EXTENSIONS, QUOTIENTS AND GENERALIZED PSEUDO-ANOSOV MAPS

ANDRÉ DE CARVALHO

Dedicated to Dennis Sullivan on his 60^{th} birthday

Abstract. We describe a circle of ideas relating the dynamics of 2-dimensional homeomorphisms to that of 1-dimensional endomorphisms. This is used to introduce a new class of maps generalizing that of Thurston's pseudo-Anosov homeomorphisms.

1. Introduction

In this paper we discuss a circle of ideas which is present in many different contexts in dynamical systems. It was first introduced by Williams in his study of expanding attractors, and has been used since by many authors. In its most basic form it can be stated as follows:

Collapsing segments of stable manifolds of a homeomorphism yields a lower dimensional endomorphism; the original homeomorphism may be recovered by taking the natural extension (inverse limit) of the quotient endomorphism.

The context we focus attention on is the interplay this creates between the dynamics of 1-dimensional endomorphisms — endomorphisms of trees and graphs — and that of 2-dimensional homeomorphisms — homeomorphisms of surfaces. In recent years, this interplay between graph endomorphisms and surface homeomorphisms has been used for different, although related, purposes. For example, it was used to give algorithmic proofs of Thurston's classification theorem for surface homeomorphisms up to isotopy. It also appeared in the contruction of models for families of surface homeomorphisms passing from trivial to chaotic dynamics as parameters are varied. And a combination of these results led to a conjecture about the way in which *forcing* organizes the braid types of horseshoe periodic orbits.

Complexification should yield a closely related discussion — which will not be treated here — linking the dynamics of endomorphisms of branched (Riemann) surfaces to the dynamics of automorphisms of $\mathbb{C}^2$. The natural extension of an endomorphism of a branched surface is usually a homeomorphism of a surface lamination. Such laminated spaces have already been found to be among the main objects in the study of complex Hénon maps. In both the real and complex cases, much is known about 1-dimensional endomorphisms, at least when the space is unbranched, whereas about homeomorphisms in dimension 2 much less is known.

This double interplay — between 1- and 2-dimensional dynamics and between the real and complex settings — seems to be an interesting approach to explore. In this paper we develop some of its aspects, mostly in the real setting.

We also introduce a different kind of quotient, taking a surface homeomorphism to another surface homeomorphism by collapsing dynamically irrelevant (wandering) domains. This produces an interesting class of surface homeomorphisms which includes torus Anosovs and Thurston's pseudo-Anosov maps. These maps — called *generalized pseudo-Anosov maps* — preserve a pair of invariant measured foliations with (possibly infinitely many) singularities, giving the underlying surface a naturally defined complex structure. Regarded from the point of view of this complex structure, the invariant foliations become the horizontal and vertical trajectories of an integrable quadratic

Date: August 2001.

differential (with possibly infinitely many poles) and the generalized pseudo-Anosov map becomes a Teichmüller mapping.

The construction of generalized pseudo-Anosov maps is done by first introducing generalized train tracks: smooth branched 1-submanifolds, with possibly infinitely many branches, of the ambient surface. We describe how to find invariant generalized train tracks for certain surface homeomorphisms which, together with measure theoretic information obtained from transition matrices, are the main ingredients in the construction of generalized pseudo-Anosov maps. This is a variant of the same circle of ideas mentioned above.

The paper is organized as follows: in Section 2 the natural extension is defined, the class of *thick graph maps* is introduced and, in Lemma 1, the circle of ideas mentioned above is closed. In Section 3 the 0-*entropy* equivalence relation, which collapses dynamically irrelevant domains, is introduced and a theorem (Theorem 2) is stated about the quotient maps so obtained — this is the first way in which generalized pseudo-Anosov maps appear. In Section 4 infinite train tracks are defined as the main tool to construct generalized pseudo-Anosov maps in the following section. The complex structure is discussed in the last subsection of Section 5.

The writing style is Sullivanian as is fitting for a paper prepared for such an occasion. Many arguments are only sketched and some are omitted entirely.

Acknowledgements: I have worked on or around the main subjects in this paper for many years. Sevaral people in the course of several conversations have influenced me. Dennis Sullivan, of course, was a major influence as my adviser. Others who have directly contributed to the ideas in this paper are Michael Handel, with whom I had many conversations about many things, including infinite train tracks and generalized pseudo-Anosovs; Fred Gardiner, who taught me Teichmüller theory and how to show a puncture is a puncture; and my friend and coworker Toby Hall, whose work was the original stimulus for me to study the subject and who wrote a good part of Section 4 when we were engaged in a project which has now been put on hold. I would also like to thank the referee for several suggestions for improvement and clarification.

2. The natural extension

In this paper a *dynamical system* will mean a continuous self-map of a topological space, at least. As we go along, we may require more of our maps, for example, that they be differentiable or diffeomorphisms.

2.1. Definition and examples.

Let $f\colon X \to X$ be a continuous surjective map of a topological space X. If f is not invertible, there is a naturally associated invertible map: set

$$\hat{X} = \{(x_0, x_1, x_2, \dots) \in \prod_0^\infty X; f(x_{i+1}) = x_i, \text{for } i = 0, 1, 2, \dots\}$$

and define $\hat{f}\colon \hat{X} \to \hat{X}$ setting $\hat{f}(x_0, x_1, x_2, \dots) = (f(x_0), x_0, x_1, \dots)$. This map is called the *natural extension* of f. The space $\hat{X}$ is also known as the *inverse* or *projective limit space* and the map $\hat{f}$ as the *inverse* or *projective limit map* associated to $f\colon X \to X$.

Here are some prototypical examples of natural extensions which come up in a variety of contexts in the study of low-dimensional dynamical systems.

Examples 1. a) Let $X = \mathbb{S}^1$ be the unit circle in the complex plane and $f\colon \mathbb{S}^1 \to \mathbb{S}^1$ be the squaring map $f(z) = z^2$ in complex notation. Then $\widehat{\mathbb{S}^1}$ is the *dyadic solenoid*. It is a fiber bundle over $\mathbb{S}^1$ with fiber a dyadic Cantor set.

b) Let $\mathbb{C}^* = \mathbb{C} \setminus 0$ and $f\colon \mathbb{C}^* \to \mathbb{C}^*$ be again $f(z) = z^2$. Then $\widehat{\mathbb{C}^*}$ is the *complex* dyadic solenoid and is a fiber bundle over $\mathbb{C}^*$ with same fiber as above.

c) A variant of the previous example may be obtained by restricting f to $X = \mathbb{C} \setminus \mathbb{D}$, the exterior of the closed unit disk. Since the action of f on X has a fundamental domain (namely, any annulus of the form $A = \{z \in \mathbb{C}; 1 < R \leq |z| \leq R^2\}$), so does the action of $\hat{f}$ on $\hat{X}$ and we can take the quotient $\mathcal{S} = \hat{X}/\hat{f}$. This is *Sullivan's Riemann Surface Lamination.* $\mathcal{S}$ is a *lamination* as were the previous examples, but, unlike them, for which collapsing transversals produces a good old space, this time the quotient is the *branched surface* obtained as the quotient A/f of the annulus by identifying its boundary under the dynamics.

d) Let $I = [0,1]$ and f be the *tent map* $f(x) = 2x$ if $x \in [0, 1/2]$ and $f(x) = 2 - 2x$ if $x \in [1/2, 1]$. The inverse limit space $\hat{I}$ is the *Knaster continuum.* Again, $\hat{I}$ is a laminated space and the quotient by collapsing transversals is I, but $\hat{I}$ is not a fiber bundle over I. We will talk more about examples like this below.

2.2. **The natural extension in 2-dimensional dynamics.** In this section we introduce a class of maps which are suitable for several applications (see [BH95, dCH01, FM93]). They are called *thick graph maps* and, as the name suggests, they are essentially graph endomorphisms that have been thickened and made into surface homeomorphisms. All of their interesting dynamics is contained in a subsurface called a *thick graph*, that is, a graph in which each point has been thickened up, either to a disk or an arc according as the point is a vertex or a regular point of the graph. The homeomorphisms are assumed to act on the thick graph in such a way that they induce endomorphisms of the underlying graph.

Definitions 1. A *thick graph* is a pair $(S, \mathbb{G})$, where S is a closed orientable surface endowed with a fixed metric compatible with its topology and $\mathbb{G}$ is a compact subsurface of S (with boundary) which is partitioned into compact *decomposition elements*, such that

i) Each decomposition element of $\mathbb{G}$ is either a *leaf* homeomorphic to $[0, 1]$, or a *junction* homeomorphic to D^2 (the unit disk in $\mathbb{R}^2$).
ii) The boundary in $\mathbb{G}$ of each junction is a finite number of disjoint arcs: if there are k such arcs, then the disk is called a *k-junction.*
iii) The set of k-junctions with $k \neq 2$ is finite.
iv) Each decomposition element which is not in the accumulation of a sequence of distinct 2-junctions is contained in a chart as depicted in Figure 1.
v) Each component of $S \setminus \mathbb{G}$ is an open disk.

If $(S, \mathbb{G})$ is a thick graph, let $\sim$ be the equivalence relation on $\mathbb{G}$ given by $x \sim y$ if and only if x and y lie in the same decomposition element. Then $G = \mathbb{G}/\sim$ is a graph, whose vertices (which may have valence 2) correspond to the junctions of $\mathbb{G}$: the canonical projection will be denoted $\pi \colon \mathbb{G} \to G$. The vertex set of G will be denoted V, and the union of the junctions of $\mathbb{G}$ will be denoted $\mathbb{V}$: thus $\mathbb{V} = \pi^{-1}(V)$. The components of $\mathbb{G} \setminus \mathbb{V}$ are called *strips*: each strip is therefore homeomorphic to $(0, 1) \times [0, 1]$. The union of the closures of the strips will be denoted by $\mathbb{E}$ and the corresponding set in the quotient, the set of *edges* of G, will be denoted by E. Thus $\mathbb{E} \cap \mathbb{V}$ consists of a collection of closed arcs which are the boundary components (in ofg) of both the junctions and the strips of $\mathbb{G}$.

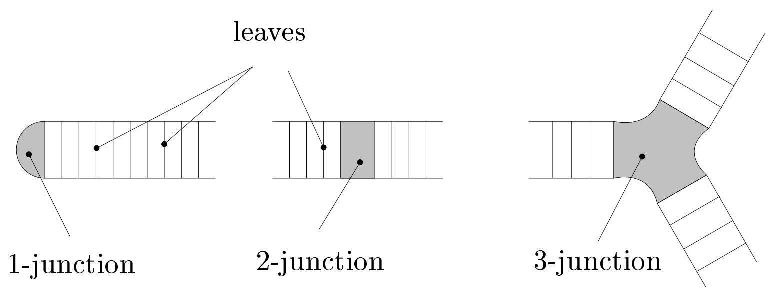

FIGURE 1. Charts in a thick graph.

Remarks 1. a) We will often refer to the thick graph as $\mathbb{G}$ alone, although it should be kept in mind that there is a surface floating around.

b) Notice that, at this point, we allow a thick graph to have infinitely many 2-junctions and, therefore, infinitely many strips (and, thus, the graph G to have infinitely many edges).

If $(S, \mathbb{G})$ is a thick graph and $F\colon (S, \mathbb{G}) \to (S, \mathbb{G})$ is a homeomorphism (i.e., a homeomorphism $F\colon S \to S$ with $F(\mathbb{G}) \subset \mathbb{G}$) under which the image of each decomposition element of $\mathbb{G}$ is contained in a decomposition element, then $F|_{\mathbb{G}}$ induces a graph endomorphism $f\colon G \to G$ such that $\pi \circ F|_{\mathbb{G}} = f \circ \pi$.

Definition 2. A *thick graph map of* $(S, \mathbb{G})$ is an orientation-preserving homeomorphism $F\colon (S, \mathbb{G}) \to (S, \mathbb{G})$ such that:

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

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

iii) The induced graph endomorphism $f\colon G \to G$ is piecewise monotone (that is, there is a finite subset L of V such that $f^{-1}(x) \cap U$ is connected for each $x \in G$ and each component U of $G \backslash L$); and is strictly monotone away from the preimages of vertices (that is, every $x \in G \setminus f^{-1}(V)$ has a neighborhood on which f restricts to an embedding).

iv) For each component U of $S \backslash \mathbb{G}$ there is a (least) positive integer n_U for which either $F^{n_U}(U) \subset \mathbb{G}$ or $F^{n_U}(U) \cap U \neq \emptyset$, in which case U contains a period n_U point p_U of F, which is a source whose immediate basin contains U: that is, $F^{-kn_U}(x) \to p_U$ as $k \to \infty$ for all $x \in U$.

Remarks 2. a) Item iv) in the definition says that the dynamics of a thick graph map in $S \setminus \mathbb{G}$ is easily understood and uninteresting.

b) If $F\colon (S, \mathbb{G}) \to (S, \mathbb{G})$ is a thick graph map, then so are all of its forward iterates F^n $(n \geq 1)$, and the graph endomorphism induced by F^n is $f^n\colon G \to G$.

c) Let $f\colon G \to G$ be the quotient of a thick graph map. A point $x \in G$ is a *critical point* if f is not a local homeomorphism at x. Since thick graphs may have infinitely many 2-junctions, the forward orbit of critical points of f may be infinite.

Example 2. The first example is Smale's horseshoe map which will be denoted by $F_1\colon (\mathbb{S}^2, \mathbb{I}) \to (\mathbb{S}^2, \mathbb{I})$ here and in what follows. It is shown in Figure 2. The thick graph in this case is a thick tree — a thick interval, in fact — and is denoted by $\mathbb{I}$. The point at infinity in $\mathbb{S}^2$ is a repeller whose basin contains all points outside $\mathbb{I}$. The horseshoe has two saddle fixed points which are labelled x_0 and x_1 (shown as $\bullet$ and $\circ$, respectively, in Figure 2) and an attracting fixed point in the 1-junction on the left denoted by x (shown as $\blacksquare$). The quotient tree is the interval and the quotient map — the 'flat top' tent map $f_1\colon I \to I$ — is also shown in the figure: the image is shown slightly separated so it is possible to see what f_1 does to the interval. This way of representing graph maps will always be used in what follows.

Example 3. In this example the map is again a sphere homeomorphism and the thick graph is again a thick tree as shown in Figure 3. It is denoted $F_2\colon (\mathbb{S}^2, \mathbb{T}) \to (\mathbb{S}^2, \mathbb{T})$. The 1- and 2-junctions of $\mathbb{T}$ contain a periodic orbit of period 6 ($\circ$) and the 3-junctions contain a periodic orbit of period 2 ($\bullet$). The quotient tree and tree endomorphism $f_2\colon T \to T$ are shown in Figure 4.

Remark 3. Both maps F_1 and F_2 above can be made to be diffeomorphisms. This will be used below when we talk about stable and unstable manifolds of their periodic points.

Let $F\colon \mathbb{G} \to \mathbb{G}$ be a thick graph map and $f\colon G \to G$ its quotient. The infinite nested intersection $\Lambda = \bigcap_{n=0}^{\infty} F^n(\mathbb{G})$ is a compact subset of $\mathbb{G}$ which is invariant under F. We now relate the dynamics of F on Λ with the natural extension of f (see [Bar86]).

Lemma 1. *The maps* $F|_\Lambda\colon \Lambda \to \Lambda$ *and* $\hat{f}\colon \hat{G} \to \hat{G}$ *are topologically conjugate, that is, there exists a homeomorphism* $\hat{\pi}\colon \Lambda \to \hat{G}$ *such that* $\hat{\pi} \circ F = \hat{G} \circ \hat{\pi}$.

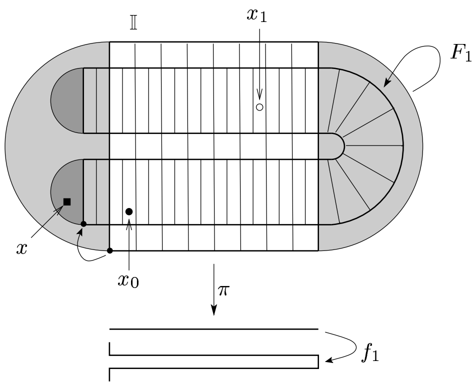

FIGURE 2. The horseshoe map.

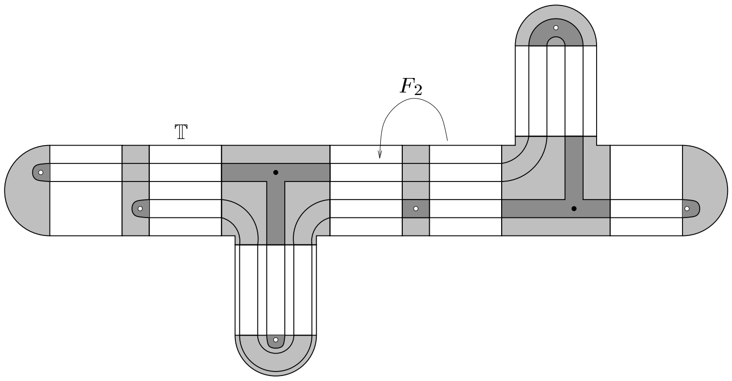

FIGURE 3. A thick tree map.

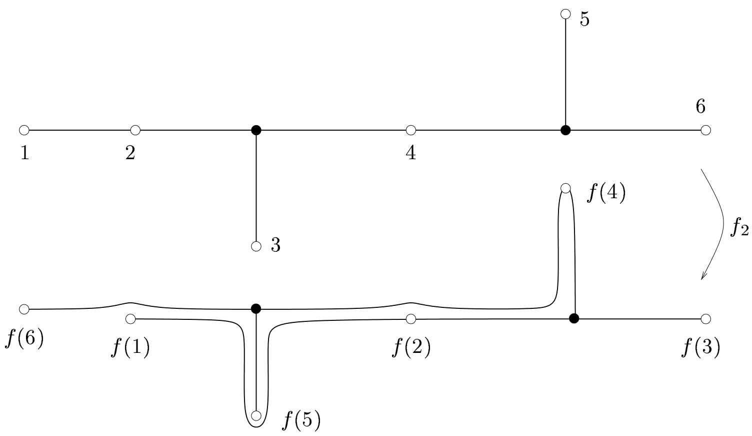

FIGURE 4. The quotient tree map of the thick tree map in Figure 3.

Proof. Let $\pi\colon \mathbb{G} \to G$ be the projection. The map $\hat{\pi}$ is defined by setting

$$\hat{\pi}(z) = (\pi(z), \pi(F^{-1}(z)), \pi(F^{-2}(z)), \dots)$$

for each $z \in \Lambda$. It is straightforward to check that $\hat{\pi}$ is well defined, continuous and surjective. Injectivity follows from the assumption that $\operatorname{diam}(F^n(\gamma)) \to 0$ as $n \to \infty$. $\square$

3. Another kind of quotient

In this section we describe a way of modifying two-dimensional maps by a semi-conjugacy which collapses 'irrelevant' dynamics: that is, parts of the space which do not carry entropy. The semi-conjugacy is defined quite generally and the space of maps it yields is quite interesting in its own right. A more thorough treatment of the equivalence relation can be found in [dCP02]. The space of quotient homeomorphisms will be treated in a future paper.

3.1. **The 0-entropy equivalence relation.** We start by recalling Bowen's definition of topological entropy [Bow71]. If X is a metric space and $F\colon X \to X$ is a uniformly continuous map, we say that $x, y \in X$ are (n, ϵ)-*separated* if it is possible to distinguish between the orbits of x and y up to $n-1$ iterates with precision ϵ. That is, $x, y \in X$ are (n, ϵ)-separated if $d(F^j(x), F^j(y)) > \epsilon$ for some $0 \leq j < n$. The topological entropy of F is defined to be the limit as $\epsilon \to 0$ of the exponential growth rate of the number of (n, ϵ)-separated orbits as $n \to \infty$. If $K \subset X$ is a compact subset and we only count those orbits which start in K, we obtain the entropy of F *in* K, denoted $h_F(K)$. More precisely, if we denote by $s(n, \epsilon, K)$ the cardinality of a maximal (n, ϵ)-separated subset of K, then

$$h_F(K) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \ln s(n, \epsilon, K)$$

and the entropy of F is defined by

$$h(F) = \sup\{h_F(K); K \subseteq X, K \text{ compact}\}.$$

Definition 3. If $F\colon X \to X$ is a homeomorphism, we define two points x and y to be 0-*entropy equivalent* if there is a continuum (that is, a compact connected set) K which contains both points and for which

$$h_F(K) = 0 = h_{F^{-1}}(K).$$

Remarks 4. a) That this indeed defines an equivalence relation follows from two facts: 1) $h_F(K \cup K') = \max\{h_F(K), h_F(K')\}$ and 2) the union of two continua containing a point in common is also a continuum.

b) Notice that if K is a proper subset of X, it is not necessarily the case that $h_F(K) = h_{F^{-1}}(K)$.

c) If F is not invertible we can consider the equivalence relation defined using only the first equality above. It would be interesting to understand this equivalence relation — and the ones mentioned below — for interval or, more generally, tree endomorphisms.

d) In general, we can consider the family of equivalence relations $\sim_\alpha$, indexed by a positive real α, declaring two points to be $\sim_\alpha$-equivalent if there is a continuum containing both and carrying entropy strictly smaller than α.

The 0-entropy equivalence relation is most interesting for two-dimensional systems.

Example 4. Let us describe its equivalence classes for the horseshoe map F_1. Denote by $\mathcal{H}^u$ and $\mathcal{H}^s$ the closures of the unstable and stable manifolds of the fixed point 0 (or indeed of any other periodic point, since their closures coincide) and let $\mathcal{H} = \mathcal{H}^s \cup \mathcal{H}^u$. Equivalence classes are of four kinds:

a) Closures of connected components of $\mathbb{S}^2 \setminus \mathcal{H}$.
b) Closures of connected components of $\mathcal{H}^u \setminus \mathcal{H}^s$ (not already contained in sets in a)).
c) Closures of connected components of $\mathcal{H}^s \setminus \mathcal{H}^u$ (not already contained in sets in a)).
d) Single points which are in none of the sets in a), b) or c).

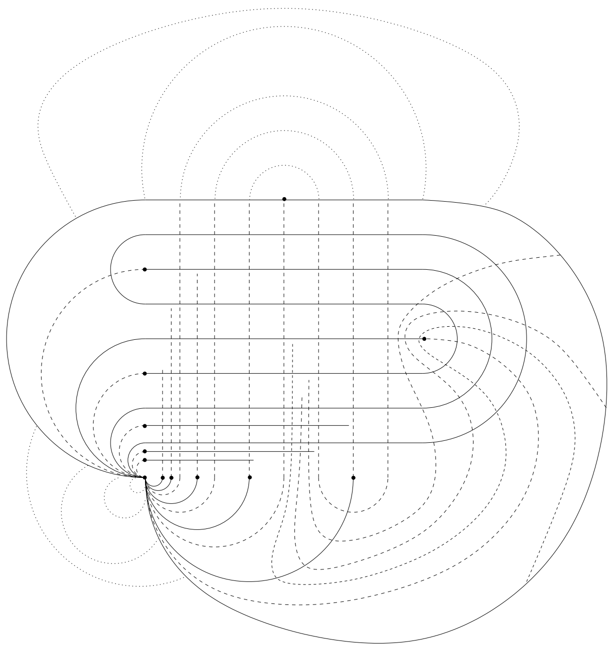

FIGURE 5. The quotient of the sphere under the 0-entropy equivalence relation for the horseshoe.

To see that these sets do not carry entropy, notice that all points in any connected component of $\mathbb{S}^2 \setminus \mathcal{H}$ (before taking the closure) converge to the attracting fixed point x. It is not hard to see that, after taking the closure nothing more exciting happens and this shows the sets in a) indeed carry no entropy. The same holds for sets of types b) and c). To see that any larger continuum must contain entropy, notice that if C is a connected set that contains two distinct sets among the ones described above, then it must intersect a Cantor set's worth of invariant manifolds, either stable or unstable (or both). It follows that one of its ω- or α-limit sets contains all the nonwandering set of the horseshoe and therefore one of $h_F(C)$ or $h_{F^{-1}}(C)$ equals $\ln 2$.

In [dCP02] it is shown that if F is a $C^{1+\epsilon}$ surface diffeomorphism, then the 0-entropy equivalence classes form an *upper semi-continuous monotone* decomposition of the surface. In particular, the quotient space is a *cactoidal* surface (roughly, a surface with nodes; see [Moo62, RS38]). Since the equivalence is dynamically defined, F projects to a homeomorphism $F/\sim$ on the quotient space. Moreover, any nontrivial continuum in the quotient space carries entropy of either the quotient homeomorphism or its inverse. The quotient map by the 0-entropy equivalence relation should be thought of as a 'tight' version of the original map in which all the wandering domains have been collapsed to points.

The quotient of the sphere by the 0-entropy horseshoe equivalence of Example 4 is shown in Figure 5. The quotient space is a sphere (we are collapsing everything outside the homoclinic tangle to a point), obtained by identifying the solid boundary in the figure along the dotted arcs from the mid-point at the top to the corner point on the lower left. The stable and unstable manifolds of the horseshoe project to two transverse foliations with singularities, represented by solid and dashed lines, respectively. In fact, these foliations carry transverse invariant measures whose product gives a measure on the sphere. The quotient map preserves both foliations, dividing one of the transverse measures by 2 and multiplying the other by 2, so that the product measure is invariant. This map is a *generalized pseudo-Anosov map*, which will be defined presently. We will also state a theorem that generalizes this construction to a class of thick graph maps.

3.2. **The Markov story and generalized pseudo-Anosov maps.** We start by introducing the basic concepts in the Perron-Frobenius theory for non-negative matrices (see [Bal00, Gan59]). Let M be a square matrix with non-negative entries. M is said to be *reducible* if, by a permutation of the index set, it is possible to put it in triangular block form:

$$M = \begin{bmatrix} A & 0 \\ B & C \end{bmatrix}.$$

Otherwise, M is said to be *irreducible.* The matrix M is said to be *irreducible and aperiodic* if there exists a positive integer k such that M^k is positive, that is, all its entries are positive. A non-negative irreducible matrix M has a unique positive eigenvector (up to scaling) and the associated eigenvalue λ — called the *Perron-Frobenius eigenvalue* of M — equals the spectral radius of M. If $\lambda = 1$ then M is a cyclic permutation matrix. Otherwise, $\lambda > 1$ and, for every i, j, there exists a power k such that the ij-entry of M^k is arbitrarily large (in fact, the entries of M grow like $(\mathrm{const})\times\lambda^k$). If M is irreducible then λ is a simple root of the characteristic polynomial of M and if M is also aperiodic, then λ is the only eigenvalue on the circle $\{z \in \mathbb{C}; |z| = \lambda\}$.

Definitions 4. A thick graph map $F\colon \mathbb{G} \to \mathbb{G}$ with quotient $f\colon G \to G$ will be called *Markov* if it satisfies the following additional conditions:

i) $\mathbb{G}$ has a finite number of strips and junctions (i.e., the graph G has finitely many edges). For each strip s we fix a homeomorphism $h_s\colon \overline{s} \to [0,1]\times[0,1]$ from the closure of the strip to the closed unit square, so that the decomposition elements of s are of the form ${h_s}^{-1}(\{x\}\times[0,1])$, for $0 < x < 1$.
ii) F is linear with respect to the structure homeomorphisms h_s, that is, 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 $\mu_{ij} < 1$ and expands horizontal coordinates uniformly by a factor $\lambda_{ij} \geq 1$.
iii) If J, J' are junctions such that $F(J) \subset J'$ then $F(\partial_{\mathbb{G}} J) \subset \partial_{\mathbb{G}} J'$.
iv) If J is a periodic junction of least period n, then J has an attracting periodic point of least period n in its interior whose basin contains $\mathrm{Int}(J)$.

We can associate a *transition matrix* $M = [m_{ij}]$ to a Markov thick graph map: letting $E = \{e_1, e_2, \ldots, e_n\}$ be the edges of G, set

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

Remarks 5. a) Notice that Markov thick graph maps can be made differentiable and we will assume, whenever we talk about them, that they are diffeomorphisms of the surface S. In particular, we will talk freely about stable and unstable manifolds of their periodic points.
b) Let $F\colon (S,\mathbb{G}) \to (S,\mathbb{G})$ be a Markov thick graph map whose transition matrix is irreducible and aperiodic, let $\Lambda = \bigcap_{n=0}^{\infty} F^n(\mathbb{G})$ and $p \in \mathbb{G}$ be any saddle periodic point of F. Then it is easy to see that $\Lambda = \overline{W^u(p)}$.

Definition 5. A surface homeomorphism $\Phi\colon S \to S$ is called a *generalized pseudo-Anosov map* if it satisfies the following conditions: there exist a pair $(\mathcal{F}^s, \mu^s)$, $(\mathcal{F}^u, \mu^u)$ of transverse measured foliations with singularities — either modeled on pronged singularities as in Figure 6 or accumulations of such, of which there are only finitely many — and a real number $\lambda > 1$ such that

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

Examples of generalized pseudo-Anosov maps include the torus Anosov maps and Thurston's pseudo-Anosov maps. We argue below that Markov thick graph maps also give rise to generalized pseudo-Anosovs. The definition above, however, probably includes many other maps and these Markov examples should form a dense set in the space of all generalized pseudo-Anosovs. A more thorough study of these issues, including a description of a uniform structure on the set

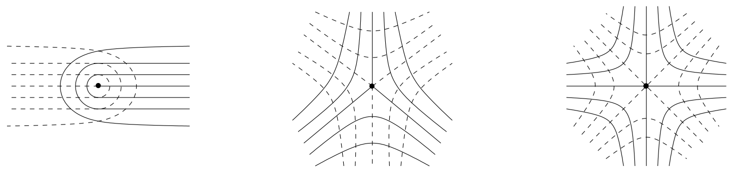

FIGURE 6. Pronged singularities of the invariant foliations.

of generalized pseudo-Anosovs, is currently under way and will hopefully appear in forthcoming papers.

Theorem 2. *Let $F\colon (S,\mathbb{G}) \to (S,\mathbb{G})$ be a Markov thick graph map with $\mathbb{G}$ of the homotopy type of S minus a point. Assume the associated transition matrix is irreducible and aperiodic. Then the quotient of S by the 0-entropy equivalence relation is homeomorphic to S and F projects to a generalized pseudo-Anosov homeomorphism $\Phi\colon S \to S$.*

A more detailed account of the 0-entropy equivalence relation and its quotients is given in [dCP02]. Below we will present an alternative construction of the generalized pseudo-Anosov quotient maps.

4. Generalized train tracks

In this section we will only deal with *finite* thick graphs, that is, thick graphs with finitely many strips and junctions. To simplify the exposition and the statements, we also assume that if $(S,\mathbb{G})$ is a thick graph then $\mathbb{G}$ has the homotopy type of the punctured surface $S \setminus \{p\}$, where $p \in S \setminus \mathbb{G}$. We refer to p as the *point at infinity* and denote it by ∞. In general, $\mathbb{G}$ has the homotopy type of the several times punctured surface $S \setminus \{p_1, \dots, p_k\}$. If this is the case, the discussion below has to be appropriately modified, but the ideas are essentially the same. The statements made here hold for the once punctured case.

4.1. **Definitions.** We start by fixing some notation and presenting some definitions. Suppose that $f\colon X \to X$ is a homeomorphism. Then f is said to be *supported* on a subset U of X if f is the identity on $X \setminus U$. A second homeomorphism $g\colon X \to X$ is *isotopic to* f if there is a continuous map $\psi\colon X \times [0,1] \to X$ such that each *slice map* $\psi_t\colon X \to X$ defined by $x \mapsto \psi(x,t)$ is a homeomorphism, and $\psi_0 = f$ and $\psi_1 = g$. The map ψ is called an *isotopy* from f to g. A *pseudo-isotopy* is a continuous map $\psi\colon X \times [0,1] \to X$ such that, for $0 \le t < 1$, the slice maps ψ_t are homeomorphisms onto their images. The isotopy or pseudo-isotopy is said to be *supported* on a subset U of X, denoted $\mathrm{supp}(\psi) = U$, if the homeomorphisms ψ_t are all equal on $X \setminus U$, and is said to be *relative* to U if $\psi_t(U) \subset \psi_0(U)$ for all $t \in [0,1]$. A map $g\colon X \to X$ is called a *near-homeomorphism* if it can be arbitrarily well approximated by homeomorphisms. Thus, if ψ is a pseudo-isotopy, the map $g = \psi_1$ is a near-homeomorphism.

Let $(S,\mathbb{G})$ be a finite thick graph, and $A \subset \mathbb{G}$ be a finite set, each of whose points lies in the interior of a junction and so that no junction contains more than one point of A. For each strip s of $\mathbb{G}$, let γ_s be an arc joining the two boundary components of s in $\mathbb{G}$ and intersecting each leaf of s exactly once. In terms of the structure homeomorphisms h_s introduced in Subsection 3.2, we can take $\gamma_s = h_s^{-1}([0,1] \times \{1/2\})$. Let $R \subseteq \mathbb{G}$ be the union of the arcs γ_s. The endpoints of the arcs γ_s are called *switches* and we denote by L the set of switches.

Definitions 6. Let $\tau \subseteq \mathbb{G} \setminus A$ be a graph with vertex set L and countably many edges, each of which intersects $\partial\mathbb{V}$ only at L, such that

i) $\tau \cap \mathbb{E} = R$, 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.

The isotopy class of τ by isotopies supported on $\mathbb{V} \setminus A$ (the set of junctions of $\mathbb{G}$ minus the points in A) is called a *generalized train track*[1] for $(\mathbb{G}, A)$. We will always refer to τ itself as the generalized train track, but it should be kept in mind that we do not distinguish between τ and τ' if it is possible to deform one to the other without crossing over points of A.

The edges of τ which are contained in $\mathbb{E}$ (that is, the connected components of R) are called *real*, and those which are contained in $\mathbb{V}$ are called *infinitesimal*. Write I for 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 *bubble* if its two endpoints coincide. An edge of τ is *homotopically trivial* if it is a bubble which bounds a disk containing no point of A, and is *homotopically non-trivial* otherwise.

Clearly a generalized train track τ for $(\mathbb{G}, A)$ is determined by its infinitesimal edges. It will sometimes be convenient to write $\tau(I)$ for the generalized train track whose set of infinitesimal edges is I, provided the thick graph and the set A are clear from the context.

A homotopy $\{\alpha_t\}$ of a path $\alpha\colon [0,1] \to X$ is said to be *relative to* $U \subset X$ if the points of $\alpha_t([0,1])$ that belong to U for $t = 0$ do not leave U for $0 \leq t \leq 1$. Let $[\alpha]$ be a homotopy class of paths in $S \setminus A$ relative to $\partial\mathbb{V}$, with endpoints in $\partial\mathbb{V}$. 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.

Remark 6. Although for our purposes a train track is a combinatorial object, we think of it as a smooth branched 1-submanifold of S. From this standpoint, the homotopy class of a path is carried by a generalized train track if there is a smooth representative in the class which is contained in the train track.

Now let $F\colon (S, \mathbb{G}, A) \to (S, \mathbb{G}, A)$ be a thick graph map such that $F(A) = A$, where $A = A_F$ is the set of attracting periodic orbits of F. On each strip s of $\mathbb{G}$ define the pseudo-isotopy $\psi_s\colon \bar{s} \times [0,1] \to \bar{s}$ to be given in coordinates by $\psi_s(x, y, t) = (x, (1-t)y + t/2)$ so that $\psi_s(\cdot, 1)$ maps $\bar{s}$ onto γ_s. Extend these pseudo-isotopies to a pseudo-isotopy $\psi_0\colon S \times [0,1] \to S$ in the following way: first extend the ψ_s to mutually disjoint disk neighborhoods $U_s \supset \bar{s}$, with $U_s \subset S \setminus A$, so that they are isotopies on $U_s \setminus \bar{s}$ and the identity on ∂U_s; then extend them to be the identity elsewhere.

Notice that, if τ is a generalized train track, then $\psi_0(F(\tau))$ satisfies the definition of train track, except possibly b), that is, it may have parallel edges.

Definitions 7. Define $F_*(\tau)$ to be the generalized train track consisting of a maximal subset of the edges of $\psi_0(F(\tau), 1)$ which contains no pair of parallel edges.

The train track $\tau = \tau(I)$ is said to be *F-invariant* if $F_*(\tau)$ is isotopic to τ in $S \setminus A$ and I is minimal (under inclusion) with this property.

Remark 7. Notice that $F_*(\tau)$ carries the homotopy class of $F(e)$ for each edge e of τ.

From the definition it follows that there exists a pseudo-isotopy ψ with the property that $\psi(F(\tau), 1) = \tau$.

Definition 8. The *train track map* $\phi\colon \tau \to \tau$ associated to $F\colon \mathbb{G} \to \mathbb{G}$ is defined by $\phi(\cdot) = \psi(F(\cdot), 1)$.

4.2. **Construction of invariant generalized train tracks.** Let $F\colon (S, \mathbb{G}, A) \to (S, \mathbb{G}, A)$ be as before. The following procedure constructs an F-invariant generalized train track τ.

Let $\tau_0 = R$, and for each $n \geq 0$ define $\tau_{n+1} = F_*(\tau_n)$. Since τ_0 is a subset of τ_1, each τ_n is naturally isotopic to a subset of τ_{n+1}: hence each τ_{n+1} can be adjusted, as it is constructed, by an isotopy relative to $A \cup \partial\mathbb{V}$ such that $\tau_n \subseteq \tau_{n+1}$. Define $\tau = \bigcup_{n \geq 0} \tau_n$. Then τ is a generalized train track by construction. It easy to see that it is F-invariant.

[1] To be painfully precise, we should talk about the isotopy class of the inclusion map $\iota\colon \tau \hookrightarrow \mathbb{G}$, but we won't do it.

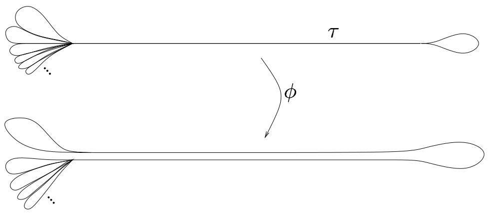

FIGURE 7. The invariant train track for the horseshoe.

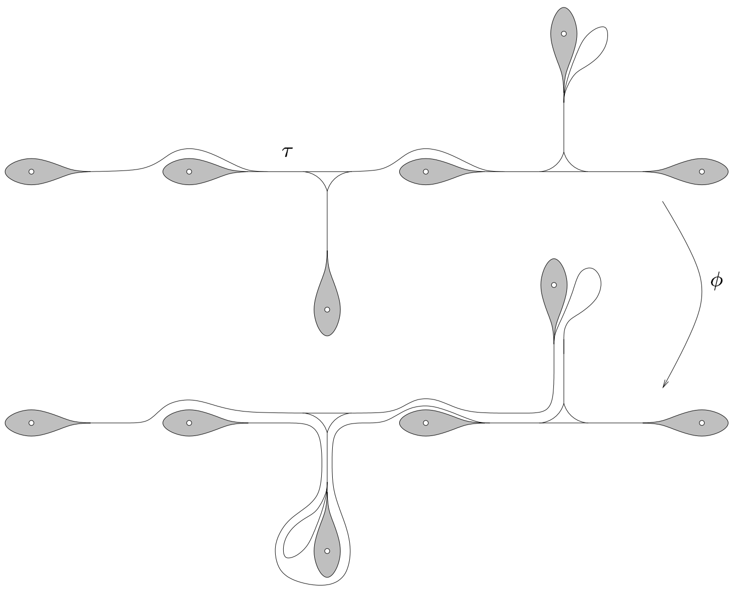

FIGURE 8. The invariant train track for the thick tree map F_2 of Example 3. Figure 9 shows a blow up of the shaded bubbles in this figure.

Remark 8. It is clear that this construction provides the minimal F-invariant generalized train track τ. Also, given any generalized sub-train track τ' of τ, it is clear that the same construction, starting with $\tau_0 = \tau'$, must generate $\tau = \bigcup_{n=0}^{\infty} F_*^n(\tau')$.

Example 5. The invariant train track for the horseshoe map F_1 is shown in Figure 7. 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.

Example 6. The invariant train track for the thick tree map F_2 of Example 3 is shown in Figures 8 and 9. Here, A consists of the union of the period 6 and period 2 orbits contained in the junctions of $\mathbb{T}$.

The invariant train track τ and the train track map ϕ should be thought of as more careful 1-dimensional representations of the thick graph $\mathbb{G}$ and the thick graph map $F\colon \mathbb{G} \to \mathbb{G}$: whereas $f\colon G \to G$ does not pay attention to junctions — they are collapsed to points — the map $\phi\colon \tau \to \tau$ gives a careful account of the behavior of the images of strips under iterates of F inside the junctions.

We now gather some useful facts that are straightforward consequences of the constructions above.

Remarks 9. a) ϕ is a near-homeomorphism, that is, it can be approximated arbitrarily well by homeomorphisms.

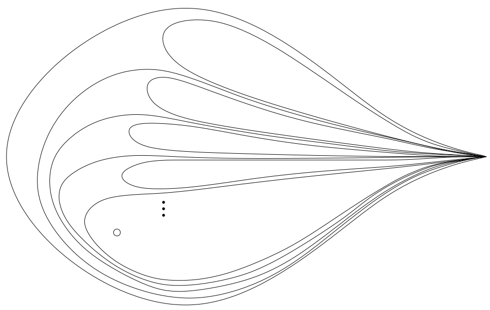

FIGURE 9. Detailed view of the bubbles in Figure 8. There and here $\circ$ represents a periodic point in the period 6 attracting orbit in Example 3.

b) ϕ maps I into itself and each infinitesimal edge is mapped homeomorphically onto another infinitesimal edge.
c) At most finitely many infinitesimal edges are mapped onto a given infinitesimal edge under ϕ.
d) If e is a real edge, then $\phi(e)$ can only intersect finitely many infinitesimal edges (except possibly at endpoints).

Definition 9. By an *infinitesimal polygon* we mean a component of the complement of τ bounded by finitely many infinitesimal edges. It is called an *n-gon* if it is bounded by n infinitesimal edges (see Figure 10).

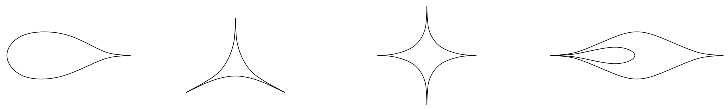

FIGURE 10. Examples of n-gons for $n = 1, 3, 4$ and a 1- and a 3-gon together.

Remarks 10. a) Bigons (2-gons) are not allowed unless they contain a point of A.
b) It is allowed that an n-gon has fewer than n vertices, as shown in the last diagram in Figure 10: two vertices in the 3-gon coincide.

The following lemma is an immediate consequence of the fact that ϕ is a near-homeomorphism on a surface.

Lemma 3. *For each integer $n \geq 1$, ϕ induces a 1-1 map on the collection of n-gons. If an n-gon contains a periodic point in its interior, then it is periodic under ϕ (with the same period). Otherwise, it belongs to a semi-infinite orbit of n-gons $\{\phi^n(\Delta); n \geq 0\}$, where Δ is an n-gon which is not the image under ϕ of any other n-gon. Two such orbits either coincide or are disjoint. Moreover, in this case, for all but finitely many n, $\phi^n(\Delta)$ has at most two vertices.*

5. Generalized pseudo-Anosov maps

In this section we describe the construction of a generalized pseudo-Anosov map using an invariant train track, as defined in the previous section. The construction follows closely the one in [BH95] for finite invariant train tracks and reduces to that in case the train track is finite. We also describe an associated complex structure on the surface with respect to which the map becomes a Teichmüller mapping. We continue to assume that thick graphs have the homotopy type of a once punctured surface.

5.1. **The construction.** Let $F\colon \mathbb{G} \to \mathbb{G}$ be a thick graph map and $\phi\colon \tau \to \tau$ its associated invariant train track map as defined in Section 4. As before, R and I denote the real and infinitesimal edges of τ. We number the real edges $\{e_i; 1 \leq i \leq n\}$ and the infinitesimal edges $\{e_i; i > n\}$ and define a (possibly infinite) transition matrix $M = [m_{ij}]$ setting, as above,

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

Since $\phi(I) \subset I$, M has block form

$$M = \begin{bmatrix} N_{n\times n} & 0 \\ B & \Pi \end{bmatrix}$$

The matrix N records transitions between real edges and Π records transitions between infinitesimal edges, whereas B records transitions from real to infinitesimal edges. The (possibly infinite) square matrix Π has only 0's and 1's in its entries. Each of its columns has exactly one non-zero entry and each row has at most finitely many non-zero entries. The matrix B (which has n columns and possibly infinitely many rows) has at most finitely many non-zero entries in each column. These observations follow from those in Remarks 9.

Example 7. For the horseshoe map, the transition matrix is infinite but is quite simple: $m_{11} = 2$, $m_{i\,i-1} = 1$ and $m_{ij} = 0$ otherwise.

Standing Assumption: It is assumed throughout this section that the matrix N is irreducible and aperiodic. This implies that its Perron-Frobenius eigenvalue $\lambda > 1$. It also follows that there is a positive integer k such that, for every real edge e, $\phi(e)$ contains all other real edges. In particular, τ is connected. This assumption is not necessary for all the results that follow, but it simplifies the discussion.

We think of M as an operator acting on the space l^1 of summable sequences of real numbers with norm $|y|_1 = \sum_{i\geq 1} |y_i|$. From the remarks above, it follows that M is a bounded operator with $\|M\|_1 \leq \max_j\{\sum_i |m_{ij}|\} < \infty$.

Let Y be an eigenvector of N associated to λ (and therefore unique, up to scale). It is possible to complete Y to an eigenvector $y = [Y\,Y']$ of M. Since the columns of Π have at most one non-zero entry which is 1, $\|\Pi\|_1 \leq 1$ and thus $\lambda I - \Pi$ is invertible. Setting $Y' = (\lambda I - \Pi)^{-1} BY$ we have

$$\begin{aligned} M\begin{bmatrix} Y \\ Y' \end{bmatrix} &= \begin{bmatrix} N & 0 \\ B & \Pi \end{bmatrix}\begin{bmatrix} Y \\ Y' \end{bmatrix} \\ &= \begin{bmatrix} NY \\ BY + \Pi Y' \end{bmatrix} \\ &= \begin{bmatrix} \lambda Y \\ \lambda Y' \end{bmatrix} \end{aligned}$$

In order to see that Y' is a positive vector, notice that

$$\begin{aligned} Y' &= \frac{1}{\lambda}(I + \frac{1}{\lambda}\Pi + \frac{1}{\lambda^2}\Pi^2 + \ldots)BY \\ &= \frac{1}{\lambda}(B + \frac{1}{\lambda}\Pi B + \frac{1}{\lambda^2}\Pi^2 B + \ldots)Y \end{aligned}$$

The matrices $\Pi^k B$ that appear above represent transitions from a real edge to an infinitesimal edge under the $(k+1)$-st iterate of the map ϕ. In fact, they represent exactly those transitions which occurred from a real to an infinitesimal edge in the first iterate and which then remained among infinitesimal edges for the next k iterates (the other ways to get from a real to an infinitesimal edge under the $(k+1)$-st iterate of ϕ are represented by matrices of the form $\Pi^{k-j-1}BN^j$). But every infinitesimal edge is the image, under some iterate of ϕ, of an infinitesimal edge of τ_1 and these are the intersection of $\phi(\tau_0)$ with $\mathbb{V}$, that is, they are transitions from real to infinitesimal edges under the first iterate of ϕ. This means that, for every $i \geq 1$, there exists $k \geq 1$ such that the i-th row of $\Pi^k B$ is non-zero. Since Y is a positive vector, it follows that every entry of Y' is non-zero.

Definition 10. A collection $\{y_i = y(e_i)\}_{i\geq 1}$ of non-negative real numbers, called *weights*, is said to satisfy the *switch conditions* if, for each switch q of τ, we have

$$y(e_{i_0}) = \sum y(e_i) + 2\sum y(e_j)$$

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

Lemma 4. *Let M be the transition matrix associated to $\phi\colon \tau \to \tau$, λ its Perron-Frobenius eigenvalue and $y = [Y\, Y'] = [y_1 y_2 \ldots]$ an eigenvector associated to λ as constructed above. Then the set of weights $\{y_i\}_{i\geq 1}$ satisfy the switch conditions.*

Proof (cf. [BH95]). Fix a large positive integer k. The equality $M^k y = \lambda^k y$ written in coordinates states that for each $i \geq 1$

$$y_i = \frac{1}{\lambda^k}\sum_j y_j \cdot (\text{number of times } \phi^k(e_j) \text{ intersects } e_i)$$

If $\phi^k(e_j)$ crosses a switch, it must cross edges on both sides except at its endpoints. Thus, the contribution to both sides of the switch is the same up to a bounded amount. Letting $k \to \infty$ yields the result. □

Now, let X be an eigenvector of N^T associated to the Perron-Frobenius eigenvalue λ. If we try to complete X to an eigenvector $[X\, X']$ for the adjoint M^*, we are forced to set $X' = 0$. The reason is that X' is a solution to the equation $\Pi^* X' = \lambda X' \Leftrightarrow (\lambda I - \Pi^*)X' = 0$. Since $\|\Pi^*\|_\infty \leq 1$, the only solution is $X' = 0$. This is why infinitesimal edges are called such.

We now give a description of the construction of the generalized pseudo-Anosov homeomorphisms corresponding to $\phi\colon \tau \to \tau$. As was mentioned before, it follows closely that presented in [BH95].

Let $x = (x_1, x_2, \ldots, x_n, 0, 0, \ldots)$ and $y = (y_1, y_2, \ldots)$ be eigenvectors of M^* and M, respectively, associated to the Perron-Frobenius eigenvalue λ as just described. To each real edge e_i, $1 \leq i \leq n$ of τ we associate a Euclidean rectangle R_i of dimensions $x_i \times y_i$ endowed with foliations by horizontal and vertical line segments. Each foliation has a transverse measure induced by Lebesgue measure. The horizontal and vertical foliations will be called *unstable* and *stable* respectively and, under the map, unstable leaves will be stretched and stable leaves will be contracted by the factors λ and $1/\lambda$, respectively. Place (homeomorphic copies of) these rectangles on S along the real edges of τ. The infinitesimal edges of τ are used to define an equivalence relation on the vertical sides of the rectangles, as follows. Let e_j be an infinitesimal edge and y_j the corresponding entry of the eigenvector y. Identify segments of length y_i along the vertical sides of the rectangles which contain the endpoints of e_j (note that these rectangles could be the same). The facts that the train track is a subset of S and that the switch conditions are satisfied imply that there is exactly one way in which these identifications can be made without self-intersections. The quotient space $\mathcal{R}$ of the rectangles under these identifications is a compact topological surface (with boundary) of the homotopy type of $S\setminus\{p\}$. It is foliated by unstable 'horizontal' lines, most of which are now infinite, that is, they are immersed copies of $\mathbb{R}$ (the exceptions being those that contain singularities of the foliations). Moreover, there is defined a map $\tilde{\Phi}\colon \mathcal{R} \to \mathcal{R}$ which stretches the unstable foliation by the factor λ, contracts the stable foliation (whose leaves are still finite segments) by the factor $1/\lambda$ and places $\tilde{\Phi}(\mathcal{R})$ inside $\mathcal{R}$ in the manner dictated by $\phi\colon \tau \to \tau$. Restricted to the interior of $\mathcal{R}$, $\tilde{\Phi}$ is a homeomorphism, but not along the boundary of $\mathcal{R}$. Notice that, by collapsing to points the stable segments, we obtain the graph G. Denoting the projection by $\tilde{\pi}\colon \mathcal{R} \to G$, $\tilde{\Phi}$ factors down to $f\colon G \to G$, that is, $f \circ \tilde{\pi} = \tilde{\pi} \circ \tilde{\Phi}$. The boundary $\partial\mathcal{R}$ has the structure of a smooth finite sided polygon, each side of which contains a periodic point. We identify segments

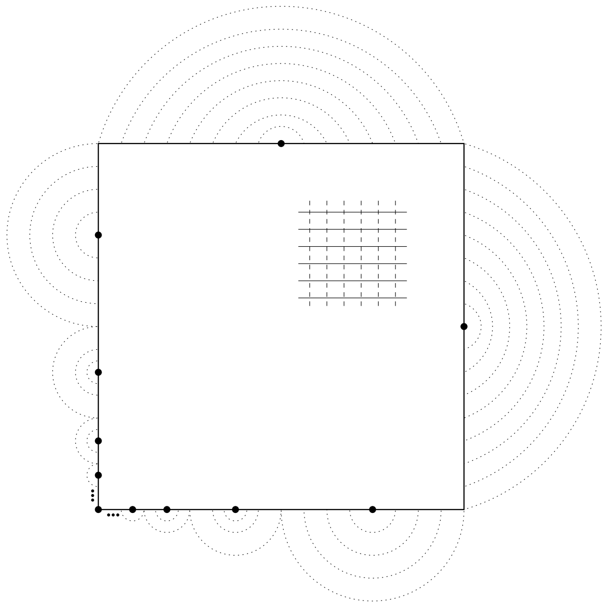

FIGURE 11. The identifications for the horseshoe.

of adjacent sides which are mapped to the same segment under some iterate of $\Phi\colon \mathcal{R} \to \mathcal{R}$. This usually leads to infinitely many identifications of smaller and smaller pieces of $\partial\mathcal{R}$ (for example, this is the case with the horseshoe, as will be seen below). The points of the periodic orbit on $\partial\mathcal{R}$ are all identified and become *the point at infinity.* Under these further identifications, the quotient of $\mathcal{R}$ is homeomorphic to S, the 'vertical' segments become a foliation of S by stable leaves, most of which are immersed copies of $\mathbb{R}$, and the induced map, denoted by $\Phi\colon S \to S$, now becomes a homeomorphism: this is the generalized pseudo-Anosov homeomorphism. The stable and unstable foliations are denoted by $\mathcal{F}^s$ and $\mathcal{F}^u$ respectively. Both are preserved by Φ, the leaves of the stable foliation being contracted and those of the unstable foliation being stretched by the factors $1/\lambda$ and λ, respectively.

Remarks 11. a) The foliations $\mathcal{F}^s$ and $\mathcal{F}^u$ are foliations with singularities: to each n-gon of τ there corresponds an n-pronged singularity of the foliations. There may be infinitely many such singularities, but they can accumulate on at most finitely many periodic orbits, which are in 1-1 correspondence with a subset of A. In the next section the orbits of 1-pronged singularities will be studied in greater detail.

b) The periodic orbits on $\partial\mathcal{R}$ described above have been found in different guises by other authors. See, for example, [BL98, NP73].

Example 8. The invariant train track for the horseshoe has only one real edge and therefore the surface ($\mathbb{S}^2$) will be constructed by identifying the edges of one rectangle. The identifications are shown in Figure 11 by dotted lines. Notice that the equivalence class of the lower left corner contains infinitely many points. The points marked with $\bullet$ lie on one orbit of 1-pronged singularities which is forward and backward asymptotic to the lower left corner, which is also the point at infinity. The quotient sphere with the two invariant foliations is shown in Figure 5.

Example 9. In Figure 12 are shown the identifications on seven rectangles dictated by the invariant train track for F_2. All the unstable identifications are indicated (dotted lines) whereas only the first set of stable identifications are shown (dashed-dotted lines). Further stable identifications are obtained from these by iterating the map backwards. The points of the period 2 orbit of infinitesimal triangles give rise to a period 2 orbit of 3-pronged singularities. There are infinitely

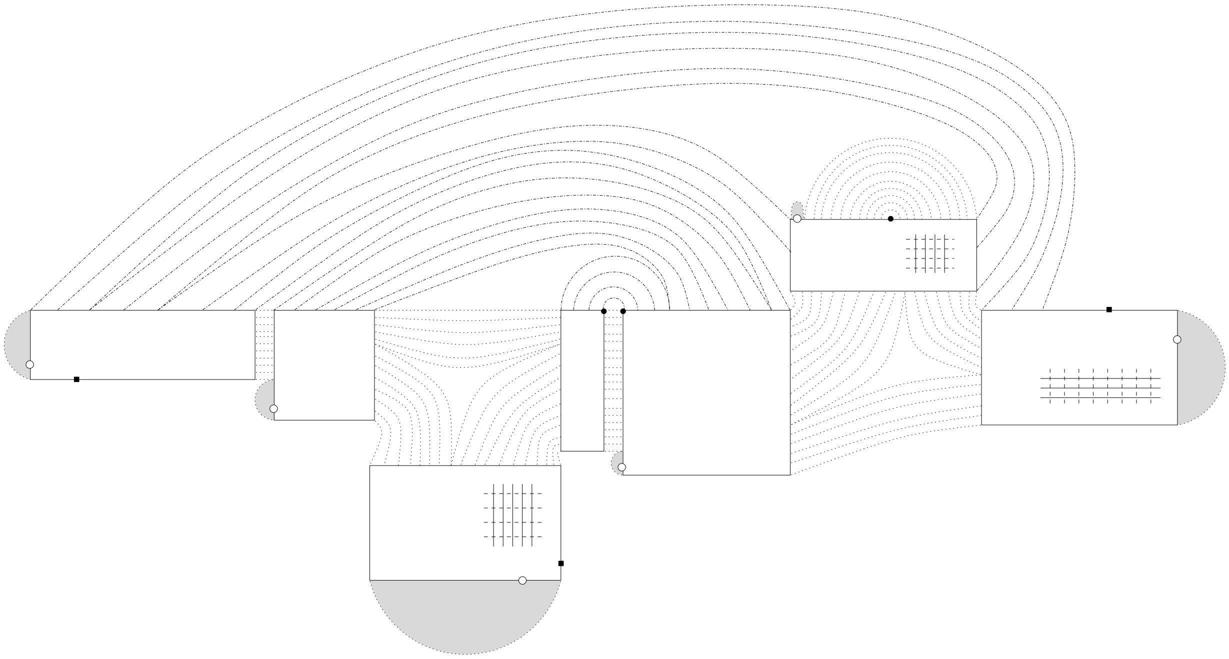

FIGURE 12. The identifications for F_2.

many 1-pronged singularities (two of which come from the points marked with $\bullet$ in the figure) converging in the future to the period 6 orbit (indicated by $\circ$) corresponding to the period 6 attracting orbit of F_2 and converging in the past to the period 3 orbit at infinity (indicated by $\blacksquare$). There is also an orbit of 3-pronged singularities with the same forward and backward fates as the orbit of 1-pronged singularities. A detailed view of the shaded regions in Figure 12 is shown in Figure 13.

5.2. **1-pronged singularities.** The generalized pseudo-Anosov maps constructed in the previous section preserve a pair of measured foliations with singularities. The 1-pronged singularities are the most important in many interrelated ways. Dynamically, they play a role analogous to that played by the critical points for endomorphisms of the interval. They also 'hold the map in place,' so to say, in the sense that isotopies relative to the set of 1-pronged singularities cannot destroy any dynamics (this has been proved for the case of finitely many singularities [Hal91, Han85, Thu88]; some infinities have been dealt with in [BH99]; the general case seems to be technically delicate).

We now describe how to find the orbits of 1-pronged singularities of the invariant foliations $\mathcal{F}^{u,s}$. They come from infinitesimal 1-gons of τ so we need to be able to determine the orbits of these. It follows from Lemma 3 that if a 1-gon contains a periodic point, it is itself periodic under ϕ and therefore corresponds to a periodic 1-pronged singularity of the invariant foliations. If a 1-gon does not belong to a periodic orbit, then there exists a *first* 1-gon whose orbit contains the given 1-gon, that is, there exists a 1-gon Δ, which is not the image of any 1-gon under ϕ and whose orbit contains the given 1-gon. In this case, there exists either a real edge e such that $\phi(e) \supset \Delta$ or there is an infinitesimal edge e', which is not a 1-gon, that is, whose boundary points are distinct switches, such that $\phi(e') = \Delta$. In either case, there must exist a real edge e, an arc $\gamma \subset e$ and a smallest integer $k \geq 1$ such that $\phi^k(\gamma) = \Delta$. If there are several such arcs (possibly contained in several distinct real edges), we use the ambient surface or thick graph to choose γ to be *innermost* among of them. By this we mean the following. To the arc γ there corresponds a *thick arc* $\Gamma = [a,b] \times [0,1] \subset s$ where s is a strip of $\mathbb{G}$. Because $\phi^k(\gamma) = \Delta$, $F^n(\Gamma)$ is contained in the junction J that contains Δ and $F^n(\{a,b\} \times [0,1])$ are contained in the same component of ∂J, that is, $F^n(\Gamma)$ 'makes a turn' inside

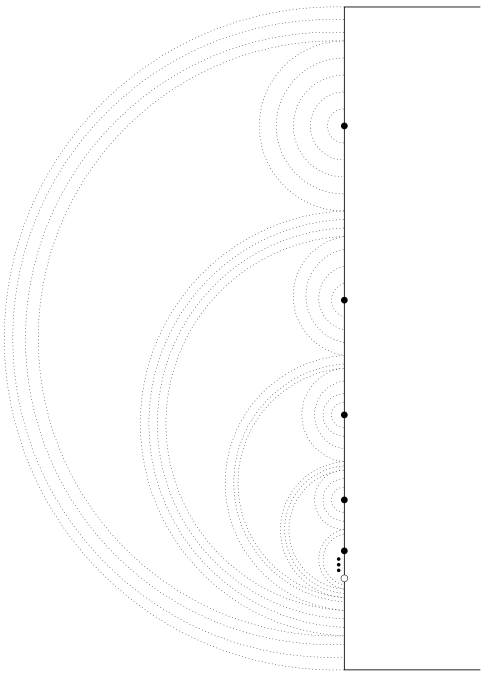

FIGURE 13. The identifications in the shaded regions in Figure 12.

J. We call γ innermost if $F^n(\Gamma)$ is an innermost turn among all thick arcs that map to J under F^n. Since infinitesimal edges are assigned 0 length in the construction of the invariant foliations, γ corresponds to a vertical segment in the rectangle R_e associated to the real edge e. Under the identifications required to make the leaves of the stable foliation infinite, R_e will be folded and one of the endpoints of this segment will become a 1-pronged singularity (the one that becomes innermost after folding). Since the 'horizontal' sides of rectangles are contained in the unstable manifold of the point at infinity, it follows that this 1-pronged singularity is backward asymptotic to the point at infinity. To summarize, we have proved the

Theorem 5. *Let $\mathcal{F}^{u,s}$ be the stable and unstable foliations of the generalized pseudo-Anosov map associated to a Markov thick graph map as constructed above. Then the 1-pronged singularities of $\mathcal{F}^{u,s}$ either belong to periodic orbits or are backward asymptotic to the point at infinity and forward asymptotic to one of finitely many periodic orbits.*

There is an easy way of tracking the orbits of 1-pronged singularities backwards using the graph map $f\colon G \to G$. Suppose $p \in G$ is a critical point of f (that is, a point at which f is not a local homeomorphism), so that $v = f(p)$ is a vertex. Assume, without loss of generality, that $f(p)$ is innermost in the junction $V = \pi^{-1}(v)$. Choose $p_{-1} \in f^{-1}(p)$ so that $f^2(p_{-1})$ is innermost in V, then choose $p_{-2} \in f^{-1}(p_{-1})$ so that $f^3(p_{-2})$ is innermost in V and proceed like this. If we reach a point p_{-n} which is not a vertex, then all subsequent ones are not vertices (since vertices map to vertices) and, in fact, at each subsequent step, there is exactly one innermost preimage (since the graph map is assumed to be onto). In this case, p_{-n} converges to the orbit at infinity. Otherwise, there are two possibilities: either we eventually return to p or we get to a point which has *no* innermost preimage. In the first case, we have found a periodic orbit of 1-pronged singularities and in the second, we were not following an orbit of 1-pronged singularities. Below we give examples explaining all three possibilities.

Examples 10. a) Consider the thick tree map F_2. In Figure 14, the quotient tree map is drawn.

To avoid having too many symbols in the figure, the images of points in T under f are denoted

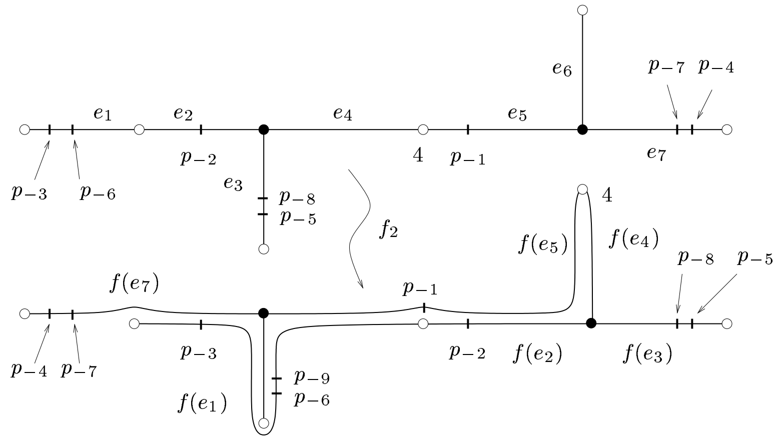

FIGURE 14. The backward orbit of 1-pronged singularities for F_2.

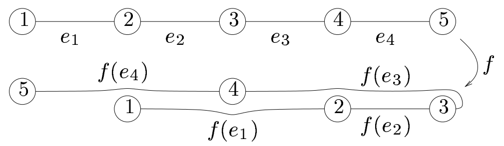

FIGURE 15. The quotient of a thick tree map whose associated generalized pseudo-Anosov has a periodic orbit of 1-pronged singularities.

by the same symbol on the image tree $f(T)$ drawn below. The only critical point is $p = 4$ and $v = f(4) = 5$. The preimage $f^{-1}(4)$ has two points: 2 and a point in edge e_5. It is the latter that becomes innermost under f^2, so $p_{-1} \in e_5$. The preimage $f^{-1}(p_{-1})$ again consists of two points, of which the one that becomes innermost under f^3 lies on edge e_2: this is p_{-2}. Again $f^{-1}(p_{-2})$ consists of two points and $p_{-3} \in e_1$. From here on things repeat, as shown in Figure 14: $p_{-4} \in e_7$, $p_{-5} \in e_3$, $p_{-6} \in e_1$, $p_{-7} \in e_7$, etc. The sequence p_n converges to the period 3 orbit at infinity.

b) In this example (see Figure 15, where only the induced tree endomorphism is drawn) $p = 3$ and $v = 5$. The preimage $f^{-1}(p)$ has two points, namely, 4 and a point on edge e_1. Under f^2 it is point 4 that becomes innermost so $p_{-1} = 4$. Again, $f^{-1}(4)$ contains two points, namely, 2 and a point on edge e_3, but it is 2 that becomes innermost under f^3 so that $p_{-2} = 2$. Continuing like this, we get $p_{-3} = 1$, $p_{-4} = 5$ and $p_{-5} = 3$. Thus, p lies gives rise to a periodic orbit of 1-pronged singularities. The associated generalized pseudo-Anosov is, in fact, a pseudo-Anosov map in the sense of Thurston.

c) The third possibility is shown in Figure 16. The top part of the figure contains parts of the graph containing the points p and $v = f(p)$ and two points in $f^{-1}(p)$; the bottom contains their images under f. Because no arc has image entering p through edge a and exiting through edge b (indicated by the dashed line in the figure), $f^{-1}(p)$ does not contain a point whose image under f^2 is innermost in v.

5.3. **The complex structure.** We now describe a natural complex structure associated to a generalized pseudo-Anosov map, with respect to which it becomes a Teichmüller mapping and the invariant foliations become the horizontal and vertical trajectories of an associated integrable quadratic differential.

As was seen above, S is obtained as the quotient of Euclidean rectangles under side identifications by Euclidean isometries. In the interior of the rectangles, the complex structure is the one

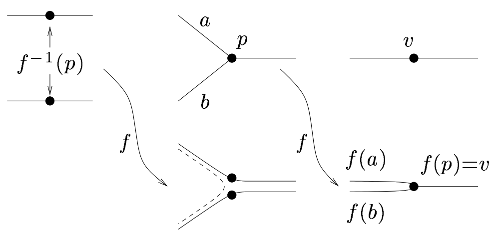

FIGURE 16. When a critical point of a tree map does not produce an orbit of 1-pronged singularities.

determined by the Euclidean structure. At k-pronged singularities, use the maps $z \mapsto z^{2/k}$ to define coordinate charts. This produces a complex structure at all points of S, except at the accumulations of singularities. These are topologically isolated and it is necessary to decide whether the complex structure in the complement regards them as punctures or as holes. We now argue that they are in fact punctures so that the complex structure extends across them uniquely. To prove this we present a sequence of concentric annuli converging to an accumulation of singularities and whose moduli add up to infinity. By length-area arguments (see [Ahl73, LV73]), the result follows[2].

The basic inequality used is

$$\mathrm{mod}(A) \geq \frac{\inf_{\gamma\in\Gamma} l(\gamma)^2}{\mathrm{Area}(A)}$$

where Γ is the set of all rectifiable curves joining the boundary components of the annular region A and $l(\gamma)$ is the length of the curve γ.

We first give the argument for the accumulation of singularities p in the generalized pseudo-Anosov associated to the horseshoe.

In Figure 17 are shown the first four concentric annular regions of a sequence $\{A_n\}_{n=1}^{\infty}$ converging to p. We argue that $\mathrm{mod}(A_n) > C/n$, for some positive constant C, implying thus that $\sum_{n=1}^{\infty} \mathrm{mod}(A_n)$ diverges. Annular region A_n is made up of three round quarter-annuli at the upper-left and right and lower-right corners, the region bounded by four circular arcs on the lower-left and $2n$ round half-annuli, n along each of the left and bottom sides of the square. All parts being (parts of) Euclidean annuli, they contribute to the extremal length of the family of all curves joining the boundary components of A_n at least like Euclidean annuli do. The infimum of the lengths of curves joining boundary components is $\propto 1/2^n$. The sum of the areas of the quarter annuli is $\propto 1/2^{2n}$, whereas the $2n$ half-annuli contribute $\propto n/2^{2n}$ to the total area. Therefore, we have

$$\begin{aligned} \mathrm{mod}(A_n) &\geq \frac{(\text{Distance between boundary components of } A_n)^2}{\mathrm{Area}(A_n)} \\ &\geq \frac{C_1}{C_2 + C_3 n} \end{aligned}$$

The general result is

Theorem 6. *Let $F\colon (S,\mathbb{G}) \to (S,\mathbb{G})$ be a Markov thick graph map whose associated transition matrix is irreducible and aperiodic. Let $\Phi\colon S \to S$ be the associated generalized pseudo-Anosov and $\{p_1,\ldots,p_k\}$ the accumulation points of pronged singularities of the Φ-invariant foliations. Then the complex structure on $S \setminus \{p_1,\ldots,p_k\}$ induced by the Euclidean structure extends uniquely to a complex structure on the compact surface S.*

Sketch of Proof. Let p be an accumulation of pronged singularities of the invariant foliations corresponding to an attracting periodic orbit of F, which we will also denote by p (the other

[2] I am grateful to Fred Gardiner for suggesting this argument (see also [EG97]).

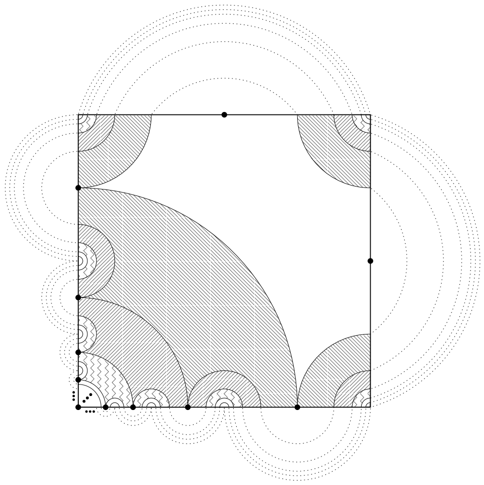

FIGURE 17. A sequence of concentric annular regions whose sum of moduli diverges.

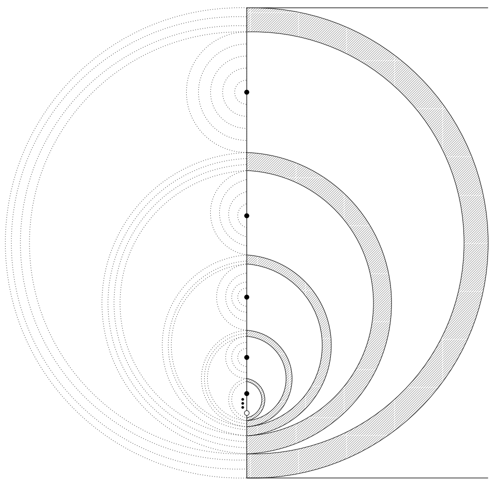

FIGURE 18. Another sequence of concentric annular regions whose sum of moduli diverges.

possibility, namely, that p is the point at infinity, is treated similarly). Passing to an appropriate iterate, we may assume that p is an attracting fixed point of F. For simplicity, we assume it lies in a 1-junction V; the other cases are analogous.

There are two possibilities to consider: either the invariant train track τ has a bubble in V which encloses p or not. If there is a bubble β enclosing p, all its iterates also do. They then give rise to an infinite collection of disjoint concentric annular regions A_n, all enclosing p, with $\mathrm{mod}(A_n) \propto 1$ (see Figure 18). In case there is no bubble enclosing p in V, then p is the point at infinity and the situation is analogous to the horseshoe example explained above. □

REFERENCES

[Ahl73] L. V. Ahlfors, *Conformal invariants: topics in geometric function theory*, McGraw-Hill Book Co., New York, 1973, McGraw-Hill Series in Higher Mathematics. MR 50 #10211

[Bal00] V. Baladi, *Positive transfer operators and decay of correlations*, World Scientific Publishing Co. Inc., River Edge, NJ, 2000. MR 1 793 194

[Bar86] M. Barge, *Horseshoe maps and inverse limits*, Pacific J. Math. **121** (1986), no. 1, 29–39. MR **87g:**58077

[BH95] M. Bestvina and M. Handel, *Train-tracks for surface homeomorphisms*, Topology **34** (1995), no. 1, 109–140. MR **96d:**57014

[BH99] P. Boyland and T. Hall, *Isotopy stable dynamics relative to compact invariant sets*, Proc. London Math. Soc. (3) **79** (1999), no. 3, 673–693. MR **2000k:**37026

[BL98] C. Bonatti and R. Langevin, *Difféomorphismes de Smale des surfaces*, Astérisque (1998), no. 250, viii+235, With the collaboration of E. Jeandenans. MR **99m:**58123

[Bow71] R. Bowen, *Entropy for group endomorphisms and homogeneous spaces*, Trans. Amer. Math. Soc. **153** (1971), 401–414. MR 43 #469

[dCH01] A. de Carvalho and T. Hall, *Pruning theory and Thurston's classification of surface homeomorphisms*, J. Eur. Math. Soc. (JEMS) **3** (2001), no. 4, 287–333. MR 1 866 160

[dCP02] A. de Carvalho and M. Paternain, *Monotone quotients of surface diffeomorphisms*, 2002, Submitted.

[EG97] C. J. Earle and F. P. Gardiner, *Teichmüller disks and Veech's $\mathcal{F}$-structures*, Extremal Riemann surfaces (San Francisco, CA, 1995), Amer. Math. Soc., Providence, RI, 1997, pp. 165–189. MR **97k:**32031

[FM93] J. Franks and M. Misiurewicz, *Cycles for disk homeomorphisms and thick trees*, Nielsen theory and dynamical systems (South Hadley, MA, 1992), Amer. Math. Soc., Providence, RI, 1993, pp. 69–139. MR **95e:**58133

[Gan59] F. R. Gantmacher, *The theory of matrices. Vols. 1, 2*, Chelsea Publishing Co., New York, 1959. MR 21 #6372c

[Hal91] T. Hall, *Unremovable periodic orbits of homeomorphisms*, Math. Proc. Cambridge Philos. Soc. **110** (1991), no. 3, 523–531. MR **92i:**58146

[Han85] M. Handel, *Global shadowing of pseudo-Anosov homeomorphisms*, Ergodic Theory Dynam. Systems **5** (1985), no. 3, 373–377. MR **87e:**58172

[LV73] O. Lehto and K. I. Virtanen, *Quasiconformal mappings in the plane*, second ed., Springer-Verlag, New York, 1973, Translated from the German by K. W. Lucas, Die Grundlehren der mathematischen Wissenschaften, Band 126. MR 49 #9202

[Moo62] R. Moore, *Foundations of point set theory*, American Mathematical Society, Providence, R.I., 1962. MR 27 #709

[NP73] S. Newhouse and J. Palis, *Hyperbolic nonwandering sets on two-dimensional manifolds*, Dynamical systems (Proc. Sympos., Univ. Bahia, Salvador, 1971), Academic Press, New York, 1973, pp. 293–301. MR 49 #4044

[RS38] J. Roberts and N. Steenrod, *Monotone transformations of two-dimensional manifolds*, Ann. of Math. **39** (1938), 851–862.

[Thu88] W. P. Thurston, *On the geometry and dynamics of diffeomorphisms of surfaces*, Bull. Amer. Math. Soc. (N.S.) **19** (1988), no. 2, 417–431. MR **89k:**57023

Institute for Mathematical Sciences, State University of New York at Stony Brook, NY 11794-3660, USA

E-mail address: andre@math.sunysb.edu