Expanding Factors for Pseudo-Anosov Homeomorphisms

E. Rykken

1. Introduction and Definitions

Thurston classified homeomorphisms on compact surfaces up to isotopy (see [3; 5]). He showed that any homeomorphism on a compact surface may be decomposed into simpler homeomorphisms on simpler compact surfaces. These simpler homeomorphisms are either periodic or pseudo-Anosov. Here we study the dynamics of the pseudo-Anosov homeomorphisms, because they are much more complicated and much richer than those of the periodic ones. In addition, pseudo-Anosov homeomorphisms on compact surfaces can be thought of as a natural extension of the study of hyperbolic toral automorphisms on the 2-dimensional torus. Using the Markov matrix, Markov partitions of these maps allow us to make a natural association with symbolic dynamics.

In the first section, we recall the basic definitions and background theorems. The second section provides several examples of pseudo-Anosov homeomorphisms on the two-dimensional sphere. In the final section, using tools from algebraic topology, we prove the following theorem, which extends a theorem concerning hyperbolic toral automorphisms on $\mathbb{T}^2$ [14].

Theorem 3.3. *Let $f: M \to M$ be a pseudo-Anosov homeomorphism on an orientable surface of genus g with oriented unstable manifolds. Let $\mathcal{P}$ be a Markov partition for f with Markov matrix $\mathcal{A}$. If f preserves the orientation of unstable manifolds, then the eigenvalues of $f_{*1}: H_1(M;\mathbb{R}) \to H_1(M;\mathbb{R})$ are the same as those of $\mathcal{A}$ including multiplicity, with the possible exception of some zeros and roots of unity.*

Hence, the expanding factor λ is an eigenvalue of $f_{*1}: H_1(M;\mathbb{R}) \to H_1(M;\mathbb{R})$. Similarly, if f reverses the orientation of unstable manifolds, then the eigenvalues of $f_{*1}: H_1(M;\mathbb{R}) \to H_1(M;\mathbb{R})$ are the same as those of $-\mathcal{A}$ including multiplicity, with the possible exception of some zeros and roots of unity. Hence $-\lambda$ is an eigenvalue of $f_{*1}: H_1(M;\mathbb{R}) \to H_1(M;\mathbb{R})$. As a consequence of this theorem, we have the following corollary.

Corollary 3.8. *Let λ be the expanding factor of a pseudo-Anosov homeomorphism f. If λ is the root of an irreducible quadratic equation over the rationals, then λ satisfies a quadratic of the form $x^2 + nx \pm 1$, where $n \in \mathbb{Z}$ and $|\lambda| \neq 1$.*

Received May 18, 1998. Revision received January 8, 1999.
Michigan Math. J. 46 (1999).

This is the same sort of quadratic equation that must be satisfied by the expanding factor of a two-dimensional hyperbolic automorphism.

DEFINITION 1.1. Let M be a compact two-dimensional orientable manifold with a flat Euclidean structure with a finite set S of isolated cone singularities [10]. A *pseudo-Anosov homeomorphism* on M is a homeomorphism $f: M \to M$ such that there exist two perpendicular foliations $W^u(f)$ and $W^s(f)$ and a real number $\lambda > 1$ with the property that the image under f of a leaf is a leaf and:

(1) $d(f(x), f(y)) = \lambda d(x, y)$ if x, y are in the same leaf of $W^u(f)$;
(2) $d(f(x), f(y)) = \frac{1}{\lambda} d(x, y)$ if x, y are in the same leaf of $W^s(f)$.

We call $W^u(f)$ the *unstable foliation*, $W^s(f)$ the *stable foliation*, and λ the *expanding factor* of f.

A pseudo-Anosov homeomorphism f is hyperbolic and locally affine. That is, in local coordinates (which are compatible with the flat Euclidean structure), the map f is affine except at points of S. Furthermore, $f(S) = S$ and a singularity of angle $k\pi$ maps to a singularity of angle $k\pi$.

DEFINITION 1.2. A *rectangle* R in M is the image of a closed Euclidean rectangle $\tilde{R}$ under a continuous map π such that π is a one-to-one Euclidean map on the interior of $\tilde{R}$.

DEFINITION 1.3. Let $W^u(x, f)$ be the leaf of the unstable foliation that contains x. Let $W^s(x, f)$ be the leaf of the stable foliation that contains x. For $x \in \operatorname{int} R$, let $W^u(x, f, R)$ be the component of $W^u(x, f) \cap R$ that contains x and define $W^s(x, f, R)$ similarly. Let the width of a rectangle R_i be given by $r_i =$ length of $W^s(x, f, R_i)$, where $x \in R_i$. Let the length of a rectangle R_i be given by $l_i =$ length of $W^u(x, f, R_i)$, where $x \in R_i$.

DEFINITION 1.4. Suppose f is a pseudo-Anosov homeomorphism on M. A *Markov partition* for f on M is a finite covering of M by rectangles, $\{R_1, R_2, \ldots, R_n\}$, such that:

(1) for $i \neq j$, $\operatorname{int} R_i \cap \operatorname{int} R_j = \emptyset$;
(2) if $x \in \operatorname{int} R_i$ and $f(x) \in \operatorname{int} R_j$, then $f(W^u(x, f, R_i)) \supset W^u(f(x), f, R_j)$ and $f(W^s(x, f, R_i)) \subset W^s(f(x), f, R_j)$.

For an example of a Markov partition for a hyperbolic toral automorphism on $\mathbb{T}^2$, see [12, Sec. 8.5.1].

DEFINITION 1.5. We define the *Markov matrix* for a Markov partition $\mathcal{P}$ with n rectangles to be the $n \times n$ matrix given by

$$M_{ij} = \#[\operatorname{int} f(W^u(x_j, f, R_j)) \cap \operatorname{int} W^s(x_i, f, R_i)],$$

where $x_i \in \operatorname{int} R_i$ and $x_j \in \operatorname{int} R_j$ for $1 \leq i, j \leq n$ and where we have used “#” to denote “the number of components of”.

The matrix does not depend on the choice of x_i or x_j, and $M_{i,j}$ should be thought of as the number of times that int $f(R_j)$ crosses int R_i.

THEOREM 1.6. *Let $f: M \to M$ be a pseudo-Anosov homeomorphism. Then f has a Markov partition.*

Proof. See the proof by Fathi and Shub [5]. □

Let $\mathcal{M}$ be the Markov matrix for a Markov partition of a pseudo-Anosov homeomorphism. Since $\mathcal{M} \geq 0$ is mixing, by the Perron–Frobenius theorem (see [9]) we have that $\mathcal{M}$ has a unique, real, and positive largest eigenvalue that exceeds the moduli of all the other eigenvalues. To this maximal eigenvalue there corresponds an eigenvector with positive coordinates. Moreover, no irreducible and nonnegative matrix can have two linearly independent nonnegative eigenvectors. Thus, if we can find an eigenvalue of $\mathcal{M}$ with positive eigenvector, then it must be the unique, real, and positive largest eigenvalue of $\mathcal{M}$. The following result is well known and an easy computation.

THEOREM 1.7. *Let f be a pseudo-Anosov homeomorphism with expanding factor λ, and let $\mathcal{P}$ be a Markov partition for f with n rectangles with Markov matrix $\mathcal{M}$. Let $r = \{r_1, r_2, \ldots, r_n\}$, where r_i is the width of the rectangle R_i, and let $l = \{l_1, l_2, \ldots, l_n\}$, where l_i is the length of the rectangle R_i. Then r is a right eigenvector for $\mathcal{M}$ with eigenvalue λ, and l is a left eigenvector for $\mathcal{M}$ with eigenvalue λ. Hence λ is the unique, real, and positive largest eigenvalue of $\mathcal{M}$.*

We can consider the *index for singularities of a line field or foliation*. For a foliation, pick any line field along the foliation. Here we consider a disk around the singularity and consider a vector in the direction of the line field. We then calculate the winding number as the vector travels counterclockwise around the boundary circle. If the line field has an even number of prongs at the singularity, then it is orientable and the index is an integer. If there are an odd number of prongs at the singularity, then the line field is not orientable and the index is a fraction. In general, if there are k prongs, then $I_p = (2 - k)/2$ (see Figure 1). This is analogous to the fixed point index for a vector field.

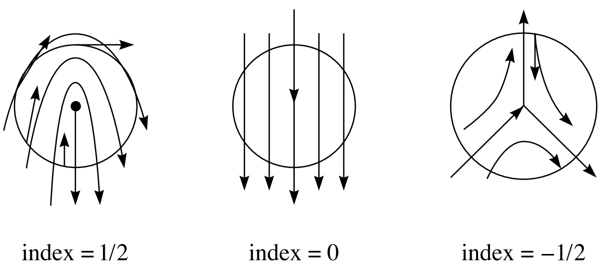

Figure 1 Index for singularities of a line field or foliation.

THEOREM 1.8.

(1) (Poincare–Hopf) *Suppose X is a smooth vector field with isolated zeros on a compact manifold M (if M has a boundary, then assume X points outward at all boundary points). Then $\sum I_p = \chi(M)$.*

(2) *Suppose M is a compact 2-dimensional manifold with a flat Euclidean structure having isolated cone singularities. Assume that there is an Euclidean foliation on M. Let $\tilde{M}$ be the branched double cover of M that orients the line field. Then $\chi(\tilde{M}) = 2\chi(M) - n$, where n is the number of singularities with an odd number of prongs.*

(3) *Suppose M is a compact 2-dimensional manifold with a flat Euclidean structure having isolated cone singularities. Assume that there is an Euclidean foliation on M. Then $\sum I_p = \chi(M)$.*

Proof.

(1) For a proof, see [11, pp. 35–41].

(2) See [4].

(3) The proof is straightforward and follows from the Poincare–Hopf theorem by using the branched double cover of the manifold that orients the line field. □

Recall that $\chi(S^2) = 2$ and that, if M is an orientable surface of genus g, then $\chi(M) = 2 - 2g$. The equality $\sum I_p = \chi(M)$ places restrictions on the combinations of singularities that can exist on a given manifold. For example, if the manifold is S^2 then there must be at least four singularities with angle π.

2. Examples of Pseudo-Anosov Homeomorphisms

We will consider examples of orientation-preserving pseudo-Anosov homeomorphisms that lie on S^2. Since orientation-preserving homeomorphisms of S^2 are homotopic to the identity, they have at least one fixed point. The group of isotopy classes of homeomorphisms of S^2 with one distinguished fixed point that permutes a specified set of n other points is isomorphic to the braid group on n symbols [1]. The braid group on n symbols is generated by permutations that permute two elements at a time. Consider a tree with n vertices and $n - 1$ edges embedded in a disk. The isotopy class of a map on this disk that permutes the vertices and pointwise fixes the boundary can be thought of as an element of the braid group. If the image of the tree lies in a small neighborhood of the tree, then we can also think of the map as a map of the tree. We will call this a map on a *thick tree structure.* A paper by Franks and Misiurewicz [8] explains how to generate pseudo-Anosov homeomorphisms by considering certain irreducible maps on thick tree structures.

To generate examples, given two vertices on the tree that are joined by an edge, we define the following map. Consider a disk containing the two vertices that lies in a small neighborhood of the tree. Consider a slightly larger disk which contains the first disk but which also lies in a small neighborhood of the tree. On the inner disk, rotate by 180° in either the clockwise or counterclockwise direction. Fix the outer disk and, on the annulus between them, interpolate in a continuous way. We will call this a *flip* (in either a clockwise or counterclockwise direction). Let σ_{AB} denote a clockwise flip on vertices A and B, and let σ_{AB}^{-1} denote the counterclockwise flip.

EXAMPLE 1. Consider a tree, with two edges (1 and 2) and three vertices (A, B, C), embedded in a disk; see Figure 2. We will generate a map on this disk by $\sigma_{AC}\sigma_{AB}^{-1}$, that is, by first flipping vertices A and B in a counterclockwise direction and then flipping vertices A and C in a clockwise direction. This gives us the picture shown in Figure 2 of a map on a tree embedded in a disk.

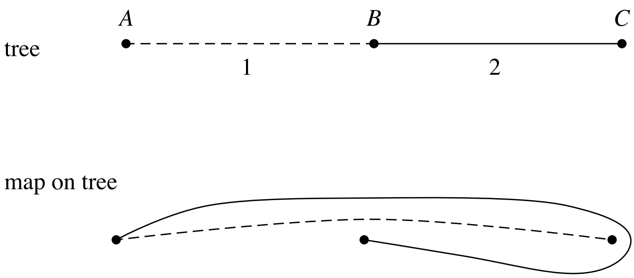

Figure 2 Map on tree for Example 1.

Because the image of the tree lies in a small neighborhood of the tree, we think of this map as a map of the tree. We have a Markov partition of this map on the tree, with each edge being a rectangle. The resulting Markov matrix for this partition is

$$M = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

In general, given such a map on a tree and its corresponding Markov matrix, we will use the following procedure to obtain a pseudo-Anosov homeomorphism on S^2. Consider the left eigenvector $(l_1, \ldots, l_n)$ and the right eigenvector $(r_1, \ldots, r_n)$ for the Perron–Frobenius eigenvalue λ of the Markov matrix. Construct n rectangles, $\{R_i\}_{i=1}^n$, with R_i having dimensions $l_i \times r_i$; then replace the edges in our tree with the rectangles. With identifications on the boundaries of the rectangles, this will form a disk.

If $|\lambda| > 1$, the map on the disk will stretch a rectangle in the horizontal or unstable direction by a factor of λ and shrink the rectangle in the vertical or stable direction by a factor of $1/\lambda$. Thus λ is called the *expanding factor*. In this example, the expanding factor is $\lambda = (3 + \sqrt{5})/2$.

The image rectangle is placed in the disk according to the map on the thick tree structure. Notice that since (l_1, l_2) is a left eigenvector, we have $\lambda l_1 = l_1 + l_2$ and $\lambda l_2 = l_1 + 2l_2$. But these are exactly what is needed for the image of R_1 to cross

R_1 and R_2 each exactly once and for the image of R_2 to cross R_1 exactly once and R_2 exactly twice. Since (r_1, r_2) is a right eigenvector, we have $r_1 + r_2 = \lambda r_1$ and $r_1 + 2r_2 = \lambda r_2$. Again, this is exactly what is needed for the images of the rectangles to fit in the horizontal direction.

We identify the edges of the rectangles that correspond to the r_i either with parts of themselves or with edges of other rectangles. For example, if a vertex is at the end of the tree then there is a single edge incident to it. Hence, there is a single rectangle incident to that vertex, and we give identification to the side of the rectangle incident to it with length r_i by folding it at its center. We make all identifications in such a fashion that the images of the rectangles will cross the rectangles in the same order as the images of the edges wrap around the edges on the thick tree. This procedure may not always work.

Finally, we identify points on the boundary by "sewing" it up, preserving arc-lengths starting at the cusps. This will produce a map on S^2. In general, if there are k cusps then there will be a fixed point on the boundary with angle $k\pi$. In this case there is only one cusp, so there will be a fixed point on the boundary that occurs at a singularity with angle π. In this example, there are exactly four singularities with angle π and none with angle more than 2π.

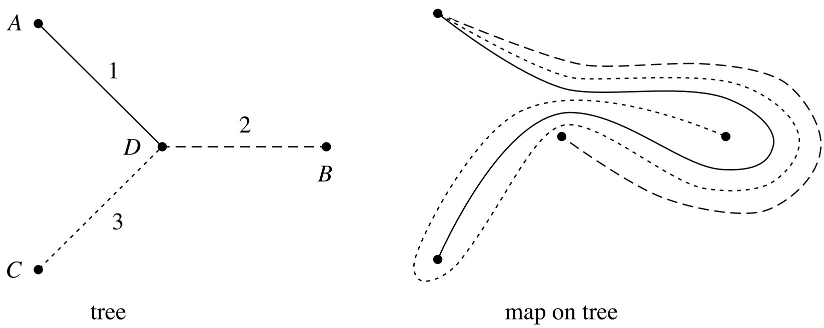

Figure 3 Map on tree for Example 2.

EXAMPLE 2. Consider a tree with three edges and four vertices (A, B, C, D) embedded in a disk (see Figure 3). The tree is such that all three edges have vertex D in common. Pairwise, they have no other vertices in common. We will generate a map on this disk by $\sigma_{CB}\sigma_{AC}\sigma_{AD}^{-1}$. This gives us the picture shown in Figure 3. The Markov matrix is

$$M = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 3 \\ 1 & 0 & 2 \end{bmatrix}.$$

The largest eigenvalue is $\lambda = 2 + \sqrt{3}$. Constructing three rectangles, we have the picture shown in Figure 4. There are four singularities with angle π. As there are two cusps, the fixed point on the boundary will have angle 2π.

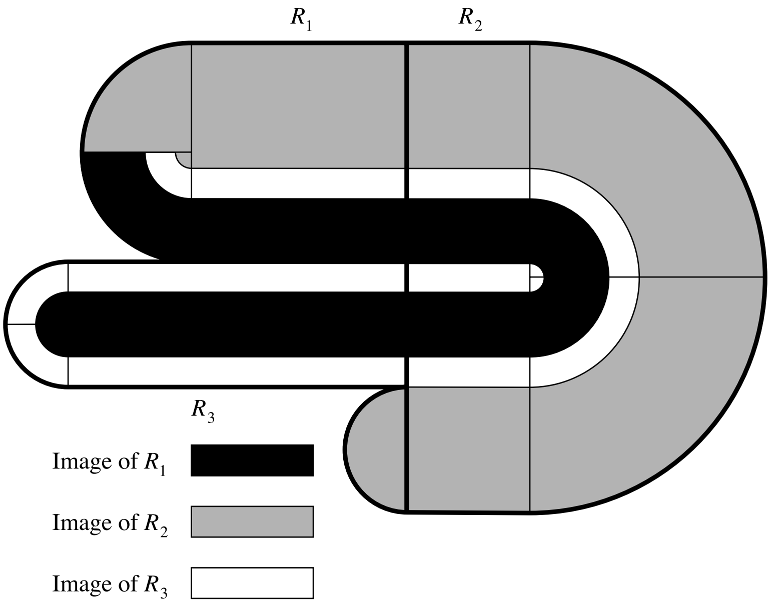

Figure 4 Pseudo-Anosov homeomorphism of Example 2.

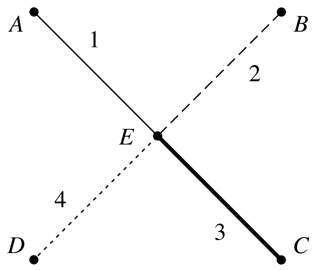

Figure 5 Tree for Example 3.

EXAMPLE 3. Consider the tree shown in Figure 5, with four edges and five vertices (A, B, C, D, E), embedded in a disk. We will generate a map on this disk by $\sigma_{BC}\sigma_{CD}\sigma_{AD}\sigma_{AE}^{-1}$. In this case the Markov matrix is

$$M = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 4 & 2 & 3 & 6 \\ 2 & 0 & 2 & 3 \\ 1 & 0 & 0 & 2 \end{bmatrix}.$$

The largest eigenvalue is

$$\frac{1}{2}\left(\frac{7}{2}+\frac{\sqrt{13}}{2}+\sqrt{-4+\left(-\frac{7}{2}-\frac{\sqrt{13}}{2}\right)^2}\right).$$

In this example, there are five singularities with angle π and the singularity on the boundary has angle 3π.

3. Lifts of Pseudo-Anosov Homeomorphisms and the Expanding Factor

Suppose f is a pseudo-Anosov homeomorphism on M. Let $M' = M \setminus \{$singularities with an odd number of prongs$\}$. Then M' has a two-fold covering space $\tilde{M}'$, which orients the line field. From algebraic topology, we have the following theorem.

THEOREM 3.1. *Let Y be connected and locally path connected, and let $g\colon (Y, y_0) \to (X, x_0)$ be continuous. If $(\tilde{X}, p)$ is a covering space of X, then there exists a unique $\tilde{g}\colon (Y, y_0) \to (\tilde{X}, \tilde{x}_0)$ (where $\tilde{x}_0 \in p^{-1}(x_0)$) lifting g if and only if $g_*\pi_1(Y, y_0) \subseteq p_*\pi_1(\tilde{X}, \tilde{x}_0)$.*

Proof. See [13, pp. 284–286]. $\square$

Since f permutes the singularities with an odd number of prongs, $f\colon M' \to M'$ is also a homeomorphism. If we let $Y = \tilde{M}'$ and $g = f \circ p$ then, by Theorem 3.1, we have that $f\colon M' \to M'$ will lift to $\tilde{f}\colon \tilde{M}' \to \tilde{M}'$; that is, the diagram

$$\begin{array}{ccc}
\tilde{M}' & \xrightarrow{\tilde{f}} & \tilde{M}' \\
\downarrow{\scriptstyle p} & & \downarrow{\scriptstyle p} \\
M' & \xrightarrow{\tilde{f}} & M'
\end{array}$$

will commute if and only if $f_*(p_*\pi_1(\tilde{M}')) \subseteq p_*\pi_1(\tilde{M}')$. But $p_*\pi_1(\tilde{M}')$ is the subgroup of $\pi_1(M')$ of homotopy classes of loops that respect orientation of the line field, and f_* will preserve this subgroup. Hence, $f_*(p_*\pi_1(\tilde{M}')) \subseteq p_*\pi_1(\tilde{M}')$ and f lifts. Since f is a homeomorphism, we can consider f^{-1}. We can also lift f^{-1} to a continuous map $(\tilde{f}^{-1})$ that satisfies $f^{-1} \circ p = p \circ (\tilde{f}^{-1})$. Hence, $p \circ (\tilde{f}^{-1}) \circ \tilde{f} = f^{-1} \circ p \circ \tilde{f} = f^{-1} \circ f \circ p = p$. Thus $(\tilde{f}^{-1}) \circ \tilde{f}$ is a lift of the identity map on M'. This implies that $(\tilde{f}^{-1}) \circ \tilde{f}$ is a covering translation α. Since covering translations are homeomorphisms, we will have $(\alpha^{-1} \circ (\tilde{f}^{-1})) \circ \tilde{f} = \mathrm{id}$ and $\tilde{f}$ is a homeomorphism. If we replace the n deleted points and extend $\tilde{f}$, we will have a homeomorphism $\tilde{f}\colon \tilde{M} \to \tilde{M}$ such that $f \circ p = p \circ \tilde{f}$. Thus we have the following theorem.

THEOREM 3.2. *If $f\colon M \to M$ is a pseudo-Anosov homeomorphism, then f lifts to a pseudo-Anosov homeomorphism $\tilde{f}\colon \tilde{M} \to \tilde{M}$, where $\tilde{M}$ is the branched double cover that orients the unstable line field of M.*

Proof. From the foregoing discussion we have that f lifts to a homeomorphism $\tilde{f}\colon \tilde{M} \to \tilde{M}$. The branched double cover $\tilde{M}$ has a pseudo-Euclidean structure. The foliations $W^u(f)$ and $W^s(f)$ will lift to two perpendicular foliations on $\tilde{M}$. Furthermore, the expanding factor for $\tilde{f}$ will be the same as the one for f. Hence $\tilde{f}$ is a pseudo-Anosov homeomorphism. $\square$

From Theorem 1.8 we have that $\chi(\tilde{M}) = 2\chi(M) - n$, where n is the number of singularities with an odd number of prongs. Thus, in Examples 1 and 2, $\chi(\tilde{M}) = 0$; in these examples, $\tilde{M}$ is the 2-dimensional torus. In Example 3, $\chi(\tilde{M}) = -2$; here $\tilde{M}$ is the orientable surface of genus 2. In general, if k is odd then singularities with angle $k\pi$ contribute $(2-k)/2$ to the Euler characteristic of M, but their lifts contribute $1-k$ to the Euler characteristic of $\tilde{M}$. If k is even then they contribute $(2-k)/2$ to the Euler characteristic of M, but their lifts contribute $2-k$ to the Euler characteristic of $\tilde{M}$. This implies that, if M has only singularities of angle π or 2π, then $\tilde{M}$ is the 2-dimensional torus. If M has singularities of angle 3π or more, we no longer lift to the torus but instead to an orientable surface of genus g with $g \geq 2$.

If the covering space is the torus then, since singularities of angle π and 2π both lift to points of angle 2π, $\tilde{f}$ will be a hyperbolic toral automorphism. The only expanding factors for hyperbolic toral automorphisms of $\mathbb{T}^2$ are roots of quadratics of the form $x^2 + nx \pm 1$, where $n \in \mathbb{Z}$ and $|\lambda| \neq 1$. Thus our original expanding factor must be a root of a quadratic of this form.

The covering space will always be an orientable surface of genus g with oriented unstable manifolds. In general, we have the following result.

THEOREM 3.3. *Let $f\colon M \to M$ be a pseudo-Anosov homeomorphism on an orientable surface of genus g with oriented unstable manifolds. Let $\mathcal{P}$ be a Markov partition for f, with Markov matrix $\mathcal{A}$. If f* preserves *the orientation of unstable manifolds then the eigenvalues of $f_{*1}\colon H_1(M;\mathbb{R}) \to H_1(M;\mathbb{R})$ are the same as those of $\mathcal{A}$ including multiplicity, with the possible exception of some zeros and roots of unity. Hence, the expanding factor λ is an eigenvalue of $f_{*1}\colon H_1(M;\mathbb{R}) \to H_1(M;\mathbb{R})$. If f* reverses *the orientation of unstable manifolds then the eigenvalues of $f_{*1}\colon H_1(M;\mathbb{R}) \to H_1(M;\mathbb{R})$ are the same as those of $-\mathcal{A}$ including multiplicity, with the possible exception of some zeros and roots of unity. Hence $-\lambda$ is an eigenvalue of $f_{*1}\colon H_1(M;\mathbb{R}) \to H_1(M;\mathbb{R})$.*

In order to prove this theorem, we will need the following two lemmas.

LEMMA 3.4. *Suppose we have the commutative diagram*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{f} & B & \xrightarrow{g} & C & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle\alpha} & & \downarrow{\scriptstyle\beta} & & \downarrow{\scriptstyle\gamma} & & \\ 0 & \longrightarrow & A & \xrightarrow{f} & B & \xrightarrow{g} & C & \longrightarrow & 0, \end{array}$$

where the rows are exact and where A, B, and C are vector spaces. Then there exists a choice of basis such that

$$\beta = \begin{bmatrix} \alpha & * \\ 0 & \gamma \end{bmatrix},$$

and hence the set of eigenvalues of β is the union of those of α and γ.

The proof is straightforward.

LEMMA 3.5. *Suppose we have the commutative diagram*

$$\begin{array}{ccccccc} A & \xrightarrow{f} & B & \xrightarrow{g} & C & \xrightarrow{h} & D \\ \downarrow{\scriptstyle\alpha} & & \downarrow{\scriptstyle\beta} & & \downarrow{\scriptstyle\gamma} & & \downarrow{\scriptstyle\delta} \\ A & \xrightarrow{f} & B & \xrightarrow{g} & C & \xrightarrow{h} & D, \end{array}$$

where the rows are exact and where A, B, C, and D are vector spaces. Then the eigenvalues of γ are a subset of those of β and δ.

Proof. Consider the short exact sequence

$$0 \longrightarrow B/\operatorname{Im} f \xrightarrow{\bar{g}} C \xrightarrow{h} \operatorname{Im} h \longrightarrow 0.$$

Since $\delta h = h\gamma$, we have $\delta(\operatorname{Im} h) \subseteq \operatorname{Im} h$. Furthermore, since $\beta f = f\alpha$, we have $\beta(\operatorname{Im} f) \subseteq \operatorname{Im} f$. Hence $\bar{\beta}\colon B/\operatorname{Im} f \to B/\operatorname{Im} f$ is well-defined. Thus we have the commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & B/\operatorname{Im} f & \xrightarrow{\bar{g}} & C & \xrightarrow{h} & \operatorname{Im} h & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle\bar{\beta}} & & \downarrow{\scriptstyle\gamma} & & \downarrow{\scriptstyle\delta|_{\operatorname{Im} h}} & & \\ 0 & \longrightarrow & B/\operatorname{Im} f & \xrightarrow{\bar{g}} & C & \xrightarrow{h} & \operatorname{Im} h & \longrightarrow & 0. \end{array}$$

By Lemma 3.4, the set of eigenvalues of γ is the union of those of $\bar{\beta}$ and $\delta|_{\operatorname{Im} h}$. However, the eigenvalues of $\delta|_{\operatorname{Im} h}$ are a subset of those of δ. And since we have the commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \operatorname{Im} f & \xrightarrow{i} & B & \xrightarrow{\pi} & B/\operatorname{Im} f & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle\beta|_{\operatorname{Im} f}} & & \downarrow{\scriptstyle\beta} & & \downarrow{\scriptstyle\bar{\beta}} & & \\ 0 & \longrightarrow & \operatorname{Im} f & \xrightarrow{i} & B & \xrightarrow{\pi} & B/\operatorname{Im} f & \longrightarrow & 0, \end{array}$$

it follows that the eigenvalues of $\bar{\beta}$ are a subset of those of β. □

Proof of Theorem 3.3. Let S be the union of the stable boundaries of the rectangles in our partition. This is a finite set of contractible intervals or prongs. Let U be the union of the interiors of the unstable boundaries of the rectangles in our partition. The set U is a finite set of contractible intervals. Let $X = M \setminus U$. Then $S \subseteq X \subseteq M$. Now $f(S) \subseteq S$ and $f(X) \subseteq X$, which gives us the following commutative diagram of pairs of spaces:

$$\begin{array}{ccccc}
(X,S) & \longrightarrow & (M,S) & \longrightarrow & (M,X) \\
\Big\downarrow{\scriptstyle f|X} & & \Big\downarrow{\scriptstyle f} & & \Big\downarrow{\scriptstyle f} \\
(X,S) & \longrightarrow & (M,S) & \longrightarrow & (M,X).
\end{array}$$

By the theorem of exact sequences of triples in algebraic topology, we have the following commutative diagram with exact rows:

$$\begin{array}{ccccccccc}
\longrightarrow & H_2(M,X) & \longrightarrow & H_1(X,S) & \longrightarrow & H_1(M,S) & \longrightarrow & H_1(M,X) & \longrightarrow \\
& \Big\downarrow{\scriptstyle f_{*2}} & & \Big\downarrow{\scriptstyle (f|X)_{*1}} & & \Big\downarrow{\scriptstyle f_{*1}} & & \Big\downarrow{\scriptstyle f_{*1}} & \\
\longrightarrow & H_2(M,X) & \longrightarrow & H_1(X,S) & \longrightarrow & H_1(M,S) & \longrightarrow & H_1(M,X) & \longrightarrow .
\end{array}$$

Consider $H_1(M,X)$. Every line segment in M with endpoints in X can be deformed within its homotopy class to a line segment X. Hence, $H_1(M,X)=0$.

Next consider $(f|_X)_{*1}\colon H_1(X,S)\to H_1(X,S)$. For each rectangle in our partition, consider a line segment across the rectangle in the unstable direction with endpoints in the stable boundaries, oriented compatibly with the orientation of the unstable manifolds. This set of line segments forms a set of generators for $H_1(X,S)$. Suppose f is orientation-preserving on unstable manifolds. With respect to this basis, f_{*1} is the matrix whose ijth entry is given by the number of times the image under f of the jth generator crosses ith rectangle. Since f preserves the orientation of unstable manifolds, the image will always cross with an orientation that is compatible with the orientation of the unstable manifolds. Hence the ijth entry will be the same as the geometric intersection number of $R_i\cap f(R_j)$ and hence f_{*1} will be the same as $\mathcal{A}$, the Markov matrix. Similarly, if f is orientation-reversing on unstable manifolds, then f_{*1} will be the same as $-\mathcal{A}$.

Finally, consider $f_{*2}\colon H_2(M,X)\to H_2(M,X)$. For each component of U, consider a small disk with boundary in X that contains this component of U and no others. This set of disks forms a set of generators for $H_2(M,X)$. Given a component $u\subseteq U$, either the closure of u contains a periodic point or not. A component cannot contain more than one periodic point. Consider an iterate of f, f^m, such that (a) the periodic points in U are fixed and (b) if the closure of U contains prongs then the prongs are also fixed. Let F be the set of these fixed points. Since under f^{-1} the set of intervals of U would have to map to themselves, each component of U must lie in the unstable manifold of one of the points of F. If the closure of u contains a fixed point p then $f^m(u)$ will intersect itself and perhaps other generators on the unstable manifold of p. If the closure of u does not contain a fixed point, then $f^{mn}(u)$ for some large n will not intersect any of the generators. Hence we can order the generators—with those having fixed points first—such that $f^m_{*2}\colon H_2(M,X)\to H_2(M,X)$ is given by the matrix

$$f^m_{*2}=\begin{bmatrix} 1 & 0 & 0 & \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & 1 & \\ * & * & * & N \end{bmatrix},$$

where N is nilpotent. Hence the eigenvalues of f_{*2} are zeros and roots of unity.

By Lemma 3.5, the eigenvalues of $f_{*1}\colon H_1(M,S)\to H_1(M,S)$ are a subset of those of $(f|_X)_{*1}\colon H_1(X,S)\to H_1(X,S)$, that is, a subset of those of $\mathcal{A}$ or $-\mathcal{A}$. Furthermore, the eigenvalues of $(f|_X)_{*1}\colon H_1(X,S)\to H_1(X,S)$ are a subset of those of $f_{*1}\colon H_1(M,S)\to H_1(M,S)$ and $f_{*2}\colon H_2(M,X)\to H_2(M,X)$. Hence, if f is orientation-*preserving* on unstable manifolds then any eigenvalue of $\mathcal{A}$ that is not zero or a root of unity must also be an eigenvalue of $f_{*1}\colon H_1(M,S)\to H_1(M,S)$. Thus the eigenvalues of $f_{*1}\colon H_1(M,S)\to H_1(M,S)$ are those of $\mathcal{A}$, with the possible exception of some zeros and roots of unity. Likewise, if f is orientation-*reversing* on unstable manifolds then any eigenvalue of $-\mathcal{A}$ that is not zero or a root of unity must also be an eigenvalue of $f_{*1}\colon H_1(M,S)\to H_1(M,S)$. Thus the eigenvalues of $f_{*1}\colon H_1(M,S)\to H_1(M,S)$ are those of $-\mathcal{A}$, with the possible exception of some zeros and roots of unity.

Since $f(S)\subseteq S$, we may apply the theorem of exact sequences of pairs in algebraic topology to the pair (M,S). Thus we have the following commutative diagram with exact rows:

$$\begin{array}{ccccccccc}
\longrightarrow & H_1(S) & \longrightarrow & H_1(M) & \longrightarrow & H_1(M,S) & \longrightarrow & H_0(S) & \longrightarrow \\
 & \big\downarrow{\scriptstyle (f|_S)_{*1}} & & \big\downarrow{\scriptstyle f_{*1}} & & \big\downarrow{\scriptstyle f_{*1}} & & \big\downarrow{\scriptstyle (f|_S)_{*0}} & \\
\longrightarrow & H_1(S) & \longrightarrow & H_1(M) & \longrightarrow & H_1(M,S) & \longrightarrow & H_0(S) & \longrightarrow .
\end{array}$$

Since S is the union of a finite number of contractible intervals or prongs, we know that $H_1(S)=0$.

Next consider $(f|_S)_{*0}\colon H_0(S)\to H_0(S)$. We can consider each component of S as a basis element for $H_0(S)$. Each basis element is either periodic or eventually periodic under $f|_S$. Hence, with the proper ordering, $f|_S$ is given by the matrix

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ * & \ddots & 0 & 0 \\ * & * & 0 & 0 \\ * & * & * & P \end{bmatrix},$$

where P is a permutation matrix. Since $(f|_S)_{*0}$ is also given by this matrix, the eigenvalues of $(f|_S)_{*0}\colon H_0(S)\to H_0(S)$ are zeros and roots of unity.

By Lemma 3.5, the eigenvalues of $f_{*1}\colon H_1(M)\to H_1(M)$ are a subset of those of $f_{*1}\colon H_1(M,S)\to H_1(M,S)$. Furthermore, the eigenvalues of $f_{*1}\colon H_1(M,S)\to H_1(M,S)$ are a subset of those of $f_{*1}\colon H_1(M)\to H_1(M)$ and $(f|_S)_{*0}\colon H_0(S)\to H_0(S)$. Hence, any eigenvalue of $f_{*1}\colon H_1(M,S)\to H_1(M,S)$ that is not zero or a root of unity must be an eigenvalue of $f_{*1}\colon H_1(M)\to H_1(M)$. Thus the eigenvalues of $f_{*1}\colon H_1(M)\to H_1(M)$ are those of $f_{*1}\colon H_1(M,S)\to H_1(M,S)$, with the possible exception of some zeros and roots of unity. If f is orientation-preserving on unstable manifolds, this implies that the eigenvalues of $f_{*1}\colon H_1(M)\to H_1(M)$ are those of $\mathcal{A}$ including multiplicity, with the possible exception of some zeros and roots of unity. Hence λ is an eigenvalue of $f_{*1}\colon H_1(M)\to H_1(M)$. If f is orientation-reversing on unstable manifolds, this implies that the eigenvalues of $f_{*1}\colon H_1(M)\to H_1(M)$ are those of $-\mathcal{A}$ including multiplicity, with the possible exception of some zeros and roots of unity. Hence $-\lambda$ is an eigenvalue of $f_{*1}\colon H_1(M)\to H_1(M)$. $\square$

Now $H_1(M)$ has a basis such that $f_{*1}\colon H_1(M) \to H_1(M)$ is given by a $2g \times 2g$ matrix with integer entries, where g is the genus of M. Since characteristic polynomials are monic, either λ or $-\lambda$ satisfies a monic polynomial with integer coefficients of degree $2g$.

We also have the following results.

PROPOSITION 3.6. *If $f\colon M \to M$ is an orientation-reversing homeomorphism of a closed oriented surface and if μ is an eigenvalue of $f_*\colon H_1(M;\mathbb{R}) \to H_1(M;\mathbb{R})$, then $-1/\mu$ is also an eigenvalue and has the same multiplicity.*

Proof. See [2]. □

Using the ideas of this proof, we can also show that the next proposition holds.

PROPOSITION 3.7. *If $f\colon M \to M$ is an orientation-preserving homeomorphism of a closed oriented surface and if μ is an eigenvalue of $f_*\colon H_1(M;\mathbb{R}) \to H_1(M;\mathbb{R})$, then $1/\mu$ is also an eigenvalue and has the same multiplicity.*

Even if the branched double cover that orients the unstable line field is not $\mathbb{T}^2$, we can still ask what sort of a quadratic equation is satisfied by a quadratic expanding factor of a pseudo-Anosov homeomorphism. Because λ or $-\lambda$ satisfies a monic polynomial with integer coefficients, if λ satisfies an irreducible quadratic then it must be a monic one with integer coefficients. Hence λ is integral over $\mathbb{Z}$. Theorem 3.3 and Propositions 3.6 and 3.7 give us the following result.

COROLLARY 3.8. *Let λ be the expanding factor of a pseudo-Anosov homeomorphism f, and suppose λ is the root of an irreducible quadratic equation over the rationals. Then λ satisfies a quadratic of the form $x^2 + nx \pm 1$, where $n \in \mathbb{Z}$ and $|\lambda| \neq 1$.*

Proof. If μ is an element of the quadratic field $\mathbb{Q}[\sqrt{d}\,]$ (where d is a square-free integer), then μ is of the form $a + b\sqrt{d}$ with $a, b \in \mathbb{Q}$. We define the conjugate of μ as $\sigma(\mu) = a - b\sqrt{d}$, and we define the norm of μ as $N(\mu) = \mu\sigma(\mu)$. If μ is integral over $\mathbb{Z}$, then both $\mu + \sigma(\mu)$ and $N(\mu)$ are integers. Hence μ will satisfy the equation $x^2 - (\mu + \sigma(\mu))x + N(\mu)$ and so λ satisfies the equation $x^2 - (\lambda + \sigma(\lambda))x + N(\lambda)$. We will show one of four possible cases. Suppose f preserves the orientation of both stable and unstable manifolds. By Theorem 3.3, λ is a root of $f_{*1}\colon H_1(M) \to H_1(M)$; by Proposition 3.7, $1/\lambda$ is also a root of f_{*1}. Now $N(\lambda)N(1/\lambda) = N(1) = 1$. Since both λ and $1/\lambda$ are integral, we must have $N(\lambda) = \pm 1$. The other three cases are similar. □

The following is an example of a pseudo-Anosov homeomorphism with a quadratic expanding factor whose branched double cover that orients the unstable manifold is an orientable surface of genus 3.

EXAMPLE 4. Consider the homeomorphism with expanding factor $2 + \sqrt{3}$ in Example 2. By the Markov matrix, this homeomorphism has five or fewer fixed points. There is a fixed point on the boundary that lies in two rectangles. Hence there are three fixed points that lie in the interiors of rectangles. One of these lies in the interior of R_1; call it P. We can cut along the stable boundary of P until we hit the unstable boundary of R_1 (see Figure 6). If we do this with two copies of the map, we can rotate one copy by 180° and sew one copy to the other along the cuts in the stable boundaries. We sew side A to side B and side B to side A (see Figure 7). This is the branched cover formed by the two-fold cover of the disk punctured at P. The fundamental group of a punctured disk is $\mathbb{Z}$. The subgroup associated to this cover is $2\mathbb{Z}$. Since isomorphisms of $\mathbb{Z}$ preserve this subgroup, the map must lift to the covering space. We can replace the punctured point and, when we sew up the boundary, we will obtain a pseudo-Anosov homeomorphism with the same expanding factor.

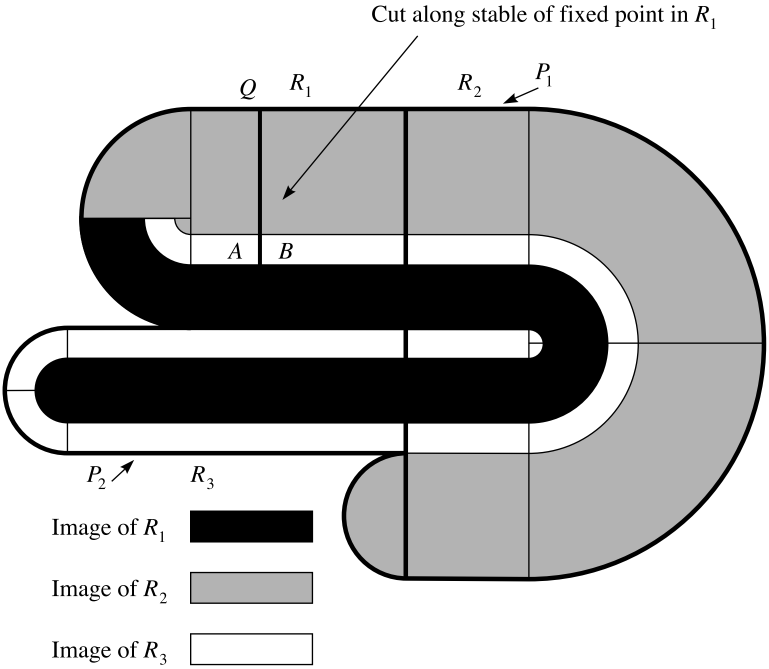

Figure 6

The unstable boundary will sew together as follows. Let P_1 and P_2 be the two points on the boundary that become identified and form a fixed point when the unstable boundary is sewn together. Let Q be the point on the unstable boundary of R_1 that lies on the stable manifold of the fixed point in R_1. Let a be the distance along the unstable boundary from Q to P_1, and let b be the distance along the unstable boundary from P_1 to P_2. Finally, let c be the distance from P_2 to Q. We have that $a + c = b$. The two copies will fit together as shown in Figure 7. This creates a pseudo-Anosov homeomorphism with the same expanding factor,

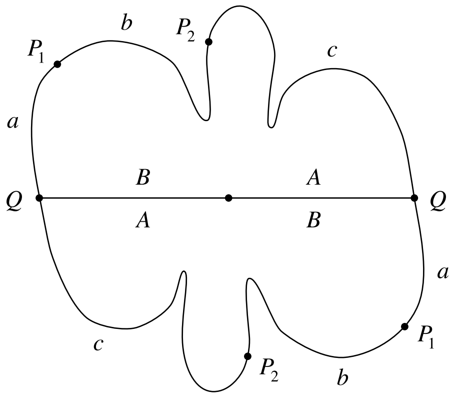

Figure 7

$2+\sqrt{3}$. There are, however, eight singularities of angle π and two with angle 4π. Hence, the Euler characteristic of branched double covering space must be -4. Thus, the branched double covering space must be an orientable surface of genus 3.

Thus, a pseudo-Anosov homeomorphism with a quadratic expanding factor must satisfy the same type of polynomial as those given by hyperbolic toral automorphisms, even though the map does not have the 2-torus as its lift.

References

[1] J. Birman, *Braids, links, and mapping class groups,* Ann. of Math. Stud., 82, Princeton Univ. Press, Princeton, NJ, 1975.
[2] P. Blanchard and J. Franks, *The dynamical complexity of orientation-reversing homeomorphisms of surfaces,* Invent. Math. 62 (1980/81), 333–339.
[3] A. Casson and S. Bleiler, *Automorphisms of surfaces after Nielsen and Thurston*, Cambridge Univ. Press, Cambridge, 1988.
[4] A. Fathi and F. Laudenbach, *Les feuilletages mesurés,* Travaux de Thurston sur les surfaces, Astérisque 66–7, Soc. Math. France, Paris, 1979.
[5] A. Fathi and M. Shub, *Some dynamics of pseudo-Anosov diffeomorphisms,* Travaux de Thurston sur les surfaces, Astérisque 66–7, Soc. Math. France, Paris, 1979.
[6] J. Franks, *Global analysis,* Proc. Sympos. Pure Math., 14, Amer. Math. Soc., Providence, RI, 1970.
[7] ———, *Homology and dynamical systems,* Amer. Math. Soc., Providence, RI, 1982.
[8] J. Franks and M. Misiurewicz, *Cycles for disk homeomorphisms and thick trees,* Neilson theory and dynamical systems (South Hadley, MA, 1992), pp. 69–139, Amer. Math. Soc., Providence, RI, 1993.
[9] F. R. Gantmacher, *The theory of matrices,* vol. 2, Chelsea, New York, 1959.

[10] H. Masur and I. Smillie, *Hausdorff dimension of sets of nonergodic measured foliation*, Ann. of Math. (2) 134 (1991), 455–543.
[11] J. Milnor, *Topology from the differentiable viewpoint,* Univ. Press of Virginia, Charlottesville, 1965.
[12] C. Robinson, *Dynamical systems,* 2nd ed., CRC Press, Boca Raton, FL, 1999.
[13] J. Rotman, *An introduction to algebraic topology,* Springer-Verlag, New York, 1988.
[14] M. Snavely, *Markov partitions for hyperbolic automorphisms of the two-dimensional torus,* Ph.D. dissertation, Northwestern Univ., Evanston, IL, 1990.
[15] M. Stafford, *Markov partitions for the doubling map,* Ph.D. dissertation, Northwestern Univ., Evanston, IL, 1989.

Department of Mathematics
Northwestern University
Evanston, IL 60208

e_rykken@math.nwu.edu

Current address
Department of Mathematical Sciences
Muhlenberg College
Allentown, PA 18104