to \Gamma$ that is homotopic to f rel $V\Gamma$ and such that the subdivision of each edge along $f_2^{-1}(V\Gamma)$ is into equal length parts with respect to $\mathcal{L}_2$. Hence $f_2\colon \Gamma \to \Gamma$ is a combinatorial graph map with respect to $\mathcal{L}_2$.

Applying Theorem A.8 to the maps $f_1, f_2\colon \Gamma \to \Gamma$ with associated linear structures Λ_1, Λ_2, we obtain a homeomorphism $h\colon \Gamma \to \Gamma$ such that $h f_1 = f_2 h$. Pulling back

$\mathcal{L}_2$ via h now gives a metric structure $\mathcal{L}$ on Γ with respect to which $f = f_1$ is a combinatorial graph map.   □

Appendix B: Clean train-tracks and full irreducibility

In this appendix we obtain a strengthened and easier-to-use version of a criterion from [39] for checking the full irreducibility of hyperbolic automorphisms of $\mathrm{Out}(F_N)$.

Recall that according to our Definition 2.11, for a train-track map $f\colon \Gamma \to \Gamma$ we allow the graph Γ to have valence-2 vertices but no valence-1 vertices. However we do require that for any vertex v of valence 2 and any $n \geq 1$ the map f^n be locally injective on some small neighborhood of v in Γ. This condition is automatically satisfied if we only require f^n to be an immersion on all edges for all n and also require the Whitehead graphs of all valence-2 vertices to be connected. Recall from Definition 2.3 that a train-track map $f\colon \Gamma \to \Gamma$ is *expanding* if for every edge e of Γ the simplicial length $|f^n(e)|$ tends to ∞ as $n \to \infty$.

Recall also from Definition 2.18 that an expanding train-track map $f\colon \Gamma \to \Gamma$ whose Whitehead graphs $\mathrm{Wh}(f, v)$ are connected for all $v \in V\Gamma$ is *clean* if there exits $m > 0$ such that $A(f^m) > 0$ and is *weakly clean* if $A(f)$ is irreducible. By definition, a clean train-track is weakly clean; here we show that the converse is also true.

Remark B.1   Note that if Δ is a finite connected subgraph of a graph Γ then the inclusion $\iota\colon \Delta \to \Gamma$ induces an injective homomorphism $\iota_*\colon H_1(\Delta; \mathbb{R}) \to H_1(\Gamma; \mathbb{R})$. For this reason if, in this situation, $f\colon \Gamma \to \Gamma$ is a graph-map which is homotopy equivalence, then $b_1(\Delta) \leq b_1(f(\Delta))$.

Proposition B.2   *Let $f\colon \Gamma \to \Gamma$ be a weakly clean train-track. Then f is clean.*

Proof   Since f is expanding and Γ is finite, for every oriented edge e of Γ there exists $n > 0$ and an oriented edge e' such that the path $f^n(e)$ contains at least three occurrences of the edge e' (with positive orientation). From irreducibility of $A(f)$ it follows that there exists $j > 0$ such that $f^j(e')$ contains either e or e^{-1}. Then $f^{n+j}(e)$ contains at least 3 occurrences of e or at least 3 occurrences of e^{-1}. Hence $f^{2(n+j)}(e)$ contains at least 3 occurrences of e. Put $k(e) := 2(n + j)$ and then put m to be the product of $k(e)$ taken over all oriented edges e of Γ. Then for any oriented edge e the path $f^m(e)$ contains at least three positive occurrences of e and the same is true for $f^{mq}(e)$ for any integer $q \geq 1$.

Now let e be an oriented edge of Γ and for $n \geq 1$ put $\Delta_0 = \{e\}$ and let Δ_n be the subgraph of Γ given by the union of all the edges in the path $f^{mn}(e)$. Then, by

construction, $\Delta_n \subseteq \Delta_{n+1}$ for all $n \geq 0$ and we put $\Delta(e) := \bigcup_{n \geq 0} \Delta_n$. Since Γ is finite, and the graphs Δ_n are connected, $\Delta(e)$ is a finite connected subgraph of Γ containing e and such that $\Delta(e) = \Delta(e^{-1})$. Also, by construction, we have $f^m(\Delta(e)) = \Delta(e)$.

Claim 1 *For any edge e of Γ and any $j > 0$ the graphs $\Delta(e)$ and $f^j(\Delta(e))$ are homotopy equivalent.*

Proof Since $f \colon \Gamma \to \Gamma$ is a homotopy equivalence, [Remark B.1](#) implies that

$$b_1(\Delta(e)) \leq b_1(\Delta'),$$

where $\Delta' = f^j(\Delta(e))$. We can find $k \geq 1$ such that $j < mk$. Then $f^{mk-j}(\Delta') = f^{mk}(\Delta(e)) = \Delta(e)$ and hence, again by [Remark B.1](#), we have $b_1(\Delta') \leq b_1(\Delta(e))$. Hence $b_1(\Delta') = b_1(\Delta(e))$, as required. $\qquad\qquad\square$

Claim 2 *For any two edges e, e' of Γ we have $b_1(\Delta(e)) = b_1(\Delta(e'))$ (and hence the graphs $\Delta(e)$ and $\Delta(e')$ are homotopy equivalent).*

Proof Indeed, let e, e' be arbitrary oriented edges. By irreducibility of $A(f)$ there exists $j > 0$ such that $f^j(e)$ contains e' or $(e')^{-1}$. Then for any $k \geq 0$,

$$f^{mk}(f^j \Delta(e)) = f^{j+mk}(\Delta(e)) = f^j(f^{mk}(\Delta(e))) = f^j(\Delta(e)).$$

Hence $f^{mk}(e')$ is contained in $f^j(\Delta(e))$ for every $k \geq 1$ and therefore, by definition, $\Delta(e') \subseteq f^j(\Delta(e))$. Hence $b_1(\Delta(e')) \leq b_1(f^j(\Delta(e))) = b_1(\Delta(e))$. A symmetric argument shows that $b_1(\Delta(e)) \leq b_1(\Delta(e'))$ and hence $b_1(\Delta(e)) = b_1(\Delta(e'))$. $\qquad\square$

Claim 3 *For any edge e of Γ we have $\Delta(e) = \Gamma$.*

Proof Indeed, suppose not, and let e be such that $\Delta(e) \neq \Gamma$. Then there exists a vertex v of $\Delta(e)$ such that v is incident to an edge of $\Delta(e)$ and to an edge of $\Gamma - \Delta(e)$.

Since the Whitehead graph $\mathrm{Wh}_\Gamma(f, v)$ is connected, there exist oriented edges e_1 and e_2 with origin v such that $e_1 \subseteq \Delta(e)$, $e_2 \nsubseteq \Delta(e)$ and such that the turn e_1, e_2 is taken by f. Hence there exist an edge e_3 and $i > 0$ such that $e_1^{-1} e_2$ is a subpath of $f^i(e_3)$. Since $f \colon \Gamma \to \Gamma$ is surjective, there exist $k \geq 1$ and an edge e_4 such that $km > i$ and such that $f^{km-i}(e_4)$ contains e_3. Hence $f^m(e_4)$ contains $e_1^{-1} e_2$ as a subpath. Therefore, by definition, we have $e_1 \subseteq \Delta(e_4)$ and $e_2 \subseteq \Delta(e_4)$ and hence $\Delta(e_1) \subseteq \Delta(e_4)$ and $\Delta(e_2) \subseteq \Delta(e_4)$.

Also, since $f^m(\Delta(e)) = \Delta(e)$ and since e_1 is an edge of $\Delta(e)$, we have $\Delta(e_1) \subseteq \Delta(e)$. In particular, e_2 is not contained in $\Delta(e_1)$.

Let v' be the terminal vertex of e_2. Then $v' \notin \Delta(e_1)$ since otherwise $\Xi := e_2 \cup \Delta(e_1)$ has the first Betti number bigger than that of $\Delta(e_1)$ which is impossible since $\Xi \subseteq \Delta(e_4)$ and $b_1(\Delta(e_1)) = b_1(\Delta(e_4))$ by Claim 2.

By the choice of m the path $f^m(e_2)$ contains at least 3 positive occurrences of e_2. Hence there exists a fixed point x of f^m in the interior of e_2 coming from a positive occurrence of e_2 in $f^m(e_2)$ which is neither initial nor terminal edge of $f^m(e_2)$. Thus f^m preserves the orientation in a small neighborhood of x. Let $[x, v']$ be the segment of e_2 from x to v'. Then $f^m([x, v']) = [x, v'] \cup \beta$, where β is a tight combinatorial edge-path of positive simplicial length with initial vertex v'. Thus for all $n \geq 0$ the path $f^{mn}([x, v']) = [x, v'] \cup \beta_n$ is a proper initial segment of the path $f^{m(n+1)}([x, v']) = [x, v'] \cup \beta_{n+1}$, where β_n is a tight combinatorial edge-path starting at v' with $|\beta_n| \to \infty$ as $n \to \infty$. Let ρ be the path in Γ obtained by taking the "union" of all $f^{mn}([x, v'])$, that is, ρ is the path starting at x such that for all $n \geq 0$ $f^{mn}([x, v'])$ is an initial segment of ρ. This ρ is sometimes called a "combinatorial eigenray" of f. By construction we have $\rho \subseteq \Delta(e_2)$. The path ρ has the form $\rho = [x, v']\rho'$ where ρ' is a semiinfinite edge-path in Γ starting at v'.

Then either ρ' has a nonempty intersection with $\Delta(e_1)$ (in which case we look at the first time when ρ' hits the graph $\Delta(e_1)$), or, if ρ' and $\Delta(e_1)$ are disjoint, then there is an oriented edge that occurs in ρ' at least twice and hence ρ' contains a nontrivial immersed loop. In both cases we see that $\Xi_1 = \Delta(e_1) \cup e_2 \cup \rho'$ has $b_1(\Xi_1) > b_1(\Delta(e_1))$. However, $\Xi_1 \subseteq \Delta(e_4)$ and hence $b_1(\Delta(e_4)) > b_1(\Delta(e_1))$, which contradicts Claim 2. $\square$

Thus we know that $\Delta(e) = \Gamma$ for every edge e of Γ. By definition of $\Delta(e)$ it follows that for every edge e of Γ there exists $n_e \geq 1$ such that $f^{n_e m}(e)$ passes through every topological edge of Γ. Since $f : \Gamma \to \Gamma$ is surjective, it follows that for every $i \geq n_e m$ the path $f^i(e)$ also passes through every topological edge of Γ. Now put M to be the maximum of $n_e m$ taken over all edges e of Γ. Then for any two oriented edges e, e' of Γ the path $f^M(e)$ contains an occurrence of e' or $(e')^{-1}$. This means that $A(f^M) > 0$, as required. $\square$

Note that if $\varphi \in \mathrm{Out}(F_N)$ is hyperbolic (ie atoroidal) and $f : \Gamma \to \Gamma$ is a train-track representative of φ with $A(f)$ irreducible, then f is expanding (since otherwise $A(f)$ is a permutation matrix and φ has finite order, which contradicts the hyperbolicity assumption on φ). Moreover, it is proved in [39] that for $\varphi \in \mathrm{Out}(F_N)$ hyperbolic, the element φ is fully irreducible if and only if some (equivalently, any) train-track representative of φ is clean. Thus combining [39, Proposition 4.4] with the equivalence between the clean and weakly clean conditions (Proposition B.2), we immediately obtain:

Theorem B.3 *Let $N \geq 3$ and let $\varphi \in \mathrm{Out}(F_N)$ be a hyperbolic element. Then the following conditions are equivalent:*

(1) *The automorphism φ is fully irreducible.*

(2) *If $f \colon \Gamma \to \Gamma$ is a train-track representative of φ then f is clean.*

(3) *There exists a clean train-track representative of φ.*

(4) *There exists a weakly clean train-track representative of φ.*

Another useful corollary of Proposition B.2 is that for hyperbolic elements of $\mathrm{Out}(F_N)$ the notions of being irreducible and fully irreducible coincide:

Corollary B.4 *Let $N \geq 3$ and let $\varphi \in \mathrm{Out}(F_N)$ be a hyperbolic element. Then φ is irreducible if and only if φ is fully irreducible.*

Proof Obviously, if φ is fully irreducible then φ is irreducible. Thus let $\varphi \in \mathrm{Out}(F_N)$ be hyperbolic and irreducible. Assume that φ is not fully irreducible. Since φ is irreducible, by a result of Bestvina and Handel [9] there exists a train-track representative $f \colon \Gamma \to \Gamma$ of φ with an irreducible transition matrix $A(f)$. Since φ is hyperbolic and thus has infinite order in $\mathrm{Out}(F_N)$, we know that $A(f)$ is not a permutation matrix. Hence $\lambda(f) > 1$, so that f is expanding. By assumption f is not fully irreducible and therefore, by Theorem B.3 above, the train-track f is not weakly clean. This means that for some vertex v of Γ the Whitehead graph $\mathrm{Wh}(f, v)$ is not connected. Proposition 4.1 of [39] (which first appears, in implicit form, in the proof of Proposition 4.5 in [9]) now implies that φ is reducible, contrary to our assumption that φ is irreducible. Thus φ is fully irreducible, as required. □

Note that the conclusion of Corollary B.4 fails without the assumption that φ is hyperbolic, as already observed by Bestvina and Handel in [9]. For example, let S be a compact connected hyperbolic surface with $m \geq 3$ boundary components, let $F_N = \pi_1(S)$ and let $\varphi \in \mathrm{Out}(F_N)$ be induced by a pseudo-Anosov homeomorphism g of S. The boundary components of S represent primitive elements of F_N and therefore such φ is not fully irreducible. However, if g transitively permutes the boundary components of S, then φ is irreducible.

References

[1] **L M Abramov**, *The entropy of a derived automorphism*, Dokl. Akad. Nauk SSSR 128 (1959) 647–650   MR0113984

[2]   **I Agol**, *Ideal triangulations of pseudo-Anosov mapping tori*, from: "Topology and geometry in dimension three", (W Li, L Bartolini, J Johnson, F Luo, R Myers, J H Rubinstein, editors), Contemp. Math. 560, Amer. Math. Soc. (2011) 1–17   MR2866919

[3]   **Y Algom-Kfir**, **E Hironaka**, **K Rafi**, *Digraphs and cycle polynomials for free-by-cyclic groups*, Geom. Topol. 19 (2015) 1111–1154   MR3336279

[4]   **Y Algom-Kfir**, **K Rafi**, *Mapping tori of small dilatation expanding train-track maps*, Topology Appl. 180 (2015) 44–63   MR3293265

[5]   **L Alsedà**, **F Mañosas**, **P Mumbrú**, *Minimizing topological entropy for continuous maps on graphs*, Ergodic Theory Dynam. Systems 20 (2000) 1559–1576   MR1804944

[6]   **M Bestvina**, **N Brady**, *Morse theory and finiteness properties of groups*, Invent. Math. 129 (1997) 445–470   MR1465330

[7]   **M Bestvina**, **M Feighn**, *A combination theorem for negatively curved groups*, J. Differential Geom. 35 (1992) 85–101   MR1152226

[8]   **M Bestvina**, **M Feighn**, *Hyperbolicity of the complex of free factors*, Adv. Math. 256 (2014) 104–155   MR3177291

[9]   **M Bestvina**, **M Handel**, *Train tracks and automorphisms of free groups*, Ann. of Math. 135 (1992) 1–51   MR1147956

[10] **R Bieri**, **W D Neumann**, **R Strebel**, *A geometric invariant of discrete groups*, Invent. Math. 90 (1987) 451–477   MR914846

[11] **R Bott**, **L W Tu**, *Differential forms in algebraic topology*, Graduate Texts in Mathematics 82, Springer, New York (1982)   MR658304

[12] **R Bowen**, *Topological entropy and axiom* A, from: "Global Analysis", Proc. Symp. Pure Math. 14, Amer. Math. Soc. (1970) 23–41   MR0262459

[13] **P Brinkmann**, *Hyperbolic automorphisms of free groups*, Geom. Funct. Anal. 10 (2000) 1071–1089   MR1800064

[14] **P Brinkmann**, **S Schleimer**, *Computing triangulations of mapping tori of surface homeomorphisms*, Experiment. Math. 10 (2001) 571–581   MR1881758

[15] **J O Button**, *Mapping tori with first Betti number at least two*, J. Math. Soc. Japan 59 (2007) 351–370   MR2325689

[16] **D Calegari**, **J Maher**, *Statistics and compression of scl*, Ergodic Theory Dynam. Systems 35 (2015) 64–110   MR3294292

[17] **M Clay**, **A Pettet**, *Twisting out fully irreducible automorphisms*, Geom. Funct. Anal. 20 (2010) 657–689   MR2720227

[18] **M Denker**, **C Grillenberger**, **K Sigmund**, *Ergodic theory on compact spaces*, Lecture Notes in Mathematics 527, Springer, Berlin (1976)   MR0457675

[19] **S Dowdall**, **I Kapovich**, **C J Leininger**, *McMullen polynomials and Lipschitz flows for free-by-cyclic groups*   arXiv:1310.7481

[20]   **S Dowdall**, **I Kapovich**, **C J Leininger**, *Dynamics on free-by-cyclic groups* (2013) arXiv:1301.7739v5

[21]   **S Dowdall**, **I Kapovich**, **C J Leininger**, *Unbounded asymmetry of stretch factors*, C. R. Math. Acad. Sci. Paris 352 (2014) 885–887   MR3268757

[22]   **T Downarowicz**, *Entropy in dynamical systems*, New Mathematical Monographs 18, Cambridge Univ. Press (2011)   MR2809170

[23]   **N M Dunfield**, *Alexander and Thurston norms of fibered* 3*–manifolds*, Pacific J. Math. 200 (2001) 43–58   MR1863406

[24]   **B Farb**, **C J Leininger**, **D Margalit**, *Small dilatation pseudo-Anosov homeomorphisms and* 3*–manifolds*, Adv. Math. 228 (2011) 1466–1502   MR2824561

[25]   **M Farber**, **R Geĭgan**, **D Shyutts**, *Closed* 1*–forms in topology and geometric group theory*, Uspekhi Mat. Nauk 65 (2010) 145–176   MR2655245   In Russian; translated in Russian Math. Surveys 65 (2010) 143–172

[26]   **A Fathi**, **F Laudenbach**, **V Poénaru** (editors), *Travaux de Thurston sur les surfaces*, Astérisque 66–67, Soc. Math. France, Paris (1979)   MR568308

[27]   **M Feighn**, **M Handel**, *Mapping tori of free group automorphisms are coherent*, Ann. of Math. 149 (1999) 1061–1077   MR1709311

[28]   **S Francaviglia**, **A Martino**, *Metric properties of outer space*, Publ. Mat. 55 (2011) 433–473   MR2839451

[29]   **D Fried**, *Flow equivalence, hyperbolic systems and a new zeta function for flows*, Comment. Math. Helv. 57 (1982) 237–259   MR684116

[30]   **D Fried**, *The geometry of cross sections to flows*, Topology 21 (1982) 353–371 MR670741

[31]   **D Gabai**, *Foliations and the topology of* 3*–manifolds*, J. Differential Geom. 18 (1983) 445–503   MR723813

[32]   **F Gautero**, *Dynamical* 2*–complexes*, Geometriae Dedicata 88 (2001) 283–319 MR1877221

[33]   **F Gautero**, *Feuilletages de* 2*–complexes*, Ann. Fac. Sci. Toulouse Math. 10 (2001) 619–638   MR1944253

[34]   **F Gautero**, *Cross sections to semi-flows on* 2*–complexes*, Ergodic Theory Dynam. Systems 23 (2003) 143–174   MR1971200

[35]   **F Gautero**, *Combinatorial mapping-torus, branched surfaces and free group automorphisms*, Ann. Sc. Norm. Super. Pisa Cl. Sci. 6 (2007) 405–440   MR2370267

[36]   **R Geoghegan**, **M L Mihalik**, **M Sapir**, **D T Wise**, *Ascending HNN extensions of finitely generated free groups are Hopfian*, Bull. London Math. Soc. 33 (2001) 292–298 MR1817768

[37]   **M Handel**, **L Mosher**, *The expansion factors of an outer automorphism and its inverse*, Trans. Amer. Math. Soc. 359 (2007) 3185–3208   MR2299451

[38]   **A Hatcher**, *Algebraic topology*, Cambridge Univ. Press (2002)   MR1867354

[39]   **I Kapovich**, *Algorithmic detectability of iwip automorphisms*, Bull. Lond. Math. Soc. 46 (2014) 279–290   MR3194747

[40]   **I Kapovich**, **M Lustig**, *Cannon–Thurston fibers for iwip automorphisms of F_N* , J. Lond. Math. Soc. 91 (2015) 203–224   MR3335244

[41]   **I Kapovich**, **A Myasnikov**, *Stallings foldings and subgroups of free groups*, J. Algebra 248 (2002) 608–668   MR1882114

[42]   **I Kapovich**, **K Rafi**, *On hyperbolicity of free splitting and free factor complexes*, Groups Geom. Dyn. 8 (2014) 391–414   MR3231221

[43]   **G Levitt**, *1–formes fermées singulières et groupe fondamental*, Invent. Math. 88 (1987) 635–667   MR884804

[44]   **D D Long**, **U Oertel**, *Hyperbolic surface bundles over the circle*, from: "Progress in knot theory and related topics", (M Boileau, M Domergue, Y Mathieu, K Millett, editors), Travaux en Cours 56, Hermann, Paris (1997) 121–142   MR1603150

[45]   **C T McMullen**, *Polynomial invariants for fibered 3–manifolds and Teichmüller geodesics for foliations*, Ann. Sci. École Norm. Sup. 33 (2000) 519–560   MR1832823

[46]   **C T McMullen**, *The Alexander polynomial of a 3–manifold and the Thurston norm on cohomology*, Ann. Sci. École Norm. Sup. 35 (2002) 153–171   MR1914929

[47]   **L Mosher**, *Dynamical systems and the homology norm of a 3–manifold, I: Efficient intersection of surfaces and flows*, Duke Math. J. 65 (1992) 449–500   MR1154179

[48]   **L Mosher**, *Dynamical systems and the homology norm of a 3–manifold, II*, Invent. Math. 107 (1992) 243–281   MR1144424

[49]   **W D Neumann**, *Normal subgroups with infinite cyclic quotient*, Math. Sci. 4 (1979) 143–148   MR562003

[50]   **U Oertel**, *Homology branched surfaces: Thurston's norm on $H_2(M^3)$*, from: "Low-dimensional topology and Kleinian groups", (D B A Epstein, editor), London Math. Soc. Lecture Note Ser. 112, Cambridge Univ. Press (1986) 253–272   MR903869

[51]   **U Oertel**, *Affine laminations and their stretch factors*, Pacific J. Math. 182 (1998) 303–328   MR1609595

[52]   **J-P Otal**, *The hyperbolization theorem for fibered 3–manifolds*, SMF/AMS Texts and Monographs 7, Amer. Math. Soc. (2001)   MR1855976

[53]   **C Pfaff**, *Constructing and classifying fully irreducible outer automorphisms of free groups*   arXiv:1205.5320

[54]   **C Pfaff**, *Ideal Whitehead graphs in $Out(F_r)$, II: The complete graph in each rank*, J. Homotopy Relat. Struct. 10 (2015) 275–301   MR3348263

[55]   **I Rivin**, *Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms*, Duke Math. J. 142 (2008) 353–379   MR2401624

[56]   **I Rivin**, *Zariski density and genericity*, Int. Math. Res. Not. 2010 (2010) 3649–3657   MR2725508

[57]   **M Scharlemann**, *Sutured manifolds and generalized Thurston norms*, J. Differential Geom. 29 (1989) 557–614   MR992331

[58]   **J Stallings**, *On fibering certain* 3–*manifolds*, from: "Topology of 3–manifolds and related topics", Prentice-Hall, Englewood Cliffs, NJ (1962) 95–100   MR0158375

[59]   **J R Stallings**, *Topology of finite graphs*, Invent. Math. 71 (1983) 551–565   MR695906

[60]   **W P Thurston**, *A norm for the homology of* 3–*manifolds*, Mem. Amer. Math. Soc. 339, Amer. Math. Soc. (1986)   MR823443

[61]   **W P Thurston**, *On the geometry and dynamics of diffeomorphisms of surfaces*, Bull. Amer. Math. Soc. 19 (1988) 417–431   MR956596

[62]   **D Tischler**, *On fibering certain foliated manifolds over S^1*, Topology 9 (1970) 153–154   MR0256413

[63]   **Z Wang**, *Mapping tori of outer automorphisms of free groups*, PhD thesis, Rutgers, NJ (2002)   MR2703150   Available at http://search.proquest.com/docview/305500756

SD: *Department of Mathematics, Vanderbilt University*
1326 Stevenson Center, Nashville, TN 37240, USA

IK,CJL: *Department of Mathematics, University of Illinois at Urbana-Champaign*
1409 West Green Street, Urbana, IL 61801, USA

spencer.dowdall@vanderbilt.edu,   kapovich@math.uiuc.edu,
clein@math.uiuc.edu

http://www.math.vanderbilt.edu/~dowdalsd/,
http://www.math.uiuc.edu/~kapovich/,   http://www.math.uiuc.edu/~clein/