ISSN 1364-0380 (on line) 1465-3060 (printed)

Geometry & Topology
Volume 8 (2004) 1127–1188
Published: 8 September 2004

Unimodal generalized pseudo-Anosov maps

André de Carvalho
Toby Hall

Departamento de Matemática Aplicada, IME - USP
Rua do Matão 1010, Cidade Universitária
05508-090 São Paulo, SP, Brazil
and
Department of Mathematical Sciences, University of Liverpool
Liverpool L69 7ZL, UK

Email: `andre@ime.usp.br` and `T.Hall@liv.ac.uk`

Abstract

An infinite family of generalized pseudo-Anosov homeomorphisms of the sphere S is constructed, and their invariant foliations and singular orbits are described explicitly by means of generalized train tracks. The complex strucure induced by the invariant foliations is described, and is shown to make S into a complex sphere. The generalized pseudo-Anosovs thus become quasiconformal automorphisms of the Riemann sphere, providing a complexification of the unimodal family which differs from that of the Fatou/Julia theory.

AMS Classification numbers Primary: 37E30

Secondary: 57M50

Keywords: Pseudo-Anosov homeomorphisms, train tracks, unimodal maps, horseshoe

Proposed: Joan Birman — Received: 10 July 2003
Seconded: David Gabai, Yasha Eliashberg — Revised: 4 February 2004

1 Introduction

1.1 Overview

Pseudo-Anosov homeomorphisms were introduced by Thurston in his classification of surface homeomorphisms up to isotopy. A surface homeomorphism $\Phi\colon S \circlearrowleft$ is pseudo-Anosov if it preserves a transverse pair of measured foliations with finitely many singularities, expanding one foliation uniformly by a factor $\lambda > 1$ and contracting the other by a factor $1/\lambda$. Pseudo-Anosov maps can be described combinatorially by train tracks — a class of graphs embedded in the surface with additional information on the vertices which specifies the turns a train riding along the track can make at each vertex — and their endomorphisms. Using information obtained from an associated transition matrix, the surface can be reconstructed by identifying sides of Euclidean rectangles foliated by horizontal and vertical line segments. The pseudo-Anosov expands the rectangles horizontally and contracts them vertically, mapping them as dictated by the train track map. Up to this point, the discussion is finite: train tracks are graphs with finitely many edges (Thurston's theorem can be proved algorithmically, the main step being to find a finite train track invariant under a given isotopy class of homeomorphisms) and train track maps are finite-to-one. It is possible, however, to forgo some of the finiteness requirements — the graphs remain finite, the endomorphisms remain finite-to-one, but the additional information at the vertices is allowed to be infinite — but otherwise to go through the construction of the maps as before. This leads to the construction of *generalized pseudo-Anosov homeomorphisms.* These are defined similarly to pseudo-Anosov homeomorphisms, except that the invariant foliations are permitted to have infinitely many singularities, provided that they accumulate on only finitely many points. The purpose of this paper is to give a detailed construction and description of an infinite family of generalized pseudo-Anosovs of the sphere for which the underlying graph and graph map are the simplest possible: an interval and a unimodal endomorphism (ie, a continuous piecewise monotone map of the interval with exactly two monotone pieces).

In [12], a complete description of the family of pseudo-Anosov maps with underlying unimodal interval endomorphisms was given. It was shown that there is a countable family of such maps parameterized by a rational number between 0 and $1/2$, called *height.* Height turns out to be a braid type invariant and this leads to the proof of weak universality results for families of plane homeomorphisms passing from trivial to chaotic dynamics as parameters are varied. Height also plays a central role in this paper and, in turn, the results presented here provide a geometric interpretation of it. The family of unimodal generalized pseudo-Anosovs extends that of unimodal pseudo-Anosovs. The height specifies the behaviour of the maps at infinity: given a rational m/n, there is an interval of kneading sequences of height m/n, whose associated generalized pseudo-Anosovs have the same behaviour at infinity. The generalized pseudo-Anosov is a pseudo-Anosov for exactly one kneading sequence in this interval.

This paper provides an explicit description of the generalized train track associated to any periodic or preperiodic kneading sequence, depending crucially on its height. The process of constructing a generalized pseudo-Anosov from a generalized train track map is similar to that of constructing a pseudo-Anosov from a train track map, but requires more care because of the more intricate nature of the identifications carried out on the sphere. The topological tool used to guarantee that the identification space is again a sphere is Moore's theorem about monotone upper semi-continuous decompositions of the sphere.

The invariant foliations of a generalized pseudo-Anosov define a complex structure on the sphere away from the accumulations of singularities. It is shown that for the unimodal generalized pseudo-Anosovs considered here, these accumulations are removable singularities of the complex structure, so that the sphere is a complex sphere, with the foliations being the horizontal and vertical trajectories of an integrable quadratic differential having infinitely many zeros and poles. The construction therefore provides a complexification of unimodal maps as quasiconformal automorphisms of the Riemann sphere, in contrast to the complexification arising via the theory of Fatou/Julia, where one thinks of the unimodal map as the real slice of an endomorphism of the Riemann sphere.

By a suitable normalization, the sphere of definition of the generalized pseudo-Anosovs can be identified canonically with the Riemann sphere, and hence the family of unimodal generalized pseudo-Anosovs, which initially are constructed on abstract topological spheres, can be regarded as a family of Teichmüller mappings of the Riemann sphere, making it possible to consider taking limits within the family. This is a necessary step in the problem of constructing a completion of the set of all pseudo-Anosov homeomorphisms of the sphere.

Section 2 describes the class of *thick interval maps*. These are homeomorphisms of the sphere which provide the starting point for the definition and construction of invariant generalized train tracks, a process which is described in Section 3. Section 4 provides a summary of necessary results on unimodal maps and Smale's horseshoe, and defines the subclass of *unimodal* thick interval maps which are used in the remainder of the paper. In Section 5, the *outside* *dynamics* of a unimodal map is defined and analysed: intuitively, this provides a description of those orbits of a unimodal map which are never lost 'inside' the fold. The main contents of the paper can be found in Sections 6, 7 and 8. The invariant generalized train track for a given unimodal thick interval map is described explicitly, a detailed account is given of how the generalized train track map can be used to construct the corresponding generalized pseudo-Anosov, and the complex structure induced by the invariant foliations is analysed.

Acknowledgements The authors are grateful for the referee's careful reading of the paper and helpful comments. This material is based upon work supported by the National Science Foundation under Grant No. 0203975. The first author is supported by FAPESP Grant No. 02/05072-5.

1.2 Definitions and notation

Let X be a metric space. An *isotopy* on X is a continuous map $\psi\colon X \times [0,1] \to X$ with the property that the slice maps $\psi_t\colon X \to X$ are homeomorphisms for $0 \le t \le 1$. A *pseudo-isotopy* is defined similarly but it is only required that the slice maps be homeomorphisms for $0 \le t < 1$. In particular, this means that the map ψ_1 is a *near-homeomorphism*, ie, it can be approximated arbitrarily closely by homeomorphisms. An isotopy or pseudo-isotopy $\psi\colon X \times [0,1] \to X$ is said to be *supported* on a subset U of X if all of its slice maps agree on $X \setminus U$.

Symbolic dynamics on both $\{0,1\}^{\mathbb{N}}$ and $\{0,1\}^{\mathbb{Z}}$ will be used in this paper. Elements of these spaces are regarded as semi-infinite or bi-infinite sequences of 0s and 1s, and in the case of $\{0,1\}^{\mathbb{Z}}$ a period is placed before the origin of the sequence (ie, the image of $0 \in \mathbb{Z}$): $\dots s_{-2}s_{-1} \cdot s_0 s_1 s_2 \dots \in \{0,1\}^{\mathbb{Z}}$. If $w \in \{0,1\}^n$ for some n, then the notation w^∞ is used to indicate semi-infinite repetition of w ($w^\infty = wwww\dots \in \{0,1\}^{\mathbb{N}}$), while $\overline{w}$ is used to indicate bi-infinite repetition ($\overline{w} = \dots www \cdot www \dots \in \{0,1\}^{\mathbb{Z}}$).

The notation S^1 and S^2 is reserved for the standard 1- and 2-dimensional spheres, and different symbols are used to denote general topological spheres.

A homeomorphism $\Phi\colon S \circlearrowleft$ of a smooth surface S is called a *generalized pseudo-Anosov map* if there exist

a) a finite Φ-invariant set Σ;

b) a pair $(\mathcal{F}^s, \mu^s)$, $(\mathcal{F}^u, \mu^u)$ of transverse measured foliations of $S \setminus \Sigma$ with countably many pronged singularities (with local charts as depicted in Figure 1), which accumulate on each point of Σ and have no other accumulation points. The transverse measures are required to be equivalent to Lebesgue measure on transversals;

c) a real number $\lambda > 1$;

such that

$$\begin{array}{rcl} \Phi(\mathcal{F}^s, \mu^s) & = & (\mathcal{F}^s, \frac{1}{\lambda}\mu^s) \\ \Phi(\mathcal{F}^u, \mu^u) & = & (\mathcal{F}^u, \lambda\mu^u). \end{array}$$

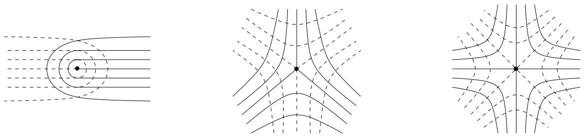


Figure 1: Pronged singularities of the invariant foliations

In particular, a generalized pseudo-Anosov is a pseudo-Anosov map if and only if there are only finitely many pronged singularities (ie, $\Sigma = \emptyset$). Note that the measures necessarily have full support and no atoms.

2 Markov thick interval maps

2.1 Thick interval maps

Thick graph maps are a class of surface homeomorphisms which have been described and used in several papers (for example [2, 5, 10, 4]). In this section a brief description of thick interval maps (where the surface is the sphere and the graph is an interval) is given. As the name suggests, a thick interval map is essentially an interval endomorphism which has been thickened up and made into a homeomorphism of the sphere, whose dynamics reflects that of the underlying interval map.

Throughout the paper, S^2 will be thought of as the one-point compactification of $\mathbb{R}^2$. The point at infinity will be denoted ∞ and the thick intervals defined below will always be assumed not to contain ∞.

Definitions 1 A *thick interval* $\mathbb{I} \subset S^2$ is a closed topological 2-disk partitioned into compact *decomposition elements*, such that

i) each decomposition element of $\mathbb{I}$ is either a *leaf* homeomorphic to $[0,1]$, or a *junction* homeomorphic to a closed 2-disk;

ii) the boundary in $\mathbb{I}$ of each junction consists of one or two disjoint arcs: if there is one (respectively two) such arc(s) the junction is called a 1-junction (respectively 2-junction);

iii) there are exactly two 1-junctions and finitely many 2-junctions;

iv) each decomposition element is contained in a chart as depicted in Figure 2.

If $\mathbb{I}$ is a thick interval, then the space obtained by collapsing each decomposition element to a point is an interval, whose vertices (the two endpoints, coming from the two 1-junctions, and finitely many valence 2 vertices) correspond to the junctions of $\mathbb{I}$. The union of the junctions of $\mathbb{I}$ is denoted $\mathbb{V}$, and the components of $\mathbb{I} \setminus \mathbb{V}$ are called *strips*: each strip is therefore homeomorphic to $(0,1) \times [0,1]$. The union of the closures of the strips is denoted $\mathbb{E}$. Thus $\mathbb{E} \cap \mathbb{V}$ is a union of closed arcs which are the boundary components (in $\mathbb{I}$) of both the junctions and the strips of $\mathbb{I}$.

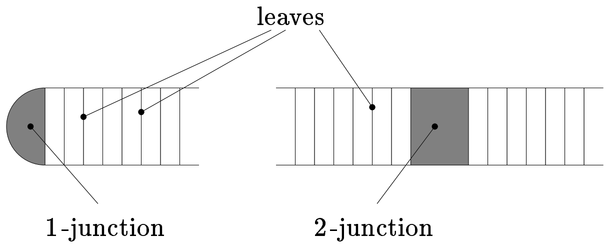


Figure 2: Charts in a thick interval

The following definition of a thick interval map is more restrictive than that used in other papers: conditions iv) to vi) have been added in order to make the definition and construction of generalized train tracks more straightforward. However, thick interval maps in the sense of [5] can be made to conform to the definition below by simple isotopies and changes in the thick interval structure which do not change the dynamics. In order to make the description more explicit, fixed orientation-preserving coordinate maps $h_s\colon \overline{s} \to [0,1] \times [0,1]$ are introduced on the closure of each strip s, with the property that the leaves of s are of the form $h_s^{-1}(\{x\} \times [0,1])$ for $0 < x < 1$.

Definition 2 A *thick interval map* is an orientation-preserving homeomorphism $F\colon (S^2, \mathbb{I}) \circlearrowleft$ such that:

i) $F(\mathbb{I}) \subset \mathrm{Int}(\mathbb{I})$.

ii) If γ is a leaf of $\mathbb{I}$, then $F(\gamma)$ is contained in a decomposition element, and $\mathrm{diam}(F^n(\gamma)) \to 0$ as $n \to \infty$. If J is a junction of $\mathbb{I}$, then $F(J)$ is contained in a junction.

iii) The point at infinity is a repelling fixed point whose basin contains $S^2 \setminus \mathbb{I}$.

iv) F is linear with respect to the coordinates h_s: in each connected component of $s_i \cap F^{-1}(s_j)$, where s_i, s_j are strips, F contracts vertical coordinates uniformly by a factor $\mu_j < 1$ and expands horizontal coordinates uniformly by a factor $\lambda_j \geq 1$;

v) If J, J' are junctions such that $F(J) \subset J'$, then $F(\partial_{\mathbb{I}} J) \subset \partial_{\mathbb{I}} J'$;

vi) If J is a junction with $F^n(J) \subseteq J$ for some (least) $n \geq 1$, then J has an attracting periodic point of least period n in its interior whose basin contains $\mathrm{Int}(J)$.

Remark 1 Item iii) in the definition says that the dynamics of a thick interval map in $S^2 \setminus \mathbb{I}$ is easily understood and uninteresting.

Let $F\colon (S^2, \mathbb{I}) \circlearrowleft$ be a thick interval map, $\tilde{I}$ be the interval obtained by collapsing each decomposition element of $\mathbb{I}$ to a point, and $\tilde{\pi}\colon \mathbb{I} \to \tilde{I}$ be the canonical projection. Then F induces a continuous map $\tilde{f}\colon \tilde{I} \circlearrowleft$: however $\tilde{f}$ is locally constant at preimages of $\tilde{\pi}(\mathbb{V})$. It is therefore convenient to collapse all intervals on which some iterate $\tilde{f}^n$ is constant. Because closed intervals are collapsed to points, the quotient of the interval $\tilde{I}$ is either a point or an interval I. In the latter case $\tilde{f}$ induces a continuous piecewise strictly monotone interval map $f\colon I \circlearrowleft$, which is called the *quotient* of $F\colon (S^2, \mathbb{I}) \circlearrowleft$. All thick interval maps considered in this paper will have an interval map for quotient (see Section 2.2). Note that if $\pi\colon \mathbb{I} \to I$ denotes the canonical projection, then $\pi \circ F|_{\mathbb{I}} = f \circ \pi$.

Example 1 The first example is Smale's horseshoe map which will be denoted $F_0\colon (S^2, \mathbb{I}) \circlearrowleft$ here and in what follows. It is shown in Figure 3. The horseshoe has three fixed points, denoted x, x_0, and x_1. The fixed point x is attracting (condition vi)), while x_0 and x_1 are saddles by conditions iv) and vi). The quotient interval map — the tent map $f_0\colon I \circlearrowleft$ — is also shown in the figure.

Example 2 Figure 4 depicts an example of a thick interval map associated to a horseshoe periodic orbit: such thick interval maps are the starting point for the construction of unimodal generalized train tracks described in Section 6. The interval endomorphism f is one with kneading sequence $\kappa(f) = (1001011)^\infty$ (see Section 4).

2.2 The MIA property

First the basic concepts of the Perron-Frobenius theory for non-negative integer matrices are described (see [11]). Let M be a square matrix with non-negative

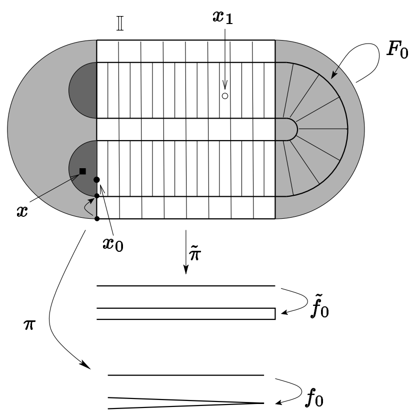


Figure 3: The horseshoe map

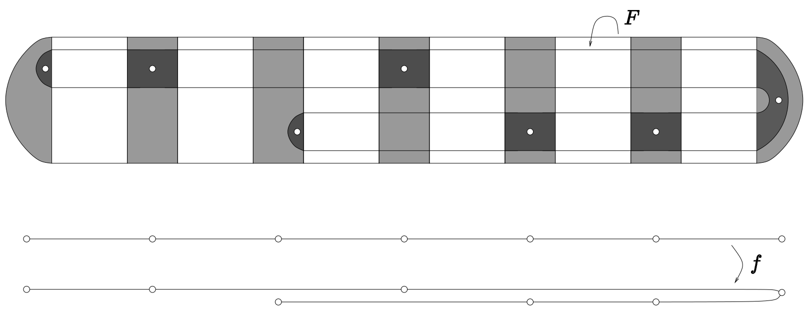


Figure 4: A thick interval map associated to a horseshoe periodic orbit

integer entries, which is not equal to the 1×1 matrix (1). M is said to be *reducible* if, by a permutation of the index set, it is possible to put it in triangular block form:

$$M = \begin{pmatrix} A & 0 \\ B & C \end{pmatrix},$$

with A and C non-trivial square matrices. Otherwise, M is said to be *irreducible*. An irreducible matrix M has a unique positive eigenvector (up to scaling), and the associated eigenvalue λ, called the *Perron-Frobenius eigenvalue* of M, is simple and is equal to the spectral radius of M. M is *irreducible and aperiodic* if there is a positive integer k such that every entry of M^k is positive. In this case the Perron-Frobenius eigenvalue λ satisfies $\lambda > 1$, and is the only eigenvalue of M on the circle $\{z \in \mathbb{C} : |z| = \lambda\}$.

Let $F\colon (S^2, \mathbb{I}) \circlearrowleft$ be a thick interval map. Associate a *transition matrix* $M =$

(m_{ij}) to F in the following way: let $s_1, s_2, \ldots, s_n$ be the strips of $\mathbb{I}$ and set

$$m_{ij} = \text{number of times } F(s_j) \text{ crosses } s_i.$$

Definition 3 A thick interval map is *MIA* if its associated transition matrix is irreducible and aperiodic (the 'M' stands for Markov).

If F is MIA then the quotient interval I is not a point, and indeed the projection $\pi(s_j)$ of each strip of $\mathbb{I}$ is a non-trivial interval. For there are integers n_j such that $F^{n_j}(s_j)$ crosses s_j at least twice, with horizontal expansion and vertical contraction bounded away from 1. Thus the set of points of s_j whose forward F^{n_j}-orbits remain in s_j contains $h_{s_j}^{-1}(C \times [0,1])$ for some Cantor set $C \subseteq [0,1]$, and $\pi(h_{s_j}^{-1}(x_1, y_1)) \neq \pi(h_{s_j}^{-1}(x_2, y_2))$ whenever x_1, x_2 are distinct points of C and $y_1, y_2 \in [0,1]$.

3 Generalized train tracks

This section contains a description of generalized train tracks associated to thick interval maps. It follows [4], where a description of generalized train tracks associated to general thick graph maps is given.

Let $\mathbb{I}$ be a thick interval and $A \subset \mathbb{I}$ be a finite *puncture* set, each of whose points lies in the interior of a junction and such that no junction contains more than one point of A. For each strip s of $\mathbb{I}$, let γ_s be (the image of) the arc $t \mapsto h_s^{-1}(t, 1/2)$, which joins the two boundary components of s in $\mathbb{I}$ and intersects each leaf of s exactly once. Let RE denote the set of these arcs. The endpoints of the arcs γ_s are called *switches* and the set of switches is denoted L.

Definitions 4 A *generalized train track* $\tau \subseteq \mathbb{I} \setminus A$ is a graph with vertex set L and countably many edges, each of which intersects $\partial\mathbb{V}$ only at L, such that

i) The edges of τ which intersect the interior of $\mathbb{E}$ are precisely the elements of RE, and
ii) No two edges e_1, e_2 contained in a given junction J are *parallel*: that is, they do not bound a disk which contains no point of A or other edges.

Two generalized train tracks τ and τ' are *equivalent*, denoted $\tau \sim \tau'$, if they are isotopic by an isotopy supported on $\mathbb{V} \setminus A$.

The edges of τ which are contained in $\mathbb{E}$ (that is, the elements of RE) are called *real*, and the others (contained in $\mathbb{V}$) are called *infinitesimal*. Let IE denote the set of infinitesimal edges of τ.

A generalized train track τ is *finite* if it has only finitely many edges. An infinitesimal edge is called a *loop* if its two endpoints coincide, and a *bubble* if in addition it bounds an open disk which is disjoint from τ. A bubble of τ is *homotopically trivial* if this disk contains no point of A, and is *homotopically non-trivial* otherwise.

Clearly τ is determined by its infinitesimal edges, and may be written $\tau = \tau(\mathrm{IE})$ when the thick interval $\mathbb{I}$ and the set A are clear from the context.

Note that τ is not required to be connected: while the generalized train tracks constructed below will always be connected, disconnected ones are needed during the construction.

Definitions 5 A homotopy $\{\alpha_t\}$ of a path $\alpha\colon [0,1] \to X$ is said to be *relative to* $U \subset X$ if, for each $s \in [0,1]$, $\alpha_0(s) \in U$ implies $\alpha_t(s) \in U$ for all t. Let $[\alpha]$ be a homotopy class of paths in $S^2 \setminus A$ relative to $\partial_{\mathbb{I}}\mathbb{V}$, with endpoints in $\partial_{\mathbb{I}}\mathbb{V}$. Then $[\alpha]$ is *carried* by a generalized train track τ if it can be realized by an edge-path in τ with alternating real and infinitesimal edges.

Generalized train tracks are normally drawn in such a way as to suggest that if a real edge γ_s and an infinitesimal edge e share a common endpoint, then their union is a smooth (branched if the endpoints of e coincide) 1-manifold. With this intuition, $[\alpha]$ is carried by τ if it can be realized by a smooth path in τ (or, even more intuitively, by a train running along τ). It is more convenient, however, to express this smoothness combinatorially as above.

The next aim is to define the image of τ under a thick interval map F. In order that this image should itself be a generalized train track, it is necessary to apply pseudo-isotopies to $F(\tau)$ so as to squash $F(\tau) \cap \mathbb{E}$ onto the real edges, and amalgamate pairs of parallel edges.

Let $(\mathbb{I}, A)$ be a thick interval together with a puncture set. On each strip s of $\mathbb{I}$ define $\psi_s\colon\ \bar{s} \times [0,1] \to \bar{s}$ using the coordinates h_s by

$$\psi_s(x,y,t) = (x, (1-t)y + t/2)$$

(thus $\psi_s(\cdot, 1)$ maps $\bar{s}$ onto γ_s). Extend these maps to a pseudo-isotopy $\psi_0\colon\ S^2 \times [0,1] \to S^2$ in the following way: first extend the ψ_s to mutually disjoint disk neighbourhoods $U_s \supset \bar{s}$, with $U_s \subset S^2 \setminus A$, so that they are isotopies on $U_s \setminus \bar{s}$, the identity on ∂U_s, and send $\mathbb{V}$ into $\mathbb{V}$; then extend them to be the identity elsewhere.

If τ is a graph in $\mathbb{I} \setminus A$ which satisfies the definition of a generalized train track except that there are finitely many pairs (e, e') of parallel edges, define a pseudo-isotopy $\psi_{e,e'}$ supported on a neighbourhood in $\mathbb{V}$ of the closed disk Δ bounded by $e \cup e'$ with the property that $\psi_{e,e'}(e, 1) = \psi_{e,e'}(e', 1)$ but no points outside Δ are identified by $\psi_{e,e'}(\cdot, 1)$. Since this pseudo-isotopy introduces no new pairs of parallel edges, it is possible to define a pseudo-isotopy $\psi_1 \colon S^2 \times [0,1] \to S^2$ (whose dependence on τ is suppressed) by composing successive pseudo-isotopies on each pair of parallel edges, with the property that $\psi_1(\tau, 1)$ is a generalized train track.

In this paper, the puncture set A will always be taken to be the set of attracting periodic points of the thick interval map F (so in particular $F(A) = A$). Conditions iv) and vi) in the definition of a thick interval map ensure that the points of A are contained in the interiors of the junctions of $\mathbb{I}$, with at most one point of A in each junction.

Definitions 6 Let $F \colon (S^2, \mathbb{I}, A) \circlearrowleft$ be a thick interval map, where A is the set of attracting periodic points of F, and let $\tau \subseteq \mathbb{I} \setminus A$ be a generalized train track. Since F restricts to an embedding $(\mathbb{V}, A) \circlearrowleft$ and the underlying interval endomorphism is piecewise monotone, there can be only finitely many pairs of parallel edges in $\psi_0(F(\tau), 1)$, and the *image* $F_*(\tau)$ of τ under F is defined as $F_*(\tau) = \psi(\tau)$, where the near-homeomorphism $\psi \colon S^2 \circlearrowleft$ is given by $\psi(x) = \psi_1(\psi_0(F(x), 1), 1)$.

A generalized train track τ is *F-invariant* if $F_*(\tau)$ is equivalent to τ.

If τ is F-invariant, then by definition there is a homeomorphism $H \colon S^2 \circlearrowleft$, isotopic to the identity by an isotopy supported on $\mathbb{V} \setminus A$, such that $H(\psi(\tau)) = \tau$. The *train track map* $\phi \colon \tau \circlearrowleft$ associated to $F \colon (S^2, \mathbb{I}, A) \circlearrowleft$ is the restriction of $H \circ \psi$ to τ, well-defined up to homotopy relative to the vertices of τ.

The following straightforward result guarantees the existence of invariant generalized train tracks, and its proof provides a method for constructing them.

Theorem 1 *Let $F \colon (S^2, \mathbb{I}, A) \circlearrowleft$ be a thick interval map, where A is the set of attracting periodic points of F. Then there exists an F-invariant generalized train track $\tau \subseteq \mathbb{I} \setminus A$.*

Proof Let $\tau_0 = \tau(\emptyset)$ be the generalized train track with no infinitesimal edges. By definition, $\tau_0 \subseteq \tau_1 = F_*(\tau_0)$. Let $\tau_2' = F_*(\tau_1)$. Then τ_1 need not be a subset

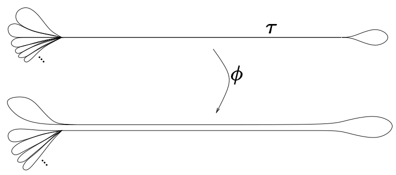


Figure 5: The invariant train track for the horseshoe

of τ_2', since τ_2' is only defined up to equivalence, but there is a generalized train track $\tau_2 \sim \tau_2'$ with $\tau_1 \subseteq \tau_2$. Continuing in this way, construct a nested sequence (τ_n) of generalized train tracks with $F_*(\tau_n) \sim \tau_{n+1}$ for all n. Then $\tau = \bigcup_{n \geq 0} \tau_n$ is a generalized train track which satisfies $F_*(\tau) \sim \tau$. □

The F-invariant generalized train track τ constructed in this proof is minimal, in the sense that it contains only those edges which arise as images of its real edges: more precisely, if τ' is also F-invariant then there is a subset IE of the infinitesimal edges of τ' such that $\tau \sim \tau(\mathrm{IE})$. For this reason τ is referred to as *the F-invariant generalized train track.*

Example 3 The invariant generalized train track for the horseshoe map F_0 is shown in Figure 5. The set A consists of the fixed point x contained in the left 1-junction of $\mathbb{I}$. No bubble encloses it and it is not shown in the figure. It is instructive to construct this train track starting from $\tau(\emptyset)$ as described in the proof of Theorem 1.

The invariant train track τ and the train track map ϕ can be thought of as more careful 1-dimensional representations of the thick interval $\mathbb{I}$ and the thick interval map $F\colon (S^2, \mathbb{I}, A) \circlearrowleft$. Whereas $f\colon I \circlearrowleft$ does not pay attention to junctions — they are collapsed to points — the map $\phi\colon \tau \circlearrowleft$ gives a careful account of the behaviour of the images of strips under iterates of F inside the junctions.

Definition 7 An *infinitesimal polygon* of τ is a component of $S^2 \setminus \tau$ bounded by finitely many infinitesimal edges. It is called an *n-gon* if it is bounded by n infinitesimal edges (see Figure 6, in which the n-gons are shaded).

Remark 2 Bigons (2-gons) can only occur if they contain a point of A (otherwise the two edges bounding the bigon would be parallel).

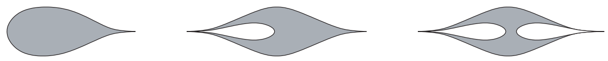


Figure 6: Examples of n-gons for $n = 1, 3, 4$

4 Unimodal maps, symbolic dynamics and the horseshoe

This section contains a summary of the theory of unimodal maps, symbolic dynamics, and the Smale horseshoe which will be used later. Detailed accounts can be found in [8, 15, 7, 13] (unimodal maps) and [8, 12] (horseshoe).

4.1 Unimodal maps

Let $I = [a, b]$. For the purposes of this paper, a *unimodal map* on I is a continuous surjection $f\colon I \circlearrowleft$ with $f(b) = a$, for which there exists $c \in (a, b)$ such that f is strictly increasing on $[a, c]$ and strictly decreasing on $[c, b]$ (and in particular $f(c) = b$). The point c is the *critical* or *turning point* of f, and $b = f(c)$ is its *critical value.*

The conditions that f be surjective and that $f(b) = a$ are not standard. However, a unimodal map f which doesn't satisfy them has trivial dynamics outside of its *dynamical interval* $[f^2(c), f(c)]$, and hence there is no loss of generality, and a gain of convenience, in adding these requirements. Note also that some authors don't require the monotonicity on $[a, c]$ and $[c, b]$ to be strict, and would refer to a map defined as above as *strictly unimodal.*

Symbolic dynamics for unimodal maps is introduced by defining the *itinerary* of a point $x \in I$ to be the sequence

$$i(x) = s_0 s_1 s_2 \ldots \in \{0, 1\}^{\mathbb{N}}$$

given by

$$s_j = \begin{cases} 0 & \text{if } f^j(x) < c \\ 1 & \text{if } f^j(x) \geq c. \end{cases}$$

This defines a map $i\colon I \to \{0, 1\}^{\mathbb{N}}$ with the property that $i \circ f = \sigma \circ i$, where $\sigma\colon \{0, 1\}^{\mathbb{N}} \circlearrowleft$ is the *shift map* given by $\sigma(s_0 s_1 s_2 \ldots) = s_1 s_2 s_3 \ldots$.

Remark 3 In other contexts it is useful to define the itinerary $i(x)$ to lie in $\{0, C, 1\}^{\mathbb{N}}$, where $i_j(x) = C$ if $f^j(x) = c$. This distinction will not be necessary here.

The *unimodal order* $\preceq$ is a total order (but not a well-ordering) on $\{0,1\}^{\mathbb{N}}$ which reflects the usual ordering of points in I. If s and t are distinct elements of $\{0,1\}^{\mathbb{N}}$, then $s \prec t$ if and only if $\sum_{j=0}^{n} s_j$ is even, where n is least such that $s_n \neq t_n$. For general $s, t \in \{0,1\}^{\mathbb{N}}$, define $s \preceq t$ if and only if $s \prec t$ or $s = t$.

The unimodal order is defined precisely in order that $i(x) \prec i(y) \implies x < y$ for all $x, y \in I$. It follows immediately that $x < y \implies i(x) \preceq i(y)$: the possibility of equality in this partial converse cannot be excluded in general. A unimodal map f is said to *have no homtervals* if the full converse holds, ie, $x < y \iff i(x) \prec i(y)$ (a *homterval* is a non-trivial interval J with $c \notin f^n(J)$ for all $n \geq 0$).

The itinerary of the critical value of a unimodal map $f\colon I \circlearrowleft$ plays a particularly important role: the *kneading sequence* $\kappa(f)$ of a unimodal map $f\colon I \circlearrowleft$ is the itinerary $\kappa(f) = i(b)$. Since $\sigma(\kappa(f)) = i(a)$, it follows that $\sigma(\kappa(f)) \preceq i(x) \preceq \kappa(f)$ for all $x \in I$, and in particular $\sigma(\kappa(f)) \preceq \sigma^n(\kappa(f)) \preceq \kappa(f)$ for all $n \geq 0$. This statement characterizes those elements of $\{0,1\}^{\mathbb{N}}$ which are the kneading sequence of some unimodal map, as expressed by the following definition and theorem (which is a translation into the language of this paper of Theorem 12.1 of [15]).

Definition 8 An element s of $\{0,1\}^{\mathbb{N}}$ is *a kneading sequence* if $\sigma(s) \preceq \sigma^n(s) \preceq s$ for all $n \in \mathbb{N}$.

Theorem 2 *An element s of $\{0,1\}^{\mathbb{N}}$ is equal to $\kappa(f)$ for some unimodal map f if and only if it is a kneading sequence.*

Definition 9 An MIA thick interval map $F\colon (S^2, \mathbb{I}) \circlearrowleft$ is *unimodal* if its quotient interval endomorphism $f\colon I \circlearrowleft$ is unimodal. In this case, define the *kneading sequence* $\kappa(F)$ of F to be $\kappa(f)$.

The quotient $f\colon I \circlearrowleft$ of a unimodal thick interval map has no homtervals. For since $F\colon (S^2, \mathbb{I}) \circlearrowleft$ is MIA, there is some n such that f^n expands each subinterval J of I for which $c \notin \bigcup_{i=0}^{n-1} f^i(J)$, the expansion being bounded away from 1 with respect to the projections of the coordinate maps h_s on the strips of $\mathbb{I}$.

Because thick intervals as defined here have only finitely many junctions, the orbits of the vertices of I under f are finite. In particular, the kneading sequence of a unimodal thick interval map is always periodic or preperiodic. The set of periodic or preperiodic kneading sequences whose associated transition matrices are irreducible and aperiodic will be denoted $\mathcal{MIA}$. Notice that given a kneading sequence in $\mathcal{MIA}$, there is a natural construction of a unimodal MIA thick interval map with that kneading sequence: contruct a piecewise affine interval map with the given kneading sequence and thicken it.

4.2 Symbolic dynamics for the horseshoe

This section contains a brief summary of the application of symbolic dynamics to the horseshoe map of Example 1: see [8] for a more detailed description. The set $\Lambda = \bigcap_{n\in\mathbb{Z}} F_0^n(\overline{s})$ (where s is the strip of $\mathbb{I}$) is a Cantor set, and an itinerary homeomorphism $i\colon \Lambda \to \{0,1\}^{\mathbb{Z}}$ is defined by setting

$$i_j(z) = \begin{cases} 0 & \text{if } F_0^j(z) \in V_0 \\ 1 & \text{if } F_0^j(z) \in V_1 \end{cases}$$

where V_0, V_1 are the left and right connected components of $F_0^{-1}(s) \cap s$ respectively. The itinerary homeomorphism conjugates $F_0|_\Lambda$ with the shift map $\sigma\colon \{0,1\}^{\mathbb{Z}} \circlearrowleft$, defined by

$$\sigma(\ldots s_{-2}s_{-1} \cdot s_0 s_1 \ldots) = \ldots s_{-1}s_0 \cdot s_1 s_2 \ldots$$

Let $f_0\colon I \circlearrowleft$ be the quotient interval map (which is conjugate to a full tent map). Then the invariant Cantor set Λ of F_0 inside s projects to the whole interval I, and this projection establishes a 1-1 correspondence between the F_0-periodic orbits in Λ and the f_0-periodic orbits in I.

Note that for all $z \in \Lambda$, $i(\pi(z))$ is obtained from $i(z)$ by deleting all symbols before the origin. The correspondence between periodic orbits is reflected in the correspondence between itineraries in the obvious way: the itinerary of a period n periodic point of F_0 is of the form $\overline{w}$ for some length n word w, and the itinerary of its projection is w^∞. This correspondence will be invoked without further comment in the remainder of the paper.

A periodic orbit P of F_0 of (least) period n is described by its *code* $c_P \in \{0,1\}^n$, which is given by the first n symbols of the itinerary $i(p)$ of its rightmost point p: thus, for example, the period 5 orbit which contains the point with itinerary $\overline{01001}$ has code 10010. A word $w \in \{0,1\}^n$ is therefore the code of a period n horseshoe orbit if and only if it is maximal in the sense of the following definition.

Definition 10 $w \in \{0,1\}^n$ is *maximal* if $\sigma^i(w^\infty) \prec w^\infty$ for $1 \le i < n$.

It follows that if w is the code of a periodic orbit of F_0, then w^∞ is a kneading sequence.

4.3 Braid types

Let $F\colon S^2 \circlearrowleft$ be a homeomorphism (which in this paper will always be orientation-preserving), and A be a finite F-invariant subset of S^2. Then the *braid type* $\mathrm{bt}(A,F)$ is defined to be the isotopy class of F relative to A up to topological change of coordinates [3]: that is, $\mathrm{bt}(A,F) = \mathrm{bt}(B,G)$ if and only if there is an orientation-preserving homeomorphism $h\colon S^2 \circlearrowleft$ with $h(A) = B$ such that $h \circ F \circ h^{-1}$ is isotopic to G relative to B. Using Thurston's classification theorem for surface homeomorphisms [17], braid types can be classified as finite order, reducible, or pseudo-Anosov.

The braid type of a periodic orbit P of the horseshoe map $F_0\colon S^2 \circlearrowleft$ is defined to be $\mathrm{bt}(P \cup \{\infty\}, F_0)$. Thus horseshoe periodic orbits can also be classified as finite order, reducible, or pseudo-Anosov.

4.4 Height

The description of unimodal generalized train tracks in Section 6 depends upon the notion of the *height* of an element s of $\mathcal{MIA}$. The height is defined using words $c_q \in \{0,1\}^{n+1}$ associated to each rational $q = m/n \in (0, 1/2]$. Motivation for the definition can be found in [12].

Definition 11 Given $q = m/n \in \mathbb{Q} \cap (0, 1/2]$, define a word $c_q \in \{0,1\}^{n+1}$ as follows. Let L_q be the straight line in $\mathbb{R}^2$ from $(0,0)$ to (n,m). For $0 \le i \le n$, let $s_i = 1$ if L_q crosses some line $y =$ integer for $x \in (i-1, i+1)$, and $s_i = 0$ otherwise. Then $c_q = s_0 s_1 \dots s_n$.

Example 4 Figure 7 shows that $c_{3/10} = 10011011001$.

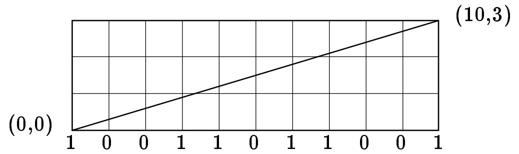


Figure 7: $c_{3/10} = 10011011001$

The words c_q are manifestly palindromic. Their general form is indicated by the examples in Table 1, in which the column headings and row headings denote the numerator and denominator of q respectively. The $n - 2m + 1$ zeros are

	1	2	3	4
3	1001			
4	10001			
5	100001	101101		
6	1000001			
7	10000001	10011001	10111101	
8	100000001		101101101	
9	1000000001	1000110001		1011111101
10	10000000001		10011011001	
11	100000000001	100001100001	100110011001	101101101101

Table 1: Examples of the words c_q

partitioned 'as even-handedly as possible' into m subwords (possibly empty), separated by 11.

The following description of the words c_q is easily shown to be equivalent: given $q = m/n$, define integers $\kappa_i(q)$ for $1 \le i \le m$ by

$$\kappa_i(q) = \begin{cases} \lfloor 1/q \rfloor - 1 & \text{if } i = 1 \\ \lfloor i/q \rfloor - \lfloor (i-1)/q \rfloor - 2 & \text{if } 2 \le i \le m \end{cases} \tag{1}$$

($\lfloor x \rfloor$ denotes the greatest integer which does not exceed x). Then

$$c_q = 10^{\kappa_1(q)}110^{\kappa_2(q)}11\dots110^{\kappa_m(q)}1.$$

Example 5 Let $q = 3/7$, so $\kappa_1(q) = \lfloor 7/3 \rfloor - 1 = 1$, $\kappa_2(q) = \lfloor 14/3 \rfloor - \lfloor 7/3 \rfloor - 2 = 4 - 2 - 2 = 0$, and $\kappa_3(q) = \lfloor 21/3 \rfloor - \lfloor 14/3 \rfloor - 2 = 7 - 4 - 2 = 1$. Thus $c_{3/7} = 10111101$.

The next lemma [12] motivates the definition of height: in particular, it will imply that the height function $q\colon \{0,1\}^{\mathbb{N}} \to [0,1/2]$ is decreasing with respect to the unimodal order $\prec$ on $\{0,1\}^{\mathbb{N}}$ and the usual order on $[0,1/2]$.

Lemma 3 *For each $q \in \mathbb{Q} \cap (0,1/2]$, the word c_q1 is maximal. Moreover, if $q, r \in \mathbb{Q} \cap (0,1/2]$ with $q < r$ then $(c_r1)^\infty \prec (c_q1)^\infty$.*

Definitions 12 Let $c \in \{0,1\}^{\mathbb{N}}$. Then the *height* $q(c) \in [0,1/2]$ of c is given by

$$q(c) = \inf\{q \in \mathbb{Q} \cap (0,1/2] \;:\; q = 1/2 \text{ or } (c_q1)^\infty \prec c\}.$$

If P is a horseshoe periodic orbit of period $n \ge 2$ with code c_P, then the *height* $q(P)$ of P is given by $q(P) = q(c_P^\infty)$.

The height of a horseshoe periodic orbit is a braid type invariant. The height of any element of $\mathcal{MIA}$ is rational, and can be computed using an algorithm described in [12].

The next result describes the kneading sequences with given rational height [12].

Definition 13 For each $m/n \in (0,1/2)$, define $w_{m/n} \in \{0,1\}^{n-1}$ to be the word obtained by deleting the last two symbols of $c_{m/n}$, and $\widehat{w}_{m/n} \in \{0,1\}^{n-1}$ to be the reverse of $w_{m/n}$.

Theorem 4 *Let $s \in \mathcal{MIA}$. Then $q(s) \in [0,1/2) \cap \mathbb{Q}$, and*

a) *$q(s) = 0$ if and only if $s = 10^\infty$.*
b) *If $m/n \in (0,1/2)$, then $q(s) = m/n$ if and only if*

$$(w_{m/n}1)^\infty \preceq s \preceq w_{m/n}01(1\widehat{w}_{m/n})^\infty.$$

Note that $w_{m/n}01(1\widehat{w}_{m/n})^\infty$ can be written more concisely as $c_{m/n}(1\widehat{w}_{m/n})^\infty$, but the former expression is more suggestive in calculations, as in the proof of Lemma 5 below. In fact, although it isn't immediately apparent from the definitions, this kneading sequence is preperiodic to $(w_{m/n}1)^\infty$, ie, there is an integer j with $\sigma^j((1\widehat{w}_{m/n})^\infty) = (w_{m/n}1)^\infty$.

The endpoints of the intervals of kneading sequences of given height will be important in the remainder of the paper, as will the kneading sequences $(c_{m/n}1)^\infty$ used to define the height. The acronym NBT in the following definitions stands for 'no bogus transitions', and reflects the original motivation of these ideas.

Definitions 14 Let $m/n \in (0,1/2)$. Then write $\mathrm{NBT}(m/n) = (c_{m/n}1)^\infty$, $\mathrm{lhe}(m/n) = (w_{m/n}1)^\infty$, $\mathrm{rhe}(m/n) = c_{m/n}(1\widehat{w}_{m/n})^\infty$, and $\mathrm{KS}(m/n)$ for the set of kneading sequences $s \in \mathcal{MIA}$ with $\mathrm{lhe}(m/n) \preceq s \preceq \mathrm{rhe}(m/n)$ (ie, the set of kneading sequences of height m/n).

Example 6 Let $m/n = 2/7$, so $c_{2/7} = 10011001$, $w_{2/7} = 100110$, and $\widehat{w}_{2/7} = 011001$. Then $\mathrm{NBT}(2/7) = (100110011)^\infty$, $\mathrm{lhe}(2/7) = (1001101)^\infty$, and $\mathrm{rhe}(2/7) = 10011001(1011001)^\infty = 10(0110011)^\infty$. An element s of $\mathcal{MIA}$ lies in $\mathrm{KS}(2/7)$ (ie, has height $2/7$) if and only if

$$(1001101)^\infty \preceq s \preceq 10(0110011)^\infty.$$

Notice that by Theorem 4,

$$\mathcal{MIA} = \{10^\infty\} \cup \bigcup_{m/n \in (0,1/2)} \mathrm{KS}(m/n).$$

Lemma 5 *Let $s \in \mathcal{MIA}$ have height $q = m/n > 0$. Then either $s = (w_q1)^\infty$ or s has c_q as an initial word. In particular, if s is periodic then either $s = \mathrm{lhe}(m/n)$ (period n), or $s = \mathrm{NBT}(m/n)$ (period $n+2$), or the period of s is at least $n+3$.*

Proof By Theorem 4, s has w_q as an initial word. If the next symbol is 0, then since w_q is odd and $s \preceq w_q01(1\widehat{w}_q)^\infty$, it follows that $s = w_q01\ldots$, ie, $s = c_q\ldots$. If the next symbol is 1, then either $s = (w_q1)^\infty$, or let $k \geq 1$ be the greatest integer such that $s = (w_q1)^kd$ for some $d \in \{0,1\}^{\mathbb{N}}$. Then $d \preceq s$ (since s is a kneading sequence), but $d \succeq (w_q1)^\infty$ (by Theorem 4), and hence $d = w_q1\ldots$, contradicting the definition of k. The proof of the final statement follows readily, and can be found in [12] (Theorem 3.5). $\square$

5 The outside dynamics of a unimodal map

The unimodal maps considered in this paper are destined to be thickened into thick interval maps, and as such have an implicit two-dimensional structure: the interval $[a,b]$ is thought of as being folded at the critical point c, and laid down over itself in such a way that points to the right of c end up above points to the left. Thus all of the points below the interval and some of the points above it (namely those whose image is to the left of $f(a)$) remain outside the interval, whereas the other points above the interval are trapped in the fold (see the left hand side of Figure 8 for clarification). The aim of this section is to formalize this intuitive idea, and to analyze the dynamics of points whose entire orbits remain outside the interval. This will play an important role in the construction of generalized pseudo-Anosovs, when it is necessary to 'sew up the outside boundary'.

The outside of the interval is represented by a circle S^1 obtained by gluing together two copies of $[a,b]$ at their endpoints, and the unimodal map $f\colon [a,b] \circlearrowleft$ induces a map $\theta\colon S^1 \to S^1 \cup [a,b]$, reflecting the two possible fates of the image of a point on the outside: to remain outside, or to be folded inside the interval.

Let $f\colon I = [a,b] \circlearrowleft$ be a unimodal map with critical point c. Let S^1 be the unit circle in $\mathbb{R}^2$, coordinatize both the upper and lower halves of S^1 with coordinates in I in such a way that $(-1,0)$ has coordinate a and $(1,0)$ has coordinate b, and let $\pi\colon S^1 \to I$ take each point of S^1 to the point of I given by its coordinate. Thus $\pi^{-1}(a)$ and $\pi^{-1}(b)$ each contain a single point, denoted $\hat{a}$ and $\hat{b}$ respectively, while for an interior point x of I, $\pi^{-1}(x)$ contains two points, denoted x_u (in the upper half circle) and x_l (in the lower half circle).

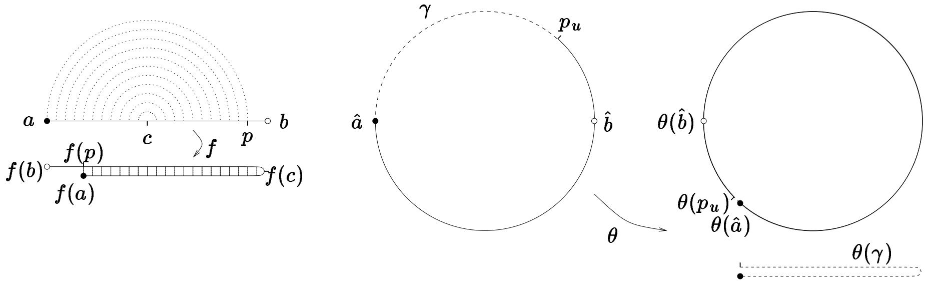


Figure 8: A unimodal map and the induced outside map

Let $p \in I$ be the point with $p > c$ and $f(p) = f(a)$. The (discontinuous) function $\theta\colon S^1 \to S^1 \cup I$ is defined by

$$\begin{aligned} \theta(\hat{a}) &= f(a)_l \\ \theta(\hat{b}) &= \hat{a} \\ \theta(x_u) &= f(x) \quad \text{if } x < p \\ \theta(x_u) &= f(x)_l \quad \text{if } x \geq p \\ \theta(x_l) &= f(x)_l \quad \text{if } x < c \\ \theta(c_l) &= \hat{b} \\ \theta(x_l) &= f(x)_u \quad \text{if } x > c, \end{aligned}$$

reflecting the intuitive notion of the action of f on the outside of the interval (see Figure 8), and in particular satisfying $\pi \circ \theta = f \circ \pi$. Note that the set of points mapped by θ into I is an open interval $\gamma = \{x_u : a < x < p\}$, and that the complement of γ is mapped strictly monotonically onto S^1, with both endpoints mapping to $f(a)_l$.

The following result describes the relationship between the dynamics on the outside and the height of $\kappa(f)$ which will be used later. If $x, y \in S^1$ then the notation (x, y) always denotes the open interval with endpoints x and y which is disjoint from γ (so $x, y \notin \gamma$), while $(x, y)_\gamma$ denotes the interval contained in γ (so $x, y \in \overline{\gamma}$). Similar notation is used for half-open and closed intervals: intervals which intersect but are not contained in γ are not used.

Theorem 6 *Let $f\colon [a,b] \circlearrowleft$ be the quotient of a unimodal thick interval map with $s = \kappa(f) \in \mathcal{MIA}$, and let $\theta\colon S^1 \to S^1 \cup I$ be the induced map on the outside. Let $q(s) = m/n > 0$. Then $\theta^i(\hat{a}) \notin \gamma$ for $0 \leq i < n$, while $\theta^n(\hat{a}) \in \overline{\gamma}$, with*

i) $\theta^n(\hat{a}) = \hat{a} \iff s = \mathrm{lhe}(m/n)$

ii) $\theta^n(\hat{a}) = c_u \iff s = \mathrm{NBT}(m/n)$
iii) $\theta^n(\hat{a}) = p_u \iff s = \mathrm{rhe}(m/n)$
iv) $\theta^n(\hat{a}) \in (\hat{a}, c_u)_\gamma \iff \mathrm{lhe}(m/n) \prec s \prec \mathrm{NBT}(m/n)$
v) $\theta^n(\hat{a}) \in (c_u, p_u)_\gamma \iff \mathrm{NBT}(m/n) \prec s \prec \mathrm{rhe}(m/n)$.

Let $\Lambda = \bigcap_{n=0}^{\infty} \theta^{-n}(S^1)$ be the set of points whose θ-orbits remain in S^1. Then Λ contains exactly one periodic orbit, which has period n and whose points are permuted as by a rigid rotation of the circle through angle $2\pi m/n$. If $\mathrm{lhe}(m/n) \prec s \prec \mathrm{rhe}(m/n)$ then there are no other points in Λ, while if $s = \mathrm{lhe}(m/n)$ (respectively $s = \mathrm{rhe}(m/n)$) then Λ is the union of this periodic orbit and the set $\{\theta^{-i}(p_u) : i \geq 0\}$ (respectively $\{\theta^{-i}(\hat{a}) : i \geq 0\}$).

Notice in particular that the periodic orbit in Λ is the orbit of $\hat{a}$ in the case when $s = \mathrm{lhe}(m/n)$, and is the orbit of p_u when $s = \mathrm{rhe}(m/n)$ (since $\theta(p_u) = \theta(\hat{a})$).

Three lemmas are used in the proof. The integers $\kappa_i(q)$ in the statement of the first are as defined by (1).

Lemma 7 *Let $q = m/n \in \mathbb{Q} \cap (0, 1/2)$. For each integer r with $1 \leq r \leq m$, the word*

$$10^{\kappa_r(q)}1^2 0^{\kappa_{r+1}(q)}1^2 \ldots 1^2 0^{\kappa_m(q)}1$$

disagrees with the word

$$10^{\kappa_1(q)-1}1^2 0^{\kappa_2(q)}1^2 \ldots 1^2 0^{\kappa_m(q)}1$$

within the shorter of their lengths, and is greater than it in the unimodal order.

Proof (See also [6], Lemma 63.) If $r = 1$ then the result is obvious, so suppose that $1 < r \leq m$. If the two words didn't disagree, then it would follow that $\kappa_1(q) = \kappa_r(q) + 1$ and that $\kappa_m(q) = \kappa_{m-r+1}(q)$, contradicting the fact that c_q is palindromic.

Observe that formula (1) gives, for each s with $1 \leq s \leq m + 1 - r$,

$$(\kappa_1(q) - 1) + \sum_{i=2}^{s} \kappa_i(q) = \left\lfloor \frac{s}{q} \right\rfloor - 2s,$$

and

$$\begin{aligned} \sum_{i=r}^{r+s-1} \kappa_i(q) &= \left\lfloor \frac{r+s-1}{q} \right\rfloor - \left\lfloor \frac{r-1}{q} \right\rfloor - 2s \\ &\geq \left\lfloor \frac{r+s-1}{q} - \frac{r-1}{q} \right\rfloor - 2s \\ &= \left\lfloor \frac{s}{q} \right\rfloor - 2s. \end{aligned}$$

Hence at the point where they first disagree, $10^{\kappa_r(q)}1^20^{\kappa_{r+1}(q)}1^2\dots1^20^{\kappa_m(q)}1$ has a longer block of 0s than $10^{\kappa_1(q)-1}1^20^{\kappa_2(q)}1^2\dots1^20^{\kappa_m(q)}1$, and so is greater in the unimodal order. $\square$

Lemma 8 *Let $s \in \mathcal{MIA}$ have height $q = q(s) > 0$, and suppose $s = c_q d$ for some $d \in \{0,1\}^{\mathbb{N}}$. Then $\sigma(d) \succeq \sigma^2(s)$ and $d \preceq 1\sigma^2(s)$.*

Proof Suppose first that the first symbol of d is 0. In this case it is obvious that $d \preceq 1\sigma^2(s)$, so, writing $d = 0e$, it is only necessary to show that $e \succeq \sigma^2(s) = \widehat{w}_q 0e$. Now $\sigma^n(s) = 10e \preceq s$, since s is a kneading sequence, and $s = 10\dots$, so $\sigma^{n+2}(s) = e \succeq \sigma^2(s)$ as required.

If $d = 1e$ for some $e \in \{0,1\}^{\mathbb{N}}$ then $\sigma(d) \succeq \sigma^2(s)$ clearly implies that $d \preceq 1\sigma^2(s)$, so again it is required to show that $e \succeq \sigma^2(s) = \widehat{w}_q 1e$.

Assume for a contradiction that $e \prec \widehat{w}_q 1e$. Now $s = c_q 1e \preceq c_q(1\widehat{w}_q)^\infty$ by Theorem 4, and so $e \succeq (\widehat{w}_q 1)^\infty$ (since $c_q 1$ is an odd word). Thus

$$(\widehat{w}_q 1)^\infty \preceq e \prec \widehat{w}_q 1e,$$

and so e has $\widehat{w}_q 1$ as an initial word. Let $k \geq 1$ be the greatest integer such that $e = (\widehat{w}_q 1)^k f$ for some $f \in \{0,1\}^{\mathbb{N}}$ (such a greatest integer exists since if $e = (\widehat{w}_q 1)^\infty$ then it is not true that $e \prec \widehat{w}_q 1e$). Since $\widehat{w}_q 1$ is an even word, this gives

$$(\widehat{w}_q 1)^\infty \preceq f \prec \widehat{w}_q 1f,$$

so $f = \widehat{w}_q 1\dots$, contradicting the definition of k. $\square$

Lemma 9 *Let $s \in \mathcal{MIA}$ have height $q = q(s) > 0$. Then every $t \in \{0,1\}^{\mathbb{N}}$ on the σ-orbit of $(w_q 1)^\infty$ of the form $t = 1^k 0\dots$, where k is odd, satisfies $t \succeq 1\sigma^2(s)$.*

Proof If $s = c_q d$ for some $d \in \{0,1\}^{\mathbb{N}}$, then $d \preceq 1\sigma^2(s) = 1\widehat{w_q}d$ by Lemma 8, so by induction using the fact that $1\widehat{w_q}$ is an even word, $d \preceq (1\widehat{w_q})^n d$ for all $n \geq 0$, and hence $d \preceq (1\widehat{w_q})^\infty$ and $1\sigma^2(s) = 1\widehat{w_q}d \preceq (1\widehat{w_q})^\infty$. If $s = (w_q 1)^\infty$, then $1\sigma^2(s) \preceq (1\widehat{w_q})^\infty$ by direct comparison. It therefore suffices to show that each such t satisfies $t \succeq (1\widehat{w_q})^\infty$. Recall that $\mathrm{rhe}(m/n)$ is preperiodic to $\mathrm{lhe}(m/n)$, ie, $(1\widehat{w_q})^\infty$ lies on the σ-orbit of $(w_q 1)^\infty$.

Writing $q = m/n$ and writing κ_i for $\kappa_i(q)$,

$$\begin{aligned}(w_q 1)^\infty &= (10^{\kappa_1}1^20^{\kappa_2}1^2\dots1^20^{\kappa_{m-1}}1^20^{\kappa_m-1}1)^\infty \quad \text{and}\\ (1\widehat{w_q})^\infty &= (10^{\kappa_1-1}1^20^{\kappa_2}1^2\dots1^20^{\kappa_{m-1}}1^20^{\kappa_m}1)^\infty,\end{aligned}$$

and hence it is required to show that

$$(10^{\kappa_m-1}1^2 0^{\kappa_1}1^2 0^{\kappa_2}1^2 \dots 1^2 0^{\kappa_{m-1}}1)^\infty \succeq (10^{\kappa_1-1}1^2 0^{\kappa_2}1^2 \dots 1^2 0^{\kappa_{m-1}}1^2 0^{\kappa_m}1)^\infty,$$

and

$$(10^{\kappa_j}1^2 \dots 1^2 0^{\kappa_m-1}1^2 0^{\kappa_1}1^2 \dots 1^2 0^{\kappa_{j-1}}1)^\infty \succeq (10^{\kappa_1-1}1^2 0^{\kappa_2}1^2 \dots 1^2 0^{\kappa_{m-1}}1^2 0^{\kappa_m}1)^\infty$$

for $1 \le j \le m-1$.

For the former, since $\kappa_m - 1 = \kappa_1 - 1$, it is equivalent to show that

$$(10^{\kappa_1}1^2 0^{\kappa_2}1^2 \dots 1^2 0^{\kappa_{m-1}}1^2 0^{\kappa_m-1}1)^\infty \succeq (10^{\kappa_2}1^2 \dots 1^2 0^{\kappa_{m-1}}1^2 0^{\kappa_m}1^2 0^{\kappa_1-1}1)^\infty,$$

which is immediate since the left hand side is $(w_q 1)^\infty$, a kneading sequence, and the right hand side is a shift of it.

For the latter, Lemma 7 gives that either $10^{\kappa_j}1^2 \dots 1^2 0^{\kappa_m-1}1$ disagrees with $10^{\kappa_1-1}1^2 \dots 1^2 0^{\kappa_{m-j+1}}1$ and is greater than it (which establishes the result), or the two are equal. If they are equal, then removing this even word from the front of each side of the inequality leaves $(w_q 1)^\infty$ on the left hand side and a shift of it on the right hand side, and the result follows. □

Proof of Theorem 6 Recall that f has no homtervals, being the quotient of an MIA unimodal thick interval map, and hence a point $x \in I$ can be specified uniquely by its itinerary $i(x)$.

Let $x \in (a,b) \setminus \{c\}$ have itinerary $i(x) = i_0 i_1 i_2 \dots$. Then $\theta(x_l) = f(x)_l$ if $i_0 = 0$, and $\theta(x_l) = f(x)_u$ if $i_0 = 1$. Similarly, $\theta(x_u) = f(x)_l$ if $x \ge p$ (ie, if $i(x) \succeq 1\sigma^2(s)$), while $x_u \in \gamma$ if $x \prec 1\sigma^2(s)$. Hence the least $n > 0$ such that $\theta^n(\hat{a}) \in \overline{\gamma}$ is the least $n > 0$ such that either $\sigma^n(i(a)) = i(a)$, or $\sigma^{n-k-1}(i(a))$ is of the form $01^k \dots$ for some odd k with $\sigma^n(i(a)) \preceq 1\sigma^2(s)$.

Suppose first that $\mathrm{lhe}(m/n) \prec s \prec \mathrm{rhe}(m/n)$, so in particular $s = c_q d$ for some $d \in \{0,1\}^{\mathbb{N}}$ by Lemma 5. Thus

$$\begin{aligned} i(a) &= 0^{\kappa_1}1^2 0^{\kappa_2}1^2 \dots 1^2 0^{\kappa_m}1d \quad \text{and} \\ 1\sigma^2(s) &= 10^{\kappa_1-1}1^2 0^{\kappa_2}1^2 \dots 1^2 0^{\kappa_m}1d \end{aligned}$$

(where $\kappa_i = \kappa_i(q)$). Since $10^{\kappa_j}1^2 \dots 1^2 0^{\kappa_m}1d \succ 1\sigma^2(s)$ for $2 \le j \le m$ by Lemma 7, it follows that $\theta^i(\hat{a}) \notin \overline{\gamma}$ for $0 < i < n$. On the other hand, $\sigma^n(i(a)) = d \preceq 1\sigma^2(s)$ by Lemma 8, and hence $\theta^n(\hat{a}) \in \overline{\gamma}$ as required.

To show the different cases for $\theta^n(\hat{a})$, it is only necessary to translate the statements on the left hand sides into statements about $i(f^n(a)) = d$, and then convert these to equivalent statements about $s = c_q d$. For instance,

$$\begin{aligned} \theta^n(\hat{a}) \in (\hat{a}, c_u)_\gamma \quad &\iff \quad d \prec i(c) = 1s \\ &\iff \quad s \prec c_q 1 s \\ &\iff \quad s \prec (c_q 1)^\infty = \mathrm{NBT}(m/n). \end{aligned}$$

Cases ii) and v) follow similarly.

Lemma 9 translates directly into the statement that θ has a periodic orbit $\tilde{P}$ above the period n orbit P of f containing the point with itinerary $(w_q 1)^\infty$ (the points of $\tilde{P}$ on the upper half-circle are exactly those with itineraries having $1^k 0$ as an initial word for some odd k). This itinerary is known to correspond to a periodic orbit of rotation type m/n.

Now let x_l be the point of $\tilde{P}$ with

$$i(x) = (0^{\kappa_1} 1^2 0^{\kappa_2} 1^2 \dots 1^2 0^{\kappa_m - 1} 1^2)^\infty,$$

and consider the interval $J_0 = [\hat{a}, x_l)$. Then J_0 contains no fixed points of θ^n: for the leftmost point y_l on the corresponding periodic orbit would have $i(a) \prec i(y) \prec i(x)$ and hence $i(y) = 0^{\kappa_1} 1^2 \dots 1^2 0^{\kappa_m - 1} {}^{0}_{1} 1 \dots$, so either $y = x$ or $i(y) = (0^{\kappa_1} 1^2 \dots 1^2 0^{\kappa_m - 1} 0 1)^\infty$, which contains an isolated 1 and therefore does not correspond to an orbit of θ. So θ^n is continuous on J_0, fixes x_l, maps $\hat{a}$ into γ, and has no fixed point in J_0, and hence every point of J_0 falls into γ under iteration of θ^n, ie, $\Lambda \cap J_0 = \emptyset$.

A similar argument shows that if y_u is the point of $\tilde{P}$ with

$$i(y) = (1 0^{\kappa_1 - 1} 1^2 0^{\kappa_2} 1^2 \dots 1^2 0^{\kappa_m} 1)^\infty$$

and $J_1 = [p_u, y_u)$, then $\Lambda \cap J_1 = \emptyset$. Thus $J = \theta(J_0) \cup \theta(J_1) = (f(x)_l, f(y)_l)$ is disjoint from Λ. In particular, the endpoints of J are consecutive points of $\tilde{P}$, and so $\bigcup_{i=0}^{n-1} \theta^i(J) = S^1 \setminus \tilde{P}$, establishing that $\Lambda = \tilde{P}$ as required.

The two special cases $s = \mathrm{lhe}(m/n)$ and $s = \mathrm{rhe}(m/n)$ can be treated similarly, but more straightforwardly since explicit expressions for s are available. The reason that $\Lambda \neq \tilde{P}$ in these cases is that $\tilde{P}$ contains the unique point $\theta(\hat{a})$ of S^1 that has two θ-preimages, $\hat{a}$ and p_u. ◻

6 Invariant unimodal generalized train tracks

In this section, the invariant generalized train tracks corresponding to elements of $\mathcal{MIA}$ are described explicitly. This is achieved by a relatively straightforward analysis of the construction of invariant generalized train tracks given in the proof of Theorem 1: the important point for what follows is the way in which the structure of the train track is governed by the height of the kneading sequence.

The case of preperiodic kneading sequences is somewhat more complicated than the periodic case, and as such the two are treated separately.

6.1 The periodic case

Let $s = \kappa(f) \in \mathcal{MIA}$ be the kneading sequence of a unimodal map f whose critical point c is periodic of period N, and let $F\colon (S^2, \mathbb{I}, A) \circlearrowleft$ be the associated thick interval map, where A is the set of attracting periodic points of F, consisting of a single periodic orbit whose points are in natural correspondence with the points of the f-orbit of c (see Figure 4 for an example in the case $s = (1001011)^\infty$). Let $\tau \subseteq \mathbb{I} \setminus A$ be the F-invariant generalized train track: the aim of this section is to describe τ explicitly. Label the junctions of $\mathbb{I}$ with integers 1 to N from left to right: thus junctions 1 and N are 1-junctions, while junctions 2 to $N-1$ are 2-junctions.

Because of the convention adopted in Section 4.1 for the itinerary of a point whose orbit contains c, the kneading sequence is given by $s = w^\infty$, where w is a word of length N whose final symbol is 1.

It will be seen (Theorem 10 below) that only four different basic configurations of infinitesimal edges can occur in any given junction of $\mathbb{I}$. These are as follows: note that the position of the puncture in the junction relative to the infinitesimal edges is also specified (indicated by a small circle in the figures).

BP A single bubble containing the puncture.

W A configuration of infinitely many loops of which infinitely many are bubbles (all homotopically trivial), and infinitely many are not, as depicted in Figure 9. There are two different versions, W^+ and W^- which are mirror images of each other.

S^+ A semi-infinite bouquet of homotopically trivial bubbles as depicted in Figure 10.

V_3 A single bigon containing W. There are two versions, V_3^+ and V_3^-, containing W^+ and W^- respectively (see Figure 11: V_3^- is the mirror image of V_3^+).

The infinitesimal edges of the train tracks described in this section are specified as follows. For each junction of $\mathbb{I}$, there is given one of the above symbols

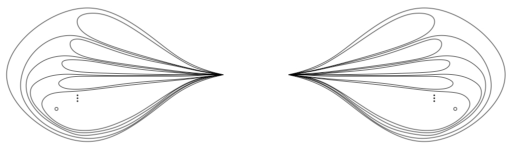


Figure 9: W^+ and W^-

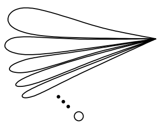


Figure 10: S^+

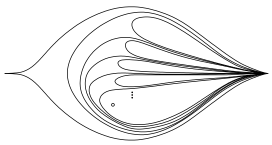


Figure 11: V_3^+

describing the basic configuration of infinitesimal edges. The junction is said to be *type* BP, *type* W^+, *type* W^-, *type* S^+, *type* V_3^+ or *type* V_3^- correspondingly. For each 2-junction, there is in addition a symbol L or R, specifying whether the loop(s) in the basic configuration are attached to the left or right switch in that junction. The 2-junction is said to be *type* L or *type* R correspondingly.

In each 2-junction not of type $V_3^{\pm}$, there is one additional infinitesimal edge, whose endpoints are the two switches on the boundary of the junction. This edge passes above (respectively below) all other infinitesimal edges in the junction if the junction is type R (respectively type L).

The reader seeking clarification as to how this notation is used can look ahead to Examples 8 and the accompanying figures.

There are two special cases: the *finite order* case $s = \mathrm{lhe}(m/n)$ (where each junction is of type S^+), and the *NBT* case $s = \mathrm{NBT}(m/n)$ (where each junction is of type BP, and hence τ is finite). In all other cases, the invariant generalized train track contains junctions of types $W^{\pm}$ and $V_3^{\pm}$.

In all cases, it can be determined which 2-junctions are type L and which are type R by applying the following simple algorithm. The idea is straightforward:

the infinitesimal edges in junction 1 are clearly attached to the right hand edge of that junction. This information is propagated around the orbit, with the type changing after each symbol 1 (corresponding to a 'flip'), and remaining unchanged after each symbol 0.

Algorithm 1 Let $s \in \mathcal{MIA}$ be a period N kneading sequence, and let $\pi = \pi_s \in S_N$ denote the induced permutation on the points of the critical orbit of a corresponding unimodal map. Then the partition $\{2, \dots, N-1\} = L \cup R$ is determined inductively as follows: $\pi(1) \in R$, and for each r with $2 \le r \le N-2$, $\pi^r(1)$ lies in the same set as $\pi^{r-1}(1)$ if $\pi^{r-1}(1) < \pi^{-1}(N)$, and in the other set if $\pi^{r-1}(1) > \pi^{-1}(N)$.

Example 7 The easiest way to determine these sets is to write down the first N symbols of s, and to place Ls and Rs above them. The LR sequence starts with an R at the third symbol, and changes after each 1.

For example, let $s = (10011001011)^\infty$, so that $\pi = (1\,3\,7\,10\,2\,5\,9\,4\,8\,6\,11)$. Then Algorithm 1 gives

$$\begin{array}{ccccccccccc} 11 & 1 & 3 & 7 & 10 & 2 & 5 & 9 & 4 & 8 & 6 \\ & & R & R & L & R & R & R & L & L & R \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1. \end{array}$$

Thus $L = \{4, 8, 10\}$ and $R = \{2, 3, 5, 6, 7, 9\}$.

The next result describes the invariant generalized train track corresponding to any periodic kneading sequence in $\mathcal{MIA}$.

Theorem 10 *Let $s \in \mathcal{MIA}$ be a periodic kneading sequence of period N with height $m/n > 0$. Then the invariant generalized train track corresponding to s has infinitesimal edges as follows. Junction i is of type L or R according as $i \in L$ or $i \in R$. The basic configurations of infinitesimal edges are:*

a) *If $s = \mathrm{lhe}(m/n)$, then all junctions are of type S^+.*
b) *If $s = \mathrm{NBT}(m/n)$, then all junctions are of type BP.*
c) *Otherwise, let $\pi \in S_N$ denote the induced permutation on the points of the periodic critical orbit of a unimodal map with kneading sequence s, and let $\epsilon \in \{+, -\}$ be given by $\epsilon = +$ if $\pi^{-1}(N) \in R$, and $\epsilon = -$ if $\pi^{-1}(N) \in L$. Then for $0 \le r \le N-1$, the junction $\pi^r(N)$ is of type W^ϵ if $r \le n+1$, and of type V_3^ϵ if $r > n+1$.*

The proof is delayed until some explanatory examples have been presented.

Examples 8 The three examples correspond to the three cases of Theorem 10.

a) Let $s = (101)^\infty = \mathrm{lhe}(1/3)$, so that $\pi = (1\,2\,3)$. Algorithm 1 gives

$$\begin{array}{ccc} 3 & 1 & 2 \\ & & R \\ 1 & 0 & 1 \end{array}$$

Thus the invariant generalized train track is described by

$$(S^+ \;;\; S^+, R \;;\; S^+)$$

(see Figure 12: note that the infinitesimal edge with endpoints the left and right switches in junction 2 passes above the other infinitesimal edges in the junction, since the junction is of type R. In this and subsequent figures, the punctures have been labelled to clarify the dynamics.)

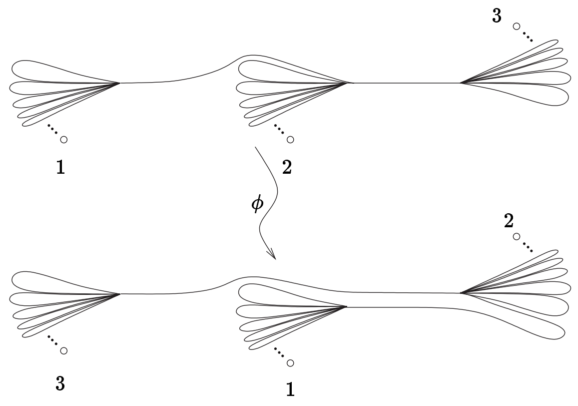


Figure 12: The generalized train track corresponding to $s = \mathrm{lhe}(1/3)$ and its image

b) Let $s = (10011)^\infty = \mathrm{NBT}(1/3)$, so $\pi = (1\,2\,4\,3\,5)$. Algorithm 1 gives

$$\begin{array}{ccccc} 5 & 1 & 2 & 4 & 3 \\ & & R & R & L \\ 1 & 0 & 0 & 1 & 1 \end{array}$$

Thus the invariant generalized train track is described by

$$(\mathrm{BP} \;;\; \mathrm{BP}, R \;;\; \mathrm{BP}, L \;;\; \mathrm{BP}, R \;;\; \mathrm{BP})$$

(see Figure 13).

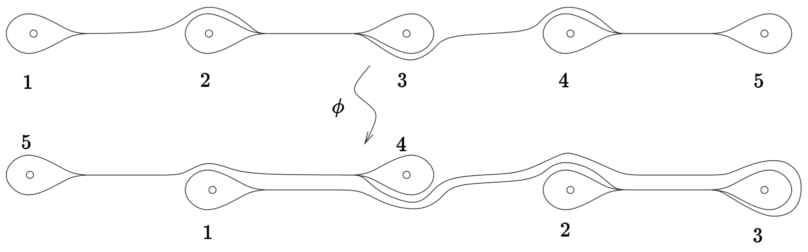


Figure 13: The generalized train track corresponding to $s = \mathrm{NBT}(1/3)$ and its image

c) Let $s = (1001011)^\infty$, so $q(s) = 1/3$ and $\pi = (1\,3\,6\,2\,5\,4\,7)$. Algorithm 1 gives

$$\begin{array}{ccccccc} 7 & 1 & 3 & 6 & 2 & 5 & 4 \\ & & R & R & L & L & R \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{array}$$

Thus $\epsilon = +$, since $\pi^{-1}(7) = 4 \in R$. The junctions $\pi^r(7)$ with $0 \le r \le 4$ are 7, 1, 3, 6, and 2, so these junctions are of type W^+, while junctions 5 and 4 are of type V_3^+. Thus the invariant generalized train track is described by

$$(W^+ \;;\; W^+, L \;;\; W^+, R \;;\; V_3^+, R \;;\; V_3^+, L \;;\; W^+, R \;;\; W^+)$$

(see Figure 14).

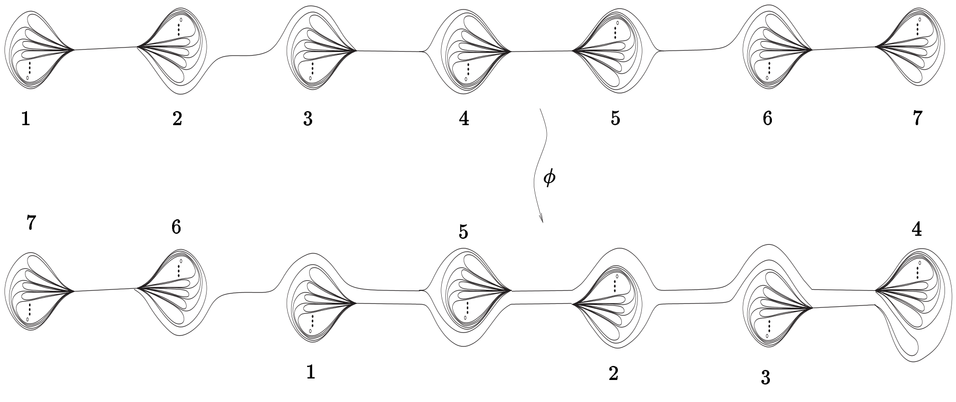


Figure 14: The generalized train track corresponding to $s = (1001011)^\infty$ and its image

Proof of Theorem 10 Let τ be the invariant generalized train track, as constructed using the algorithm of Theorem 1. Let $A \subseteq \{2, \ldots, N-1\}$ (respectively

$B \subseteq \{2, \dots, N-1\}$) be the set of 2-junctions which contain an infinitesimal edge of τ joining the two switches in the junction and passing above (respectively below) the puncture. There are two ways such edges can arise during the construction. First, as images of real edges: this requires that $r \in A$ (respectively $r \in B$) if $r \geq \pi(1)$ and $\pi^{-1}(r) < \pi^{-1}(N)$ (respectively $r \geq \pi(1)$ and $\pi^{-1}(r) > \pi^{-1}(N)$). Second, as images of other such infinitesimal edges: this requires that if $\pi^{-1}(r) < \pi^{-1}(N)$ then r lies in A (respectively B) if $\pi^{-1}(r)$ lies in A (respectively B); while if $\pi^{-1}(r) > \pi^{-1}(N)$ then r lies in B (respectively A) if $\pi^{-1}(r)$ lies in A (respectively B).

To express this symbolically, partition $\{1, \dots, N\} = O \cup I$, with $r \in O$ if $r < \pi^{-1}(N)$ (ie, if r corresponds to the symbol 0), and $r \in I$ otherwise. Then A and B are the smallest subsets of $\{2, \dots, N-1\}$ such that

i) If $r \geq \pi(1)$, then $r \in A$ if $\pi^{-1}(r) \in O$, and $r \in B$ if $\pi^{-1}(r) \in I$.

ii) If $\pi^{-1}(r) \in O$ then $r \in A$ if $\pi^{-1}(r) \in A$, and $r \in B$ if $\pi^{-1}(r) \in B$. If $\pi^{-1}(r) \in I$ then $r \in B$ if $\pi^{-1}(r) \in A$, and $r \in A$ if $\pi^{-1}(r) \in B$.

By condition ii), $A \cap B$ is a final segment of the sequence $(\pi^j(1))_{j=1}^{N-2}$. By conditions i) and ii), the least j such that $\pi^j(1) \in A \cap B$ is the least j with $\pi^j(1) > \pi(1)$ and $\pi^{j-1}(1) \in B$. The main step of the proof is to establish that this least j is equal to $n+1$. In particular, by Lemma 5, $A \cap B = \emptyset$ if and only if either $s = \mathrm{lhe}(m/n)$ or $s = \mathrm{NBT}(m/n)$.

By Lemma 5, either $s = \mathrm{lhe}(m/n)$ or $s = c_{m/n} d$ for some $d \in \{0,1\}^{\mathbb{N}}$. Assume that the latter holds: the proof that $A \cap B = \emptyset$ in the former case is similar. Now since $1 \in O$, condition i) gives $\pi(1) \in A$, and the assignments to A and B given by condition ii) are as follows:

$$s = \begin{array}{ccccccccccccc} & A & A & B & A & A & B & & A & B & A & A & \\ 1 & 0^{\kappa_1} & 1 & 1 & 0^{\kappa_2} & 1 & 1 & \dots & 1 & 1 & 0^{\kappa_m} & 1 & d \end{array}$$

Using $i(\pi(1)) = 0^{\kappa_1 - 1} 1^2 0^{\kappa_2} 1^2 \dots 1^2 0^{\kappa_m} 1 d = \sigma^2(s)$ and applying Lemma 7 gives $j \geq n+1$. Since the first symbol of d corresponds to an element of B, it only remains to show that $\sigma(d) \succeq \sigma^2(s)$. This is given by Lemma 8.

Notice that for $r \notin A \cap B$, comparing condition ii) with Algorithm 1 gives that $r \in A$ if and only if $r \in R$.

Now consider how loops can arise during the construction of τ. Again, there are two ways. First, as the image of an infinitesimal edge joining the two switches of junction $\pi^{-1}(N)$: this gives rise to a homotopically trivial bubble in junction N (respectively a loop containing the puncture in junction N) if the edge passes above (respectively below) the puncture in junction $\pi^{-1}(N)$. Second, as the image of another loop.

If $s = \mathrm{NBT}(m/n)$, ie,

$$s = \begin{array}{cccccccccccccc} & A & A & B & A & A & B & & A & B & A & A & B & \\ 1 & 0^{\kappa_1} & 1 & 1 & 0^{\kappa_2} & 1 & 1 & \dots & 1 & 1 & 0^{\kappa_m} & 1 & 1 & \dots, \end{array}$$

then any infinitesimal edge joining the switches of junction $\pi^{-1}(N)$ passes below the puncture, and hence gives rise to a loop in junction N which contains the puncture. Statement b) follows easily. Similarly, if $s = \mathrm{lhe}(m/n)$, ie,

$$s = \begin{array}{ccccccccccccc} & A & A & B & A & A & B & & A & B & A & A & \\ 1 & 0^{\kappa_1} & 1 & 1 & 0^{\kappa_2} & 1 & 1 & \dots & 1 & 1 & 0^{\kappa_m - 1} & 1 & \dots, \end{array}$$

then any infinitesimal edge joining the switches of junction $\pi^{-1}(N)$ passes above the puncture, and hence gives rise to a homotopically trivial bubble in junction N. Statement a) follows.

In any other case, $\pi^{-1}(N) \in A \cap B$, so that both homotopically trivial bubbles and homotopically non-trivial loops are created in junction N, leading under iteration of the construction to a configuration of infinitesimal edges of type W in each junction. The orientation of these infinitesimal edges bubbles depends on whether $\pi^{-1}(N) \in R$ or $\pi^{-1}(N) \in L$ (which determines whether a new homotopically trivial bubble arises below or above the existing loops in junction N during the construction). Junction r is of type $V_3^{\pm}$ if $r \in A \cap B$, but of type $W^{\pm}$ if r is in only one of A and B. Statement c) follows. □

6.2 The preperiodic case

The preperiodic case is somewhat more complicated than the periodic case, though there are no new ideas introduced: as such, the treatment is more informal and some of the details are omitted. The most substantial modifications arising because the critical point is not periodic are:

a) There is a unique loop which arises during the construction as the image of an infinitesimal edge which is not a loop, and this loop is a homotopically trivial bubble. All other loops are images of this one, and hence all loops are homotopically trivial bubbles.

b) Only finitely many infinitesimal edges occur in those junctions corresponding to the strictly preperiodic part of the orbit of the critical point.

c) The parity of the periodic part of the kneading sequence plays an important role. If there are an odd number of 1s, then there is a flip each time around the periodic orbit, and hence bubbles are attached to both the left and right switches of 2-junctions.

d) In the periodic case, the configuration of infinitesimal edges changes (from W to V_3) at a point in the orbit given by the denominator of the height. In the preperiodic case, separate cases arise according as this point lies in the preperiodic or periodic part of the orbit.

The basic configurations of infinitesimal edges are entirely different from the periodic case. There are two distinct types: those which arise in junctions corresponding to the preperiodic part of the orbit (B and V_0), which contain only finitely many infinitesimal edges and have no associated puncture; and those which arise in periodic junctions, which contain infinitely many infinitesimal edges and have an associated puncture. The configurations are:

B A single (homotopically trivial) bubble.

V_0 A bigon containing a single bubble, as depicted in Figure 15.

S^- The mirror image of S^+.

V_1 A decorated bigon as depicted in Figure 16. There are two versions, V_1^+ and V_1^-, which are mirror images of each other. In addition, an additional external bubble may be present in the position indicated in Figure 18. In this case, the configuration is denoted $V_1^\pm B$.

V_2 A decorated bigon as depicted in Figure 17. There are two versions, V_2^+ and V_2^-: V_2^+ is the mirror image of V_2^-. In addition, there may be two additional external bubbles in the positions indicated in Figure 18. In this case, the configuration is denoted $V_2^\pm B^2$.

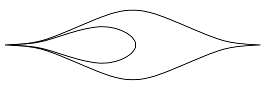
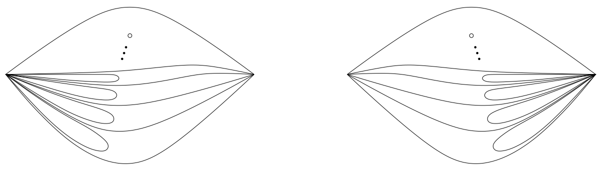


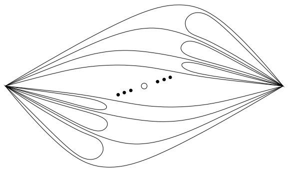
Figure 15: V_0

For each junction of $\mathbb{I}$, one of the above symbols is given describing the basic configuration of infinitesimal edges: the junction is said to be *type* V_0, *type* V_1^-, *type* $V_2^+ B^2$, etc. Each 2-junction which is not of type V_2 is either *type* L or *type* R, according as the bubbles are attached to the left or right switch of the junction. For each 2-junction of type B or S^-, there is one additional



V_1^- V_1^+

Figure 16: V_1^- and V_1^+



V_2^-

Figure 17: V_2^-

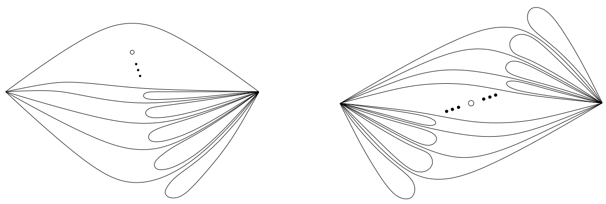


Figure 18: Extra bubbles: the configurations $V_1^+ B$ and $V_2^- B^2$

infinitesimal edge, whose endpoints are the two switches on the boundary of the junction. This edge passes above (respectively below) all other infinitesimal edges in the junction if the junction is type R (respectively type L).

Throughout this subsection s denotes a strictly preperiodic kneading sequence, written $s = vw^\infty$, where v is taken to be as short a word as possible (so the final symbol of v is not equal to the final symbol of w). The lengths of v and w will be denoted k and l respectively (so $k, l \geq 1$), and the size of the orbit of s under the shift map is denoted $N = k+l$. The N points of this orbit are labelled with the integers 1 through N according to their relative unimodal ordering (so point r lies in junction r). The shift map σ thus induces a map $\rho\colon \{1, \dots, N\} \circlearrowleft$. Note that ρ is not a permutation: the point N (corresponding to the first symbol of s) has no preimage, and one point (corresponding to the first symbol of w) has two preimages.

The following algorithm is the analogue of Algorithm 1 for the preperiodic case. Notice that $L \cup R$ need not be a partition of $\{2, \ldots, N-1\}$ (reflecting the possibility that the word w may be odd): if $r \in L \cap R$, then bubbles are attached to both switches in junction r (configuration V_2).

Algorithm 2 The subsets L and R of $\{2, \ldots, N-1\}$ are determined inductively as follows. Let $\rho^2(N) \in R$, and for each r with $3 \le r \le N+l-1$, let $\rho^r(N)$ lie in the same set as $\rho^{r-1}(N)$ if $s_{r-1} = 0$, and in the other set if $s_{r-1} = 1$.

If w is an even word, then the final l steps in this algorithm are unnecessary, and $L \cap R = \emptyset$.

Examples 9 a) Let $s = 1000(101)^\infty$. Then $v = 1000$, $w = 101$, and

$$\rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 4 & 5 & 6 & 6 & 3 & 1 \end{pmatrix}$$

(ie, the ρ-orbit of 7 is $7124(635)^\infty$). Thus

$$\begin{array}{cccccccc} 7 & 1 & 2 & 4 & & 6 & 3 & 5 \\ & & R & R & & R & L & L \\ 1 & 0 & 0 & 0 & & 1 & 0 & 1 \end{array},$$

and so $L = \{3,5\}$ and $R = \{2,4,6\}$.

b) Let $s = 100101(10)^\infty$. Then $v = 100101$, $w = 10$, and

$$\rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 5 & 6 & 7 & 6 & 3 & 2 & 1 \end{pmatrix}$$

(ie, the ρ-orbit of 8 is $814725(63)^\infty$). Thus

$$\begin{array}{cccccccccccc} 8 & 1 & 4 & 7 & 2 & 5 & & 6 & 3 & & 6 & 3 \\ & & R & R & L & L & & R & L & & L & R \\ 1 & 0 & 0 & 1 & 0 & 1 & & 1 & 0 & & 1 & 0 \end{array},$$

and so $L = \{2,3,5,6\}$ and $R = \{3,4,6,7\}$. Note that since w is odd, the two repeating points 3 and 6 of the orbit lie in both L and R.

For preperiodic kneading sequences, there is only one special case: that in which $s = \mathrm{rhe}(m/n)$.

Theorem 11 *Let $s = vw^\infty \in \mathcal{MIA}$ be a strictly preperiodic kneading sequence of height $m/n > 0$. Let k and l be the lengths of v and w (chosen as small as possible). Let ρ be the induced map on $\{1, \dots, N = k + l\}$. Then the invariant generalized train track corresponding to s has infinitesimal edges as follows. If $i \in L \setminus R$ (respectively $i \in R \setminus L$) then junction i is of type L (respectively R). The basic configurations of infinitesimal edges are:*

a) *If $s = \mathrm{rhe}(m/n)$, then the junction $\rho^r(N)$ is of type B for $0 \le r < k$, and of type S^- for $k \le r < N$.*

b) *Otherwise, let $\epsilon \in \{+, -\}$ be given by $\epsilon = +$ if $\rho^{k-1}(N) \in R$, and $\epsilon = -$ if $\rho^{k-1}(N) \in L$. Then for each r with $0 \le r < N$:*

   i) *If $r \le n + 1$, then junction $\rho^r(N)$ is of type*
      - *B, if $r \le k$.*
      - *$V_1^\epsilon B$ or $V_2^\epsilon B^2$ (according as w is even or odd), if $r > k$.*

   ii) *If $r > n + 1$, then junction $\rho^r(N)$ is of type*
      - *V_0, if $r \le k$.*
      - *V_1^ϵ or V_2^ϵ (according as w is even or odd), if $r > k$.*

Examples 10 The first example illustrates the case $s = \mathrm{rhe}(m/n)$. The other two examples show the two possibilities w even and w odd.

a) Let $q = 1/3$ and $s = \mathrm{rhe}(q)$ so $s = 1001(101)^\infty = 10(011)^\infty$. Thus $v = 10$, $w = 011$, and

$$\rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 2 & 1 \end{pmatrix}$$

(ie, the ρ-orbit of 5 is $51(234)^\infty$). Thus

$$\begin{array}{cccccc} 5 & 1 & \quad 2 & 3 & 4 & \\ & & \quad R & R & L & , \\ 1 & 0 & \quad 0 & 1 & 1 & \end{array}$$

and so $R = \{2, 3\}$ and $L = \{4\}$.
Since $k = 2$ and $l = 3$, junctions $N = 5$ and $\rho(N) = 1$ are of type B, while the other junctions are of type S^-.
Thus the invariant generalized train track is described by

$$(B \; ; \; S^-, R \; ; \; S^-, R \; ; \; S^-, L \; ; \; B)$$

(see figure 19: here all junctions have been labelled to clarify the dynamics).

b) Let $s = \overset{7}{1}\overset{1}{0}\overset{2}{0}\overset{4}{0}(\overset{6}{1}\overset{3}{0}\overset{5}{1})^\infty$. Then $L = \{3,5\}$ and $R = \{2,4,6\}$ (see Example 9 a)).
Since $k = 4$ and $l = 3$, $\rho^{k-1}(N) = \rho^3(7) = 4 \in R$, and hence $\epsilon = +$. s has height $q(s) = 1/4$, so $n = 4$. Thus junctions $N = 7$, $\rho(N) = 1$, $\rho^2(N) = 2$, and $\rho^3(N) = 4$ are of type B; junctions $\rho^4(N) = 6$ and $\rho^5(N) = 3$ are of type $V_1^+ B$; and junction $\rho^6(N) = 5$ is of type V_1^+. Thus the invariant generalized train track is described by

$$(B \;;\; B,R \;;\; V_1^+ B,L \;;\; B,R \;;\; V_1^+,L \;;\; V_1^+ B,R \;;\; B)$$

(see Figure 20).

c) Let $s = \overset{8}{1}\overset{1}{0}\overset{4}{0}\overset{7}{1}\overset{2}{0}\overset{5}{1}(\overset{6}{1}\overset{3}{0})^\infty$. Then $L = \{2,3,5,6\}$ and $R = \{3,4,6,7\}$ (see Example 9 b)).
Since $k = 6$ and $l = 2$, $\rho^{k-1}(N) = \rho^5(8) = 5 \in L$, and hence $\epsilon = -$. s has height $q(s) = 1/3$, so $n = 3$. Thus junctions $N = 8$, $\rho(N) = 1$, $\rho^2(N) = 4$, $\rho^3(N) = 7$, and $\rho^4(N) = 2$ are of type B; junction $\rho^5(N) = 5$ is of type V_0; and junctions $\rho^6(N) = 6$ and $\rho^7(N) = 3$ are of type V_2^-. Thus the invariant generalized train track is described by

$$(B \;;\; B,L \;;\; V_2^- \;;\; B,R \;;\; V_0,L \;;\; V_2^- \;;\; B,R \;;\; B)$$

(see Figure 21).

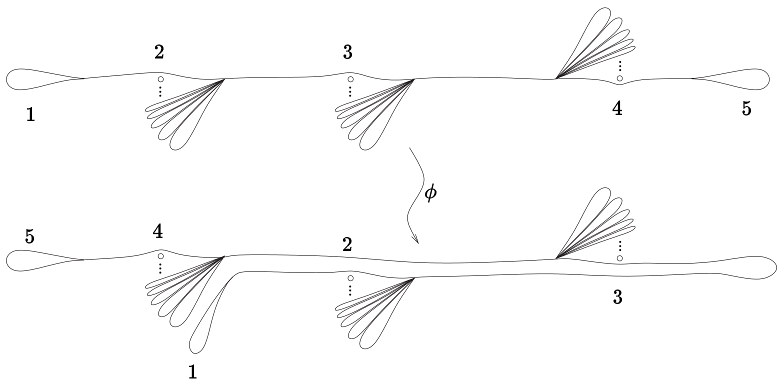


Figure 19: The generalized train track corresponding to $s = \mathrm{rhe}(1/3)$ and its image

Sketch proof of Theorem 11 The proof is very similar to that of Theorem 10. Defining sets A and B as there, and noting that $s = c_{m/n}d$ for some $d \in \{0,1\}^{\mathbb{N}}$, the same analysis shows that $A \cap B = \emptyset$ if and only if $s = \mathrm{rhe}(m/n)$,

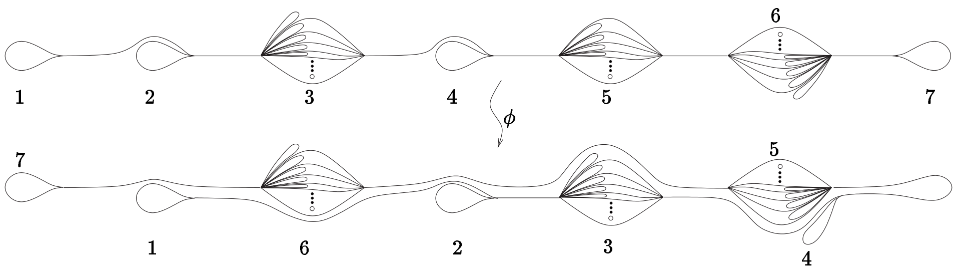


Figure 20: The generalized train track corresponding to $s = 1000(101)^\infty$ and its image

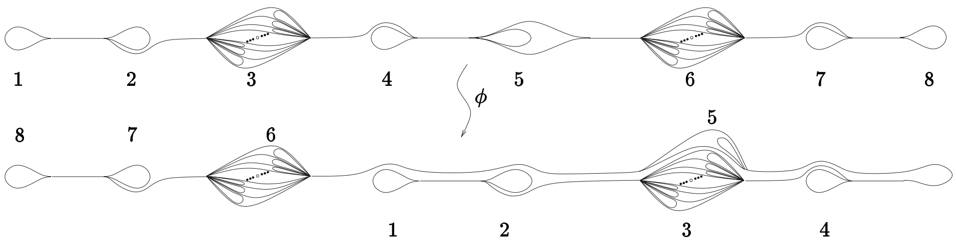


Figure 21: The generalized train track corresponding to $s = 100101(10)^\infty$ and its image

and in all other cases $A \cap B = \{\rho^j(N) \,:\, n+2 \le j \le N-1\}$. This is again the essential part of the proof.

When $s = \mathrm{rhe}(m/n)$, it is then straightforward to check case a). In other cases, it is immediate that for $r \le k$ (ie, in the preperiodic part of the orbit) r has type V_0 if $r \in A \cap B$, and type B otherwise.

All of the bubbles arise in the construction as images of the bubble in junction N (which itself is created as the image of part of a real edge). In particular, they are all homotopically trivial. The combinatorics of the different cases can be checked routinely. $\square$

7 Unimodal generalized pseudo-Anosov maps

In this section the construction of generalized pseudo-Anosov maps starting from the invariant train tracks of Section 6 is described.

7.1 Preliminaries

Let $F\colon (S^2, \mathbb{I}, A) \circlearrowleft$ be a unimodal MIA thick interval map, where A is the set of attracting periodic points of F. Let $f\colon I \circlearrowleft$ be the quotient unimodal map, and $\phi\colon \tau \circlearrowleft$ be the associated generalized train track map. Let $e_1, e_2, \dots, e_{N-1}$ denote the real edges of τ, and $e_N, e_{N+1}, \dots$ denote the infinitesimal edges. Define the transition matrix $M = (m_{ij})$ by

$$m_{ij} = \text{number of times } \phi(e_j) \text{ crosses } e_i.$$

Thus M is an infinite matrix, unless $\kappa(F) = \mathrm{NBT}(m/n)$ for some $m/n \in (0, 1/2)$.

From the construction of the invariant generalized train track in Section 6, it is immediate that each $\phi(e_j)$ is a finite length edge path; that each e_i is contained in the edge path $\phi(e_j)$ for only finitely many j; and that each infinitesimal edge is mapped by ϕ homeomorphically onto an infinitesimal edge. These observations yield the following facts about the strucure of the transition matrix:

a) M is of the form
$$M = \begin{pmatrix} A & 0 \\ B & \Pi \end{pmatrix},$$
where A is an $(N-1)\times(N-1)$ matrix recording transitions between real edges, Π records transitions between infinitesimal edges, and B records transitions from real to infinitesimal edges.

b) All entries of Π are either 0 or 1. Each column of Π has exactly one non-zero entry, and each row has only finitely many non-zero entries.

c) B has only finitely many non-zero entries in each column.

In particular, if M is regarded as an operator acting on l^1, then M is bounded with $\|M\|_1 \le \max_j\{\sum_i |m_{ij}|\} < \infty$. Let λ be the Perron-Frobenius eigenvalue of A, and $Y \in \mathbb{R}^{N-1}$ the associated positive eigenvector. Then Y can be extended to a positive eigenvector $y = (Y\,Y') \in l^1$ of M, called the Perron-Frobenius eigenvector of M, by setting

$$Y' = \frac{1}{\lambda}(B + \frac{1}{\lambda}\Pi B + \frac{1}{\lambda^2}\Pi^2 B + \dots)Y.$$

Note that $\|\Pi\|_1 \le 1$ so that the series above converges in l^1, implying that $y = (Y\,Y') \in l^1$. Note also that the matrices $\Pi^k B$ represent transitions from a real edge to an infinitesimal edge under ϕ, and then to another infinitesimal edge under a further k iterations of ϕ. By the construction of τ described in the proof of Theorem 1, every infinitesimal edge of τ is the image, under some iterate of ϕ, of some infinitesimal edge of $F_*(\tau(\emptyset))$, and each of these infinitesimal edges is crossed by $\phi(e)$ for some real edge e. Thus for all $i \geq 1$ there exists $k \geq 1$ such that row i of $\Pi^k B$ is non-zero. Since Y is a strictly positive vector, this establishes that every entry of Y' is also positive.

Definition 15 A sequence (y_i) of positive real numbers is said to satisfy the *switch conditions for* τ if, for each switch q of τ,

$$y_{i_0} = \sum y_i + 2 \sum y_j,$$

where e_{i_0} is the real edge with endpoint q, and the first and second sums range over the sets of indices of infinitesimal edges having one and both endpoints at q respectively.

The proof of the following lemma is essentially the same as that of the corresponding fact for finite matrices in Section 3.4 of [2].

Lemma 12 *Let the transition matrix M of $\phi\colon \tau \circlearrowleft$ have Perron-Frobenius eigenvector $y = (y_1\, y_2\, \ldots)$. Then y satisfies the switch conditions for τ.*

Proof For $r, s, k \geq 1$ denote by $m_{rs}^{(k)}$ the number of times that $\phi^k(e_s)$ crosses e_r. Thus $m_{rs}^{(k)} = (M^k)_{rs}$. The fact that y is an eigenvector of M^k with eigenvalue λ^k thus gives that for any k and any r,

$$y_r = \frac{1}{\lambda^k} \sum_{s=1}^{\infty} m_{rs}^{(k)} y_s.$$

Now if q is a switch of τ which is an endpoint of the real edge e_{i_0}, and if I and J are the sets of indices of infinitesimal edges having one and two endpoints at q respectively, then for any s,

$$\left(\sum_{i \in I} m_{is}^{(k)} + 2 \sum_{j \in J} m_{js}^{(k)} \right) - 2 \leq m_{i_0 s}^{(k)} \leq \left(\sum_{i \in I} m_{is}^{(k)} + 2 \sum_{j \in J} m_{js}^{(k)} \right) + 2.$$

This is because, except at its endpoints, $\phi^k(e_s)$ must cross both e_{i_0} and an edge indexed in $I \cup J$ each time it intersects q. Multiplying by y_s, summing over s, and dividing by λ^k gives

$$\left| \sum_{i \in I} y_i + 2 \sum_{j \in J} y_j - y_{i_0} \right| \leq \frac{2 \sum_{s=1}^{\infty} y_s}{\lambda^k},$$

which gives the result on letting $k \to \infty$. □

Now let $X \in \mathbb{R}^{N-1}$ be an eigenvector of A^T associated to the Perron-Frobenius eigenvalue λ. If $x = (X\, X') \in l^\infty$ is a completion of X to an eigenvector of the adjoint M^* then $X' = 0$, since it satisfies the equation $(\lambda I - \Pi^*)X' = 0$, and $\|\Pi^*\|_\infty \leq 1 < \lambda$.

7.2 The construction

The construction will be illustrated by means of a running example, where $F\colon (S^2, \mathbb{I}, A) \circlearrowleft$ is a unimodal MIA thick interval map with $\kappa(F) = (1001011)^\infty$, as depicted in Figure 4. The associated generalized train track map $\phi\colon \tau \circlearrowleft$ is depicted in Figure 14.

Let $x = (x_1, x_2, \ldots, x_{N-1}, 0, 0, \ldots)$ and $y = (y_1, y_2, \ldots)$ be the eigenvectors of M^* and M associated to the Perron-Frobenius eigenvalue λ of A as described above. For definiteness these vectors are scaled so that $\|x\|_2 = 1$ and $\sum x_i y_i = 1$, although this normalization can be chosen arbitrarily. In the running example, to 3 decimal places:

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}, \quad \lambda = 1.686, \quad X = \begin{pmatrix} 0.543 \\ 0.104 \\ 0.191 \\ 0.724 \\ 0.175 \\ 0.322 \end{pmatrix}, \quad Y = \begin{pmatrix} 0.368 \\ 0.291 \\ 0.509 \\ 0.536 \\ 0.490 \\ 0.620 \end{pmatrix}$$

Associate a Euclidean rectangle R_i of width x_i and height y_i, foliated by horizontal and vertical line segments, to each real edge e_i of τ. The idea of the construction is as follows: a topological sphere S is constructed by identifying the vertical sides of these rectangles in the manner dictated by the infinitesimal edges of τ and then identifying the horizontal sides in such a way that the train track map ϕ defines a homeomorphism $\Phi\colon S \circlearrowleft$. The foliations of the rectangles project to singular foliations of S with transverse measures induced by the Euclidean metric: these foliations are preserved by Φ, the horizontal and vertical foliations being stretched and contracted by factors λ and $1/\lambda$ respectively, so that Φ is a generalized pseudo-Anosov map.

When τ is finite (ie, when $\kappa(F) = \mathrm{NBT}(m/n)$ for some m/n), this construction is identical to that of [2], and produces a pseudo-Anosov homeomorphism. When τ is infinite, however, the construction involves infinitely many sets of identifications, and it is therefore necessary to describe them carefully and to justify that the quotient space is a sphere. The tool used to do this is Moore's theorem [16] on monotone upper semi-continuous (m.u.s.c.) decompositions of the sphere.

Definitions 16 A partition or *decomposition* $\mathcal{G}$ of a topological space into compact sets is *upper semi-continuous* if for every decomposition element $\zeta \in \mathcal{G}$ and every open set $U \supset \zeta$, there exists an open set $V \subset U$ with $\zeta \subset V$ such that every $\zeta' \in \mathcal{G}$ with $\zeta' \cap V \neq \emptyset$ has $\zeta' \subset U$. The decomposition is *monotone* if its elements are connected.

The proof of the following lemma is routine.

Lemma 13 *A decomposition $\mathcal{G}$ of a compact metric space X is upper semi-continuous if and only if whenever $x_n \to x$ and $y_n \to y$ are convergent sequences in X such that x_n and y_n belong to the same element of $\mathcal{G}$ for all n, then x and y belong to the same element of $\mathcal{G}$.* □

Theorem 14 (Moore) *Let $\mathcal{G}$ be a monotone upper semi-continuous decomposition of a topological sphere S such that no decomposition element separates S. Then the quotient space obtained by collapsing each decomposition element to a point is a topological sphere.*

The side identifications will be realized by *arc bands* with which m.u.s.c. decompositions of the sphere will be constructed. These arc bands are of two types:

- A *rectangular arc band η of size y* is the image of an embedding $\psi\colon [0,1]\times[0,y] \to S^2$ which restricts to a Euclidean isometry on $\{0\}\times[0,y]$ and on $\{1\}\times[0,y]$, foliated by the leaves $\psi([0,1]\times\{t\})$ for $0 \le t \le y$. The *vertices* of η are the points $\psi(0,0)$, $\psi(0,y)$, $\psi(1,0)$, and $\psi(1,y)$.
- A *semi-circular arc band η of size y* is the image of an embedding
$$\psi\colon \{(s,t)\in\mathbb{R}^2 \,:\, t\in[-y,y],\, s\in[0,\sqrt{y^2-t^2}]\} \to S^2$$
which restricts to a Euclidean isometry on $\{0\}\times[-y,y]$, foliated by the leaves
$$\psi\left(\{(\sqrt{r^2-t^2},t) \,:\, t\in[-r,r]\}\right)$$
for $0 \le r \le y$. The *vertices* of η are the points $\psi(0,-y)$ and $\psi(0,y)$, and its *centre* is $\psi(0,0)$.

The decomposition of an arc band η into the leaves of its foliation is clearly m.u.s.c.

7.2.1 Identifications of vertical sides

Place isometric copies of the rectangles R_i in S^2 along the real edges of τ, as illustrated for the running example in Figure 22.

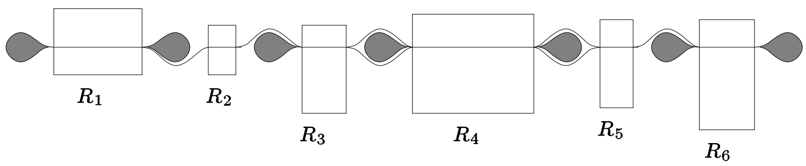


Figure 22: The rectangles R_i placed along the real edges of τ

For each infinitesimal edge e_i of τ, place an arc band η_i of size y_i in S^2, joining the segments of length y_i on the vertical sides of the rectangles corresponding to the real edges adjacent to the endpoints of e_i. If e_i is a bubble then these segments are adjacent on the same vertical side of the same rectangle, and η_i is a semi-circular arc band; otherwise, η_i is rectangular. The arc bands are placed on the sphere in the configuration dictated by τ, and are chosen to be mutually disjoint except possibly at their vertices. By Lemma 12, the arc bands incident on any given vertical side of a rectangle precisely cover that side. When there are infinitely many infinitesimal edges in a given junction, then some care is needed in the placement of these arc bands to ensure that Moore's theorem can be applied (despite the fact that the arc bands only serve to indicate how the vertical sides of the rectangles should be identified, and from this point of view the details of their placement is irrelevant).

$S^{\pm}$ Choose the semi-circular arc bands with diameters converging to 0, which is possible since their sizes converge to 0.

$W^{\pm}$, $V_3^{\pm}$ Choose the semi-circular arc bands with diameters converging to 0. It is then possible to choose the rectangular arc bands so that their diameters converge to 0 (see Figure 23 for the arc bands in the junction between R_3 and R_4 in the running example). Note that each bounded complementary component of the union of the arc bands in the junction is a 3-gon.

$V_1^{\pm}$, $V_1^{\pm} B$ Choose the semi-circular arc bands with diameters converging to 0. The rectangular arc bands can then be chosen so that they converge in the Hausdorff topology to a single arc v joining the two switches of the junction, and separating all but one of the arc bands from the puncture. Note that the bounded complementary components of the union of the arc bands comprise one 2-gon (containing the puncture), and infinitely many 3-gons.

$V_2^\pm$, $V_2^\pm B^2$ Choose the semi-circular arc bands with diameters converging to 0, and the rectangular arc bands so that the sequence of arc bands above (respectively below) the puncture converges to an arc v^+ (respectively v^-) joining the two switches, and v^+ and v^- are disjoint except at their endpoints. Again, the bounded complementary components of the union of the arc bands comprise one 2-gon containing the puncture, and infinitely many 3-gons.

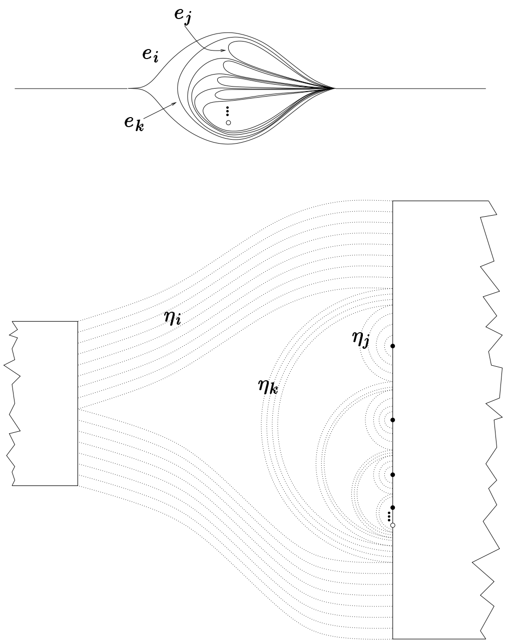


Figure 23: The arc bands in the junction between R_3 and R_4

Let R denote the closure of $\bigcup_{i<N} R_i \cup \bigcup_{i\geq N} \eta_i$. The complement of R in S^2 consists of one unbounded component C_∞ (containing ∞), finitely many open 2-gons containing punctures, and countably many open 3-gons. Define a decomposition of the sphere with elements of four types:

i) the closure of each bounded complementary component of R;

ii) the connected components of the intersection of the closure of C_∞ with the closure of the union of the arc bands;

iii) the arcs in each arc band except those contained in elements of i) or ii);

iv) single points in the rectangles and in C_∞ which are not contained in elements of i), ii), or iii).

Decomposition elements of type ii), which would not otherwise be decomposition elements of type iii), arise in the following circumstances: in 2-junctions of type B, BP, $W^{\pm}$, and $V_1^{\pm}B$ there is one such element which is the union of two arcs with a common endpoint; in 2-junctions of type $V_2^{\pm}B^2$, there are two such elements, each the union of two arcs with common endpoints; and in junctions of type $S^{\pm}$ there is one such element, which is the closure of the union of infinitely many arcs.

Lemma 15 *This decomposition of S^2 is m.u.s.c. and none of its elements separates S^2.*

Proof It is clear that the decomposition elements don't separate the sphere. Thus it is required to show (using Lemma 13) that if ζ_n is a sequence of decompositions elements, and $x_n \to x$, $y_n \to y$ are convergent sequences in S^2 with $x_n, y_n \in \zeta_n$ for all n, then x and y lie in the same decomposition element.

If all but finitely many of the ζ_n are single points then $x = y$, so passing to subsequences it can be assumed that none of the ζ_n is a single point. Since there are only finitely many decomposition elements of type ii), each compact, it can be assumed that none of the ζ_n is of type ii). If infinitely many of the ζ_n belong to a given arc band then the result follows by the upper semi-continuity of each arc band, so it can be assumed that no two of the ζ_n belong to the same arc band. If infinitely many of the ζ_n are semi-circular arcs, then $x = y$ since the diameters of the semi-circular arc bands tend to 0, so it can be assumed that no ζ_n is a semi-circular arc. If infinitely many of the ζ_n are the closure of the same complementary component of R, then the result follows by the compactness of these closures. Thus there are only two cases left to consider: where (ζ_n) is a sequence of arcs from distinct rectangular arc bands, and where (ζ_n) is a sequence of distinct closures of complementary components of R.

In either case, it can be assumed that the rectangular arc bands or complementary components are all in a single junction of type $V_1^{\pm}$ or $V_2^{\pm}$, since the diameters of rectangular arc bands and complementary components in junctions of other types tend to 0. Then by construction, x and y lie on the limiting arcs v or $v^{\pm}$ in that junction, which are contained in the closure of the complementary component of R containing the periodic point in that junction. □

It follows by Theorem 14 that the quotient space obtained by collapsing each decomposition element to a point with the induced quotient topology is a topological sphere $\tilde{S}$. The projection of the complement of C_∞ is a closed topological disk which is denoted $\mathcal{R}$. The horizontal line segments in the rectangles project to a singular foliation of $\mathcal{R}$ at all but finitely many points (the accumulations of singularities). There is a 3-pronged singularity at each point of the quotient corresponding to a bounded complementary 3-gon, and a 1-pronged singularity at each point corresponding to the centre of a semi-circular arc band. The disk $\mathcal{R}$ and parts of its foliation in the case of the running example is depicted in Figure 24, and details of the shaded regions of Figure 24 are shown in Figure 25.

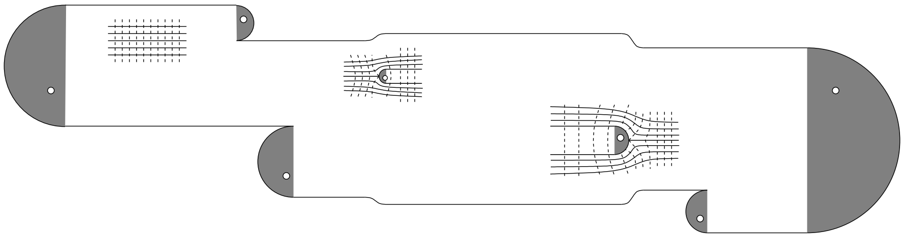


Figure 24: The quotient disk $\mathcal{R}$ and its foliation

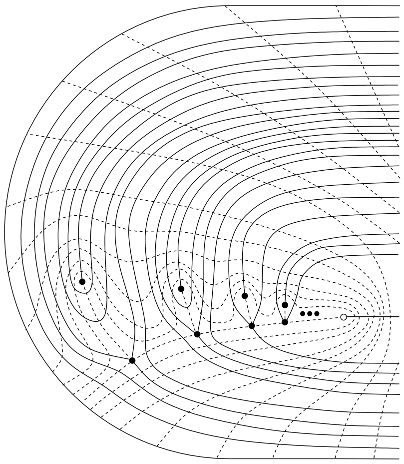


Figure 25: Detailed view of the shaded regions of Figure 24

The train track map $\phi\colon \tau \circlearrowleft$ induces a map $\tilde{\Phi}\colon \mathcal{R} \circlearrowleft$: unstable leaves (coming from horizontal segments) are stretched by the factor λ, stable leaves (coming from vertical segments) are contracted by the factor $1/\lambda$, the (projections of) rectangles are mapped as dictated by the action of ϕ on the real edges of τ and the identifications are mapped according to the action of ϕ on the infinitesimal edges of τ.

7.2.2 Identifications of horizontal sides

The identifications of the horizontal sides of $\mathcal{R}$ are carried out in such a way as to restore the injectivity of $\tilde{\Phi}$ on $\partial\mathcal{R}$. In order to describe them, it is therefore necessary to understand the structure of $\partial\mathcal{R}$ and the action of $\tilde{\Phi}$ on it. This was the purpose of the analysis of the outside map of the unimodal map f in Section 5: the dynamics of $\tilde{\Phi}|_{\partial\mathcal{R}}$ is modelled by the dynamics of the outside map, in a sense made precise by the following lemma.

Lemma 16 *Let $\theta\colon S^1 \to S^1 \cup I$ be the outside map of f. Then there is a homeomorphism $h\colon S^1 \to \partial\mathcal{R}$ with the property that $\theta(x) \in S^1$ if and only if $\tilde{\Phi}(h(x)) \in \partial\mathcal{R}$, and in this case $h(\theta(x)) = \tilde{\Phi}(h(x))$.*

Proof An analysis of the different possible configurations of infinitesimal edges in the junctions of $\mathbb{I}$, as given by Theorems 10 and 11, shows that for any kneading sequence $s \in \mathcal{MIA}$ the two corners on the left of R_1 are identified, the two corners on the right of R_{N-1} are identified, and for $1 \le j < N-1$ the top (respectively bottom) right corner of R_j is identified with the top (respectively bottom) left corner of R_{j+1}. The corners of the rectangles therefore define $2N-2$ points on $\partial\mathcal{R}$, which correspond naturally with the $2N-2$ points on S^1 defined by the critical orbit of f (namely $\hat{a}$ and $\hat{b}$, together with x_l and x_u for each other point x in the critical orbit).

Define $h\colon S^1 \to \partial\mathcal{R}$ to map each of these $2N-2$ points on S^1 to the corresponding point on $\partial\mathcal{R}$, and to map each interval of S^1 between two such points affinely onto the corresponding interval of $\partial\mathcal{R}$. Then h is a homeomorphism, the points $x \in S^1$ with $\theta(x) \notin S^1$ are precisely those for which $\tilde{\Phi}(h(x)) \notin \partial\mathcal{R}$, and away from these points both θ and $\tilde{\Phi}$ are affine on each of the $2N-2$ intervals, and send corresponding endpoints to corresponding endpoints. □

Abusing notation, let $\gamma \subseteq \partial\mathcal{R}$, $a, c, p \in \partial\mathcal{R}$ be defined as $h(\gamma)$, $h(\hat{a})$, $h(c_u)$, $h(p_u)$ respectively (see Figure 26).

Theorem 17 *Let $s = \kappa(F)$, and suppose that $q(s) = m/n \in (0, 1/2) \cap \mathbb{Q}$. Then $\tilde{\Phi}|_{\partial\mathcal{R}}$ has a unique periodic orbit, which has period n and rotation number m/n. The boundary $\partial\mathcal{R}$ has the structure of an n-sided polygon, with vertices the points $\tilde{\Phi}^j(a)$ for $1 \le j \le n$. Moreover, $v = \tilde{\Phi}^n(a) \in \overline{\gamma}$ satisfies:*

i) *$v = a$ (so that a is periodic under $\tilde{\Phi}$) if and only if $s = \mathrm{lhe}(m/n)$*

ii) *$v = c$ if and only if $s = \mathrm{NBT}(m/n)$*

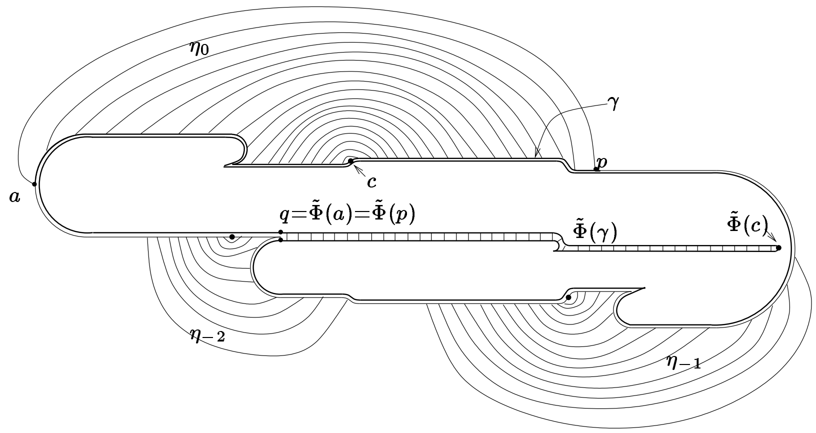


Figure 26: $\mathcal{R}$ and its image under $\tilde\Phi$

iii) *$v = p$ (so that p is periodic under $\tilde\Phi$) if and only if* $s = \mathrm{rhe}(m/n)$

iv) *v is strictly between a and c if and only if* $\mathrm{lhe}(m/n) \prec s \prec \mathrm{NBT}(m/n)$

v) *v is strictly between c and p if and only if* $\mathrm{NBT}(m/n) \prec s \prec \mathrm{rhe}(m/n)$

Proof The statements about the periodic orbit on $\partial\mathcal{R}$ and the position of v are direct translations of Theorem 6. An analysis of the different possible configurations of infinitesimal edges in the junctions of $\mathbb{I}$ shows that there are singularities on $\partial\mathcal{R}$ corresponding to each junction with configuration $S^\pm$, and cusps on $\partial\mathcal{R}$ corresponding to each 2-junction with configuration BP, B, $W^\pm$, $V_1^\pm B$, or $V_2^\pm B^2$. By Theorems 10 and 11, there are exactly n such junctions, and these junctions are precisely those containing the first n images under F of the left hand 1-junction. □

Remark 4 Observe that if s is not one of the endpoints of $\mathrm{KS}(m/n)$, then the vertices of $\partial\mathcal{R}$ correspond exactly to the vertices of the unbounded complementary component of τ.

The m.u.s.c. decomposition giving the identifications on the horiztonal edges can now be described: as before, it is constructed using rectangular and semi-circular arc bands joining points on the horizontal edges which will be identified. These arc bands will be labelled η_j for $j \le 0$ (in fact η_j will often denote a union of two arc bands), with the arcs of η_j joining precisely those points of $\partial\mathcal{R}$ which are identified by $\tilde\Phi^{1-j}$ but not by $\tilde\Phi^{-j}$: thus $\tilde\Phi$ will project to a homeomorphism of the quotient space. In particular, η_0 is a semi-circular arc band centred at c whose arcs join pairs of points of $\overline{\gamma}$ which are identified by $\tilde{\Phi}$. This is illustrated for the running example in Figure 26, where the image of γ has been depicted slightly separated for clarity. Notice that $\tilde{\Phi}|_{\partial\mathcal{R}}^{-1}$ is discontinuous at $q = \tilde{\Phi}(a) = \tilde{\Phi}(p)$ and is not surjective since it misses γ, but away from q it is a continuous injection (the injectivity is what makes it possible to construct the arc bands η_j in such a way that they intersect $\partial\mathcal{R}$ with disjoint interiors).

The preimages $\gamma_j = \tilde{\Phi}^j(\gamma)$ are connected for $-n+1 \leq j < 0$, since $\overline{\gamma_{-n+1}}$ is the first preimage of $\overline{\gamma}$ which contains the point of discontinuity q by Theorem 17. Thus the arc bands η_j for $-n+1 \leq j < 0$ are semi-circular, centred on $c_j = \tilde{\Phi}^j(c)$, and can be chosen to be disjoint.

For $j \leq -n$ there are three different cases, according as s is an endpoint of $\mathrm{KS}(m/n)$, is equal to $\mathrm{NBT}(m/n)$, or is neither. As in the statement of Theorem 17, let $v = \tilde{\Phi}^n(a)$ denote the vertex of $\partial\mathcal{R}$ which lies in $\overline{\gamma}$.

a) *s is an endpoint of* $\mathrm{KS}(m/n)$*:* Then v coincides with an endpoint of γ and is periodic of period n. Thus q is one of the endpoints of γ_{-n+1} so the preimages γ_j are connected for all $j < 0$. The arc bands η_j are semi-circular centred at c_j for all $j \leq 0$. Since their sizes tend to 0, they can be chosen to be disjoint away from their vertices and to have diameters tending to 0 as $j \to -\infty$. They cover $\partial\mathcal{R}$ by Theorem 6, accumulate on the periodic orbit on $\partial\mathcal{R}$, and their vertices lie on the orbit of preimages of this periodic orbit (described in Theorem 6). Figure 27 depicts these arc bands in the case $s = \mathrm{lhe}(1/3)$: the white dots depict the points of the period 3 orbit on $\partial\mathcal{R}$. The m.u.s.c. decomposition consists of the

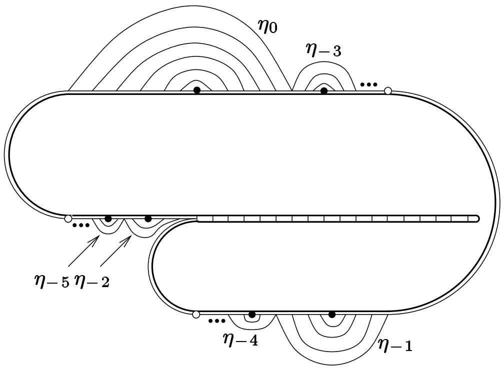


Figure 27: The arc bands in case a): here $s = (101)^\infty = \mathrm{lhe}(1/3)$

following elements:

i) the closure of the complement of $\mathcal{R} \cup \bigcup_{j=0}^{-\infty} \eta_j$, which is a closed topological disk bounded by infinitely many arcs, containing the periodic orbit on its boundary and the point at infinity in its interior;

ii) the arcs of each arc band $\eta_j, j \leq 0$, except for those contained in elements of i);

iii) the points of $\mathrm{Int}(\mathcal{R})$.

b) $s = \mathrm{NBT}(m/n)$: Then $v = c$, so $q = \tilde{\Phi}^{-n+1}(c)$ and the point of discontinuity of $\tilde{\Phi}|_{\partial\mathcal{R}}^{-1}$ coincides with the centre of η_{-n+1}. Thus $\overline{\gamma_{-n}}$ is the union of two arcs, and each point in each arc is paired with exactly one point in the other. Inductively, the same is true for $\overline{\gamma_j}$ for all $j \leq -n$: thus the η_j for $j \leq -n$ can be chosen to be rectangular arc bands, one of whose boundaries coincides with the outer boundary of η_{j+n} and the other with the inner boundary of η_{j-n}. The sizes of these arc bands decrease exponentially, and they cover $\partial\mathcal{R}$ by Theorem 6. They are depicted in Figure 28 for the case $s = \mathrm{NBT}(1/3)$. The m.u.s.c. decomposition consists of the following elements:

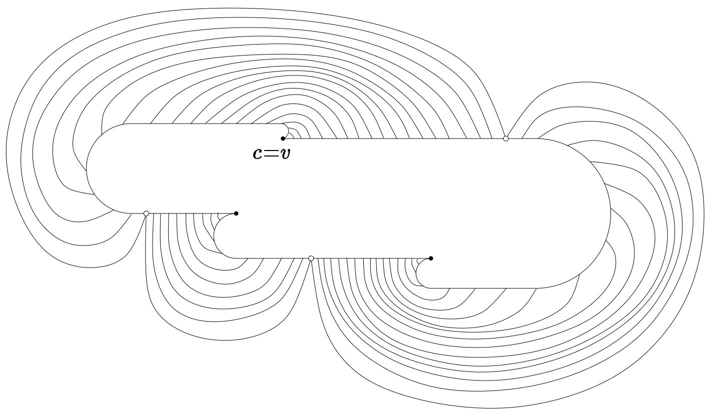


Figure 28: The arc bands in the NBT case: here $s = (10011)^\infty$

i) the closure of the complement of $\mathcal{R} \cup \bigcup_{j=0}^{-\infty} \eta_j$, which is a closed topological disk bounded by n arcs, having the points of the periodic orbit on $\partial\mathcal{R}$ for vertices, and containing the point at infinity in its interior;

ii) the arcs of each arc band $\eta_j, j \leq 0$ (in this case, no such arc is contained in an element of i));

iii) the points of $\mathrm{Int}(\mathcal{R})$.

c) *s is neither* $\mathrm{NBT}(m/n)$ *nor an endpoint of* $\mathrm{KS}(m/n)$*:* As in case b), γ_j has two connected components for $j \leq -n$. However, because $v \neq c$, one of the components of each such γ_j is divided into two subarcs: the points of one subarc are paired with other points in the same subarc, while those in the other are paired with points in the other component of γ_j. In Figure 29, $\gamma_{-n+1} = \alpha_1 \cup \alpha_2 \cup \alpha_3$, points of $\tilde{\Phi}^{-1}(\alpha_1)$ are paired to points of $\tilde{\Phi}^{-1}(\alpha_3)$ whereas points of $\tilde{\Phi}^{-1}(\alpha_2)$ are paired to points of $\tilde{\Phi}^{-1}(\alpha_2)$ itself. Thus η_j is the union of a semi-circular arc band and a rectangular arc band; the diameters of the semi-circular bands tend to 0 as $j \to -\infty$, as do the sizes of the rectangular bands. The arc bands cover $\partial\mathcal{R}$ by Theorem 6.
The m.u.s.c. decomposition consists of the following elements:
   i) the closures of the (infinitely many) complementary components of $\mathcal{R} \cup \bigcup_{j\leq 0} \eta_j$; there is one exterior n-gon containing ∞ and having the points in the periodic orbit on $\partial\mathcal{R}$ for vertices, and infinitely many 3-gons;
   ii) the arcs of each arc band in $\eta_j, j \leq 0$, except for those contained in elements of i);
   iii) the points of $\mathrm{Int}(\mathcal{R})$.

The complete decomposition for the running example, which satisfies $\mathrm{lhe}(1/3) \prec s \prec \mathrm{NBT}(1/3)$, is shown in Figure 30.

Lemma 18 *Each of these decompositions is m.u.s.c. and none of its elements separates* $\tilde{S}$*.*

Proof In case a), where the diameters of the arc bands tends to 0, and case b), where the arc bands can be regarded as a finite number of semi-circular arc bands, the proof is straightforward. For case c), let ζ_n be a sequence of decomposition elements, and $x_n \to x$, $y_n \to y$ be convergent sequences with $x_n, y_n \in \zeta_n$ for all n. Then it is required to show that x and y lie in the same decomposition element. As in the proof of Lemma 15, it can be assumed that either every ζ_n is an arc of a distinct rectangular arc band, or every ζ_n is a distinct complementary component. In both cases, both x and y lie on the boundary of the complementary component containing ∞. □

By Theorem 14 the quotient space is a topological sphere S, and by construction $\tilde{\Phi}$ projects to a homeomorphism $\Phi\colon S \circlearrowleft$. The following theorem asserts that Φ is a generalized pseudo-Anosov map, and describes the orbits of 1-pronged singularities, which are the most important from the point of view of

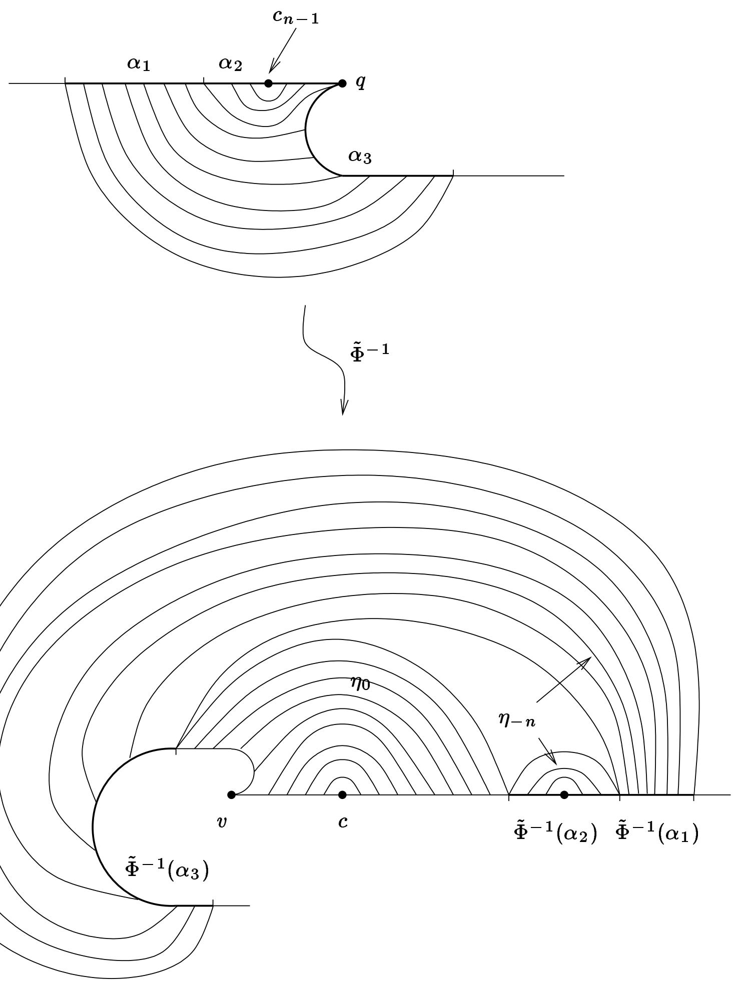


Figure 29: The arc bands in case c)

dynamics. Recall that an orbit $\{f^n(x) : n \in \mathbb{Z}\}$ of a homeomorphism $f\colon X \circlearrowleft$ is *homoclinic* to a fixed point p of f if $f^n(x) \to p$ as $n \to \pm\infty$.

Theorem 19 *Let $s \in \mathcal{MIA}$ have height $q(s) = m/n \in (0, 1/2)$ and let $\Phi_s\colon S \circlearrowleft$ denote the associated homeomorphism constructed above. Then Φ_s is a generalized pseudo-Anosov map. Moreover, it has exactly one orbit of 1-pronged singularities which*

i) *is homoclinic to the point at infinity if and only if s is an endpoint of the interval $\mathrm{KS}(m/n)$;*

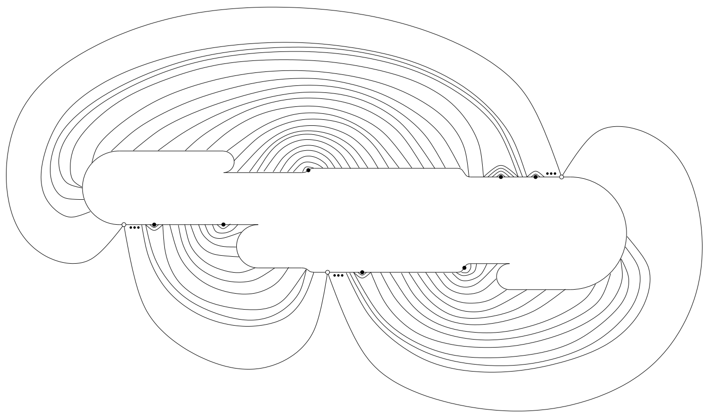


Figure 30: The complete second stage decomposition for the running example

ii) *is finite if and only if* $s = \mathrm{NBT}(m/n)$*;*

iii) *is backward asymptotic to the point at infinity and forward asymptotic to the periodic orbit determined by* s *if and only if* $\mathrm{lhe}(m/n) \prec s \prec \mathrm{NBT}(m/n)$ *or* $\mathrm{NBT}(m/n) \prec s \prec \mathrm{rhe}(m/n)$*.*

Proof Let $\pi\colon S^2 \to S$ denote the projection map given by the construction, and let $Q = \bigcup_{i=1}^{N-1} R_i$. Then $\pi(Q) = S$, since every fibre of π intersects Q; and $\pi|_{\mathrm{Int}(Q)}$ is an embedding, since every fibre which intersects $\mathrm{Int}(Q)$ is trivial. Thus the foliations of Q by horizontal and vertical line segments project to regular foliations of $\pi(\mathrm{Int}(Q))$, which are equipped with transverse measures induced by Euclidean distance in Q. The arc bands describe isometric identifications of segments of ∂Q, with vertical (respectively horizontal) segments always identified with other vertical (respectively horizontal) segments. Thus the measured foliations are also regular at points of $\pi(\partial Q)$ other than those which are the projections of vertices of arc bands, or centres of semi-circular arc bands. The foliations have a 1-pronged singularity at $\pi(c)$ whenever c is the centre of a semi-circular arc band. If v is the vertex of an arc band, there are more possibilities.

a) If v is the vertex of an arc band on the interior of a vertical segment of ∂Q corresponding to a junction of any type other than $S^{\pm}$, then $\pi(v)$ is a 3-pronged singularity.

b) If v is the vertex of an arc band on a vertical segment of ∂Q corresponding to a junction of type $S^{\pm}$, then $\pi(v)$ is an essential (∞-pronged) singularity.

c) If v is a vertex of an arc band on a horizontal segment in the case where s is an endpoint of $\mathrm{KS}(m/n)$, then $\pi(v)$ is an essential singularity.

d) If v is a vertex of an arc band on a horizontal segment in the case where $s = \mathrm{NBT}(m/n)$, then $\pi(v)$ is an n-pronged singularity if v is a point of the periodic orbit on ∂Q, and is a regular point otherwise.

e) If v is a vertex of an arc band on a horizontal segment in the case where $s \neq \mathrm{NBT}(m/n)$ and s is not an endpoint of $\mathrm{KS}(m/n)$, then $\pi(v)$ is an essential singularity if v is a point of the periodic orbit on ∂Q, and is a 3-pronged singularity otherwise.

By considering the singularities corresponding to junctions of each given type, it can easily be seen that junctions of types BP, B, and V_0 give rise to only finitely many singularities, whereas all other types give rise to infinitely many singularities which accumulate on the periodic point in the associated junction. If $s \neq \mathrm{NBT}(m/n)$ then the singularities arising from arc bands on horizontal segments of ∂Q accumulate on the point $\pi(P)$, where P is the periodic orbit on ∂Q; while if $s = \mathrm{NBT}(m/n)$ then there are only finitely many such singularities.

Thus the singularities of the foliations of S accumulate at only finitely many points. By construction, these foliations are invariant under Φ, with the horizontal (unstable) foliation being expanded by a factor $\lambda > 1$ and the vertical (stable) foliation being contracted by a factor $1/\lambda$. Hence Φ is a generalized pseudo-Anosov map as required.

To show that the one-pronged singularities arising from arc bands on the vertical sides of ∂Q lie on a single orbit, it is necessary to consider the action of the train track map $\phi\colon \tau \circlearrowleft$ on the bubbles of τ. If $s = \mathrm{NBT}(m/n)$, then it is clear that the $n+2$ bubbles are permuted cyclically by ϕ. In all other cases where s is periodic, the proof of Theorem 10 shows that there is a bubble B in junction N which is the image of an infinitesimal edge in junction $\pi^{-1}(N)$ joining the two switches of the junction and passing above the puncture, and all other bubbles are images of this one. Hence the forwards Φ-orbit of the 1-pronged singularity corresponding to B contains precisely all of the 1-pronged singularities arising from arc bands on the vertical sides of ∂Q. Similarly, if s is preperiodic then all of the bubbles of τ are images of the unique bubble B of τ in junction N.

It is immediate from the construction of Section 7.2.2 that the one-pronged singularities arising from arc bands on the horizontal sides of ∂Q are precisely the points of the backwards Φ-orbit of the point $\pi(c)$, since the centres of the arc bands on $\partial\mathcal{R}$ are at the points $\tilde{\Phi}^{-j}(c)$ for $j \geq 0$. Since $\tilde{\Phi}(c)$ lies on the right hand edge of Q, it follows that $\Phi(\pi(c))$ is the unique one-pronged singularity corresponding to a bubble in junction N whose preimage has not already been assigned, and hence the 1-pronged singularities form a single orbit as required. That the asymptotics of this orbit are as given in the statement of the theorem follows directly from the facts about 1-pronged singularities established earlier in the proof. □

Remark 5 $s = 10^\infty$ is the only element of $\mathcal{MIA}$ with height 0, and was excluded from the construction above in order not to have to introduce exceptions into each statement about $\mathrm{KS}(m/n)$. It can, however, be treated in exactly the same way as other elements of $\mathcal{MIA}$: the invariant generalized train track, depicted in Figure 5, is described by

$$(S^+ \ ; \ B)$$

(and hence is unique among train tracks corresponding to strictly preperiodic kneading sequences in having a junction of type S^+), and the same construction can be applied starting with this train track to obtain a 'tight horseshoe'. This has been described in [4].

8 The complex structure

In this section it is shown that the topological sphere S constructed above carries a natural complex structure which makes it into a complex sphere. With respect to this structure the Φ-invariant foliations become the horizontal and vertical trajectories of an integrable quadratic differential which is meromorphic away from finitely many essential singularities, with respect to which Φ is a Teichmüller mapping. This is a special case of results sketched in [4].

Theorem 20 *Let $s \in \mathcal{MIA}$, and let $\Phi_s\colon S \circlearrowleft$ be the associated generalized pseudo-Anosov map. Then the Euclidean structure on the rectangles used in the construction of S induces a complex structure on S, with respect to which it is a complex sphere.*

Proof As in the proof of Theorem 19, let $\pi\colon S^2 \to S$ be the projection map given by the construction of S and $Q = \bigcup_{i=1}^{N-1} R_i$, where the R_i are the rectangles corresponding to the real edges of the invariant train track τ. Since $\pi|_{\mathrm{Int}(Q)}$ is a homeomorphism onto its image, the Euclidean structure on $\mathrm{Int}(Q)$ induces a natural complex structure on $\pi(\mathrm{Int}(Q))$. Moreover, the conformal mappings

$z \mapsto z^{2/k}$ (on the slit plane) induce complex structures on the neighbourhoods of each k-pronged singularity. Thus S has a complex structure away from the finite set Σ of accumulations of singularities of the invariant foliations. To establish the theorem, it is necessary to show that the complex structure on $S \setminus \Sigma$ regards each point of Σ as a puncture rather than a hole. Using the theory of extremal length [1, 14], this can be accomplished by constructing a nested sequence of annuli converging to each point of Σ, the sum of whose moduli is divergent. The standard estimate

$$\mathrm{Mod}(A) \geq \frac{\mathrm{Width}(A)^2}{\mathrm{Area}(A)}$$

for the modulus of an annular region A will be used, where $\mathrm{Width}(A)$ is the minimum distance between the two boundary components of A.

Let $q(s) = m/n$. If $s = \mathrm{NBT}(m/n)$ then Σ is empty, and there is nothing to show. If $\mathrm{lhe}(m/n) \prec s \prec \mathrm{NBT}(m/n)$ or $\mathrm{NBT}(m/n) \prec s \prec \mathrm{rhe}(m/n)$ then there is one point of Σ corresponding to each junction of type W, V_1, V_2, or V_3 (ie, each junction containing infinitely many infinitesimal edges), and one point of Σ corresponding to the point at infinity. Finally, if $s = \mathrm{lhe}(m/n)$ or $s = \mathrm{rhe}(m/n)$ then Σ contains a single point. This case is harder, since the annuli cannot be constructed locally in Q, and is treated first.

Case 1. $s = \mathrm{lhe}(m/n)$ *or* $s = \mathrm{rhe}(m/n)$

Suppose that $s = \mathrm{lhe}(m/n)$: the case $s = \mathrm{rhe}(m/n)$ is entirely analogous. Let M be the transition matrix for the train track map $\phi\colon \tau \circlearrowleft$, λ be its Perron-Frobenius eigenvalue, and y be its Perron-Frobenius eigenvector. A construction similar to that described below can be found in [9].

By Theorem 10, τ has $n-2$ infinitesimal edges which are not loops, one in each junction other than the leftmost and rightmost. Carrying out the identifications corresponding to these edges yields a Euclidean polygon E with n vertical and n horizontal sides. Infinitely many other identifications are applied to produce S, each of which involves identifying the two halves of an *identification interval* on the boundary of E. The identification intervals on the vertical sides of E are given by (the eigenvector entries corresponding to) the bubbles of τ, while those on the horizontal sides are given by the map $\tilde{\Phi}$ and its backwards iterates, as described in the proof of Theorem 17.

The following notation will be used:

- w_v is half the length of the longest identification interval on the vertical sides of E: in other words, w_v is the largest entry of y which corresponds to a bubble of τ.

- w_h is half the length of the longest identification interval on the horizontal sides of E (this interval is denoted γ in the proof of Theorem 17).
- $w = \min(w_v, w_h)$, $W = \max(w_v, w_h)$.
- $a^0, a^1, \ldots, a^{n-1}$ are the points of the periodic orbit on the boundary of E, labelled so that a^0 is the rightmost and $a^{i+1 \bmod n}$ is the image of a^i. Thus a^1 is the leftmost point of the orbit, denoted a in the proof of Theorem 17. Observe that each a^i is a vertex of E.
- Each side of E is the union of a sequence of adjacent identification intervals which converge to one of the points a^i. Denote by $(u^i_j)_{j=0}^\infty$ (respectively $(v^i_j)_{j=0}^\infty$) this sequence on the vertical (respectively horizontal) side of E with endpoint a^i.
- a^i_j and a^i_{j+1} (respectively b^i_j and b^i_{j+1}) are the endpoints of the identification interval u^i_j (respectively v^i_j). Thus
  i) The vertical (respectively horizontal) side of E containing the vertex a^i has a^i_0 (respectively b^i_0) as its other vertex.
  ii) (a^i_j) and (b^i_j) are both sequences converging to a^i. All of the points of $\bigcup_{i,j}\{a^i_j, b^i_j\} \cup \bigcup_i \{a^i\}$ are identified to the point at infinity in S.
  iii) $|u^0_0| = 2w_v$ (where $|I|$ denotes the length of an interval I), and $|v^{i_0}_0| = 2w_h$ for some i_0 (with the property that a^{i_0} is the topmost point of the periodic orbit on the boundary of E).
  iv)
$$\begin{aligned} |u^i_j| &= \frac{2w_v}{\lambda^{i+nj}} \qquad \text{and} \\ |v^i_j| &= \frac{2w_h}{\lambda^{(i-i_0 \bmod n)+nj}} \end{aligned}$$
  for $i = 0, 1, \ldots, n-1$ and $j \geq 0$.
- For each $j \geq 0$, let
$$r_j = \frac{w}{\lambda^{(n-1)+nj}}.$$
  That is, r_j is half the length of the smallest of the $2n$ identification intervals u^i_j and v^i_j.

A sequence of nested annuli X_k, converging to the point at infinity, will now be constructed. Each X_k will be given as the union of finitely many regions of three types.

Half-annuli For $0 \leq i < n$ and $1 \leq j \leq k$, let $A^i_{j,k}$ (respectively $B^i_{j,k}$) be the half-annulus centred at a^i_j (respectively b^i_j) with inner radius r_{k+1} and outer radius r_k. (These are half-annuli bounded by Euclidean semicircles and segments of the boundary of E: see Figure 31.)

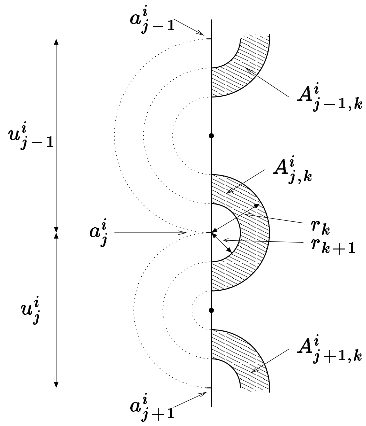


Figure 31: The half-annuli $A^i_{j,k}$

The identification intervals which share a^i_j as an endpoint are u^i_{j-1} and u^i_j and $r_k \leq |u^i_j|/2 < |u^i_{j-1}|/2$, so $A^i_{j,k}$ meets the identification intervals u^i_{j-1} and u^i_j between a^i_j and the centres of the intervals, as in Figure 31. The same argument applies to the half-annuli $B^i_{j,k}$.

After the identifications on the boundary of E have been carried out, each $\bigcup_{j=1}^k A^i_{j,k}$ and each $\bigcup_{j=1}^k B^i_{j,k}$ is a strip of width $r_k - r_{k+1}$ and area $k\pi(r_k^2 - r_{k+1}^2)/2$.

Rectangular quarter-annuli Figure 32 shows the half-annuli $A^i_{k,k}$ and $B^i_{k,k}$. The endpoints in u^i_k of the boundary semicircles of $A^i_{k,k}$ lie in the half of u^i_k which is further from a^i, and these endpoints are identified with two points in the half of u^i_k nearer to a^i, which are denoted z_1 and z_2 in the figure. There are analogous points z_3 and z_4 identified with endpoints of the half-annulus $B^i_{k,k}$. Since the internal angle of E at a_i is $\pi/2$, it is possible to define a rectangular quarter-annulus Z^i_k as shown in the figure, each of whose boundary components is the union of a horizontal and a vertical arc, joining either z_1 to z_3 or z_2 to z_4.

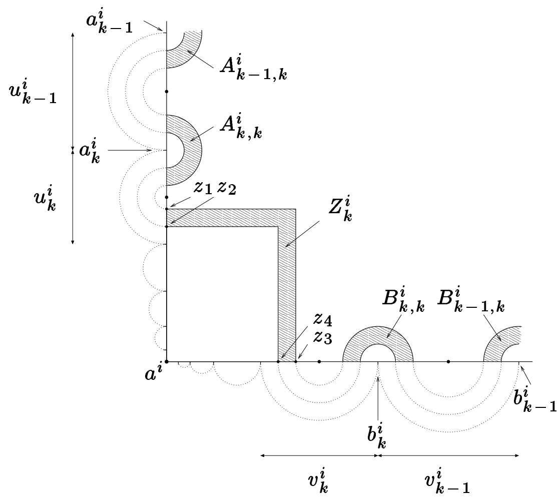


Figure 32: The rectangular quarter-annulus Z^i_k

A rectangular quarter-annulus is used since the distance between a^i and z_1 need not be equal to the distance between a^i and z_3. The former distance is less than $\sum_{j=k}^{\infty}|u^i_j|$ and the latter is less than $\sum_{j=k}^{\infty}|v^i_j|$. Hence, for each i and each k, the region

$$Y^i_k = Z^i_k \cup \bigcup_{j=1}^{k} A^i_{j,k} \cup \bigcup_{j=1}^{k} B^i_{j,k}$$

is (once the identifications have been carried out) a strip of width $r_k - r_{k+1}$ and area bounded above by

$$k\pi(r_k^2 - r_{k+1}^2) + (r_k - r_{k+1})\left(\sum_{j=k}^{\infty}|u^i_j| + \sum_{j=k}^{\infty}|v^i_j|\right).$$

Circular vertex annuli The other vertices $a^i_0 = b^{i'}_0$ of E have internal angle either $\pi/2$ or $3\pi/2$. At each such vertex, construct a (circular) quarter-annulus or three-quarter-annulus D^i_k with radii r_k and r_{k+1}, whose boundary components join points identified with endpoints of $A^i_{1,k}$ to points identified with endpoints of $B^{i'}_{1,k}$. These partial annuli have areas bounded above by $\pi(r_k^2 - r_{k+1}^2)$.

After the identifications have been carried out,

$$X_k = \bigcup_{i=0}^{n-1}(Y^i_k \cup D^i_k)$$

is an annulus of width $r_k - r_{k+1}$ and area bounded above by

$$n\left((k+1)\pi(r_k^2 - r_{k+1}^2) + (r_k - r_{k+1})\left(\sum_{j=k}^{\infty}|u_j^i| + \sum_{j=k}^{\infty}|v_j^i|\right)\right).$$

The annuli X_1 and X_2 in the case $s = \text{lhe}(1/3) = (101)^\infty$ are depicted in Figure 33.

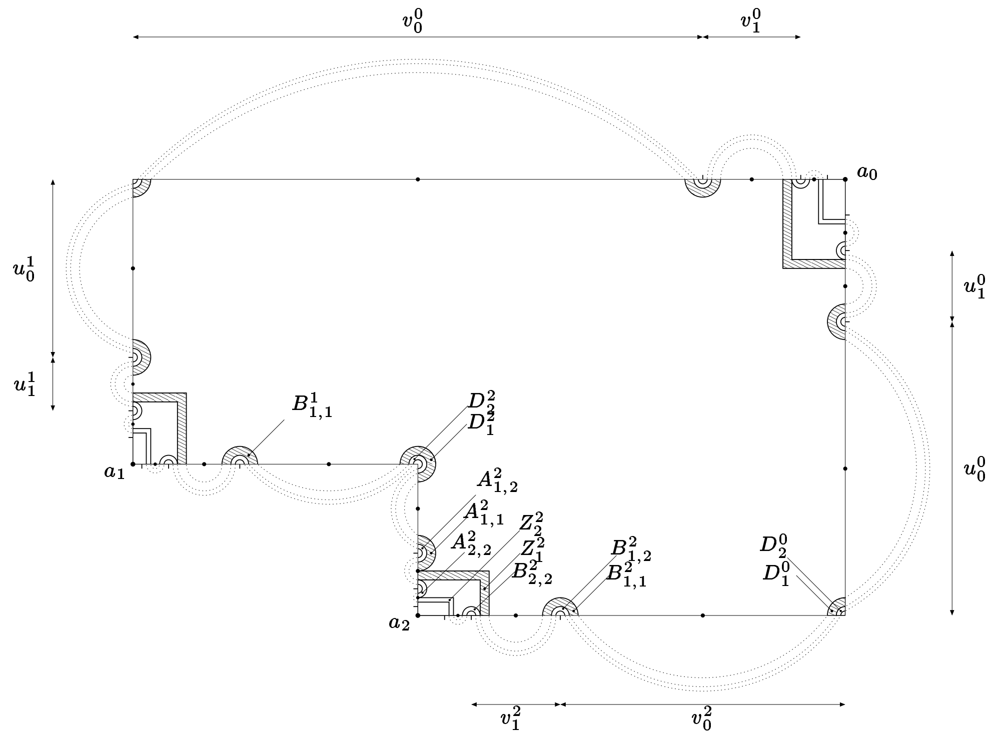


Figure 33: Annuli with divergent moduli sum surrounding the accumulation of singularities of $\Phi_{\text{lhe}(1/3)}$

Now

$$r_k \pm r_{k+1} = \frac{w}{\lambda^{(n-1)+nk}}(1 \pm \frac{1}{\lambda^n})$$

and

$$\sum_{j=k}^{\infty}|u_j^i| + \sum_{j=k}^{\infty}|v_j^i| \le \sum_{j=k}^{\infty}\frac{4W}{\lambda^{nj}} = \frac{4W\lambda^n}{\lambda^{nk}(\lambda^n - 1)},$$

so

$$\mathrm{Mod}(X_k) \geq \frac{\mathrm{Width}(X_k)^2}{\mathrm{Area}(X_k)} \geq \frac{r_k - r_{k+1}}{n\left((k+1)\pi(r_k + r_{k+1}) + \frac{4W\lambda^n}{\lambda^{kn}(\lambda^n - 1)}\right)} = \frac{C_1/\lambda^{nk}}{(C_2 k + C_3)/\lambda^{nk}} = \frac{C_1}{C_2 k + C_3},$$

where C_1, C_2, and C_3 depend only on n, λ, w, and W. Hence $\sum_{k\geq 1} \mathrm{Mod}(X_k)$ diverges as required.

Case 2. $\mathrm{lhe}(m/n) \prec s \prec \mathrm{NBT}(m/n)$ *or* $\mathrm{NBT}(m/n) \prec s \prec \mathrm{rhe}(m/n)$

Consider first the accumulation of singularities in a junction of type $W^{\pm}$. In Figure 34 the identifications on Q are indicated with dotted lines, and the shaded half-annuli A_k project to nested annular regions in S which converge to the accumulation of singularities p. If the width of the rectangle in which these half-annuli lie is less than half its height then not all of them can be constructed there: in this case, simply start with the first half-annulus A_{k_0} which can be constructed.

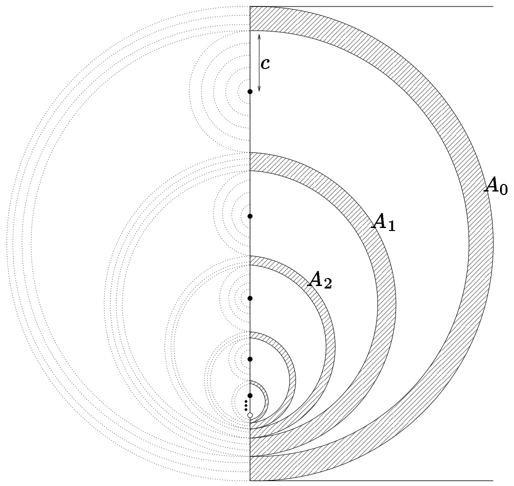


Figure 34: Annuli with divergent moduli sum surrounding an accumulation of singularities of type W.

Let $\mu = \lambda^{-N}$, where N is the period of p, let w be the width of the annulus A_0, and denote by c the size of the largest semi-circular arc band. Then $\mathrm{Width}(A_k) = w\mu^k$, and the k^{th} semi-circular arc band has size $c\mu^k$.

The outer radius R_k of A_k is one half of the distance between the highest and lowest points of A_k, in other words

$$\begin{aligned} R_k &= \frac{1}{2}\left(2w\sum_{j=k}^{\infty}\mu^j + 2c\sum_{j=k}^{\infty}\mu^j\right) \\ &= \frac{(w+c)\mu^k}{1-\mu}. \end{aligned}$$

Thus

$$\mathrm{Area}(A_k) = \frac{\pi}{2}\,\mathrm{Width}(A_k)(2R_k - \mathrm{Width}(A_k)) = C\mu^{2k},$$

where C depends only on μ, w, and c. Hence

$$\mathrm{Mod}(A_k) \geq \frac{\mathrm{Width}(A_k)^2}{\mathrm{Area}(A_k)} = \frac{w}{C}$$

so that $\sum_{k\geq k_0}\mathrm{Mod}(A_k)$ diverges as required.

Similar arguments apply to accumulations of singularities of types V. For a junction of type V_1, let R and R' denote the adjacent rectangles, where R is the rectangle adjacent to the switch to which all of the bubbles in the junction are attached. Then a sequence of half-annuli is constructed in R as for junctions of type W, but rather than being completed to annuli by the identifications they are completed by corresponding half-annuli in R'. The moduli of these annuli are again bounded below by a positive constant. Junctions of type V_2 are double covers of junctions of type W, and a sequence of annuli with moduli bounded below by a positive constant at an accumulation of type W thus induces a similar sequence at an accumulation of type V_2. Accumulations of type $V_3^{\pm}$ (respectively V_1B, V_2B^2) are identical to those of type $W^{\pm}$ (respectively V_1, V_2) in a neighbourhood of the accumulation point.

Finally, consider the point at infinity. The pattern of identification of horizontal sides, described in Section 7.2.2, is an n-fold cover of the pattern in junctions of type W: there are n points on the horizontal boundary of $\mathcal{R}$ which are identified to the point at infinity, and the identifications along the horizontal sides are given by an alternating sequence of rectangular and semi-circular arc bands whose sizes decay exponentially. □

References

[1] **L V Ahlfors**, *Conformal invariants: topics in geometric function theory*, McGraw-Hill Series in Higher Mathematics, McGraw-Hill Book Co. New York (1973)

[2] **M Bestvina**, **M Handel**, *Train-tracks for surface homeomorphisms*, Topology 34 (1995) 109–140

[3] **P Boyland**, *Topological methods in surface dynamics*, Topology Appl. 58 (1994) 223–298

[4] **A de Carvalho**, *Extensions, quotients and generalized pseudo-Anosov maps*, to appear in: "Graphs and patterns in mathematics and theoretical physics (Stony Brook, NY, 2001)", (AMS Proc. Sympos. Pure Math.)

[5] **A de Carvalho**, **T Hall**, *Pruning theory and Thurston's classification of surface homeomorphisms*, J. Eur. Math. Soc. (JEMS) 3 (2001) 287–333

[6] **A de Carvalho**, **T Hall**, *Braid forcing and star-shaped train tracks*, Topology 43 (2004) 247–287

[7] **W de Melo**, **S van Strien**, *One-dimensional dynamics*, Springer Erbebnisse series (3) 25, Springer-Verlag, Berlin (1993)

[8] **R Devaney**, *An introduction to chaotic dynamical systems*, second edition, Addison-Wesley Publishing Company Advanced Book Program, Redwood City, CA (1989)

[9] **C Earle**, **F Gardiner**, *Teichmüller disks and Veech's $\mathcal{F}$-structures*, from: "Extremal Riemann surfaces (San Francisco, CA, 1995)", Contemp. Math. 201, Amer. Math. Soc. Providence, RI (1997) 165–189

[10] **J Franks**, **M Misiurewicz**, *Cycles for disk homeomorphisms and thick trees*, from: "Nielsen theory and dynamical systems (South Hadley, MA, 1992)", Contemp. Math. 152, Amer. Math. Soc. Providence, RI (1993) 69–139

[11] **F R Gantmacher**, *The theory of matrices. Vols. 1, 2*, (translated by K A Hirsch), Chelsea Publishing Co. New York (1959)

[12] **T Hall**, *The creation of horseshoes*, Nonlinearity 7 (1994) 861–924

[13] **A Katok**, **B Hasselblatt**, *Introduction to the modern theory of dynamical systems*, Encyclopedia of Mathematics and its Applications 54, Cambridge University Press, Cambridge (1995), with a supplementary chapter by Katok and L Mendoza

[14] **O Lehto**, **K I Virtanen**, *Quasiconformal mappings in the plane*, second edition, (translated from the German by K W Lucas), Grundlehren series 126, Springer-Verlag, New York (1973)

[15] **J Milnor**, **W Thurston**, *On iterated maps of the interval*, from: "Dynamical systems (College Park, MD, 1986–87)", Springer Lecture Notes series 1342, Springer, Berlin (1988) 465–563

[16] **R Moore**, *Foundations of point set theory*, AMS Colloquium Publications XIII, Amer. Math. Soc. Providence, RI (1962)

[17] **W P Thurston**, *On the geometry and dynamics of diffeomorphisms of surfaces*, Bull. Amer. Math. Soc. 19 (1988) 417–431