Smale diffeomorphisms and surface topology

STEVE BATTERSON AND JOHN SMILLIE

Department of Mathematics, Emory University, Atlanta, Georgia 30322, *USA*;
Department of Mathematics, Herbert H. Lehman College, CUNY, Bronx, NY 10468,
USA and *Mathematical Sciences Research Institute, University of California, Berkeley,
CA* 94720, *USA*

Abstract. This paper concerns Smale diffeomorphisms of compact oriented surfaces. Relationships are found between the isotopy class of the map and the dynamics of its basic sets. The form of the dynamical properties involves restrictions on periods and reduced zeta functions.

0. *Introduction*

In this paper we discuss relations between the topology and dynamics of Smale diffeomorphisms of surfaces. The dynamics of a Smale map can be largely characterized by describing the chain recurrent set $\mathcal{R}$. This is a set composed of basic sets $\Sigma_1 \cdots \Sigma_k$ each of which is topologically conjugate to a topologically transitive subshift of finite type. Blanchard-Franks [**B-F**] and Fried [**Fr**] characterize the reduced zeta functions of subshifts which can occur on a surface of a given genus. In this paper we extend these results in two directions. We describe restrictions on basic sets which can occur for maps in particular isotopy classes. These restrictions involve reduced zeta function and period conditions. Theorem 1 describes restrictions on Smale maps isotopic to the identity. Theorem 4 describes restrictions on Smale maps in pseudo-Anosov isotopy classes.

In another direction we find restrictions on which collections of subshifts can occur as basic sets of a given map. In theorem 2 we describe the situation for S^2 giving necessary and sufficient conditions on reduced zeta functions and information about periods. In theorem 3 we show that there are new compatibility restrictions on surfaces of each genus and in many isotopy classes.

We believe that the techniques used in this paper are capable of describing precisely which collections of reduced zeta functions can occur for collections of basic sets in any given isotopy class. We thank Paul Blanchard for his useful comments.

1. *Periods, subshifts, and filtrations*

The reader should refer to [**F**] for further definitions and background. Let $\sigma: \Sigma \to \Sigma$ be a subshift of finite type. If σ is not mixing then we can write Σ as a disjoint union $\Sigma = \Sigma_1 \cup \cdots \cup \Sigma_k$ where σ cyclically permutes these sets and $\sigma^k|\Sigma_1$ is a mixing subshift. This decomposition is unique.

Definition. The sets $\Sigma_1 \cdots \Sigma_k$ are the *mixing components of* Σ. The number of mixing components of Σ is the *period of* σ.

The following propositions relate the period k of a subshift to other invariants. We denote the zeta function of a subshift σ by $\zeta(t)$. The corresponding function for $\sigma^k|\Sigma_1$ is denoted by $\zeta_k(t)$.

PROPOSITION. *The period of σ is the greatest common divisor of the periods of all the periodic points.*

PROPOSITION. *If σ is a subshift of period k then $\zeta_k(t^k) = \zeta(t)$.*

Note that the second proposition implies that k divides each power in the polynomial $[\zeta(t)]^{-1}$.

Proof. $\zeta(t) = \prod_\gamma [1 - t^{p(\gamma)}]^{-1}$ where the product is over all periodic orbits γ and $p(\gamma)$ is the period of γ.

The period of each periodic orbit is a multiple of k. For each n there is a one to one correspondence between the period n orbits of $\sigma^k|\Sigma_1$ and the period nk orbits of σ. This yields the following relations in which the first product is over the periodic orbits of $\sigma^k|\Sigma_1$.

$$\zeta_k(t^k) = \prod_\psi (1 - t^{kp(\psi)})^{-1} = \prod_\gamma (1 - t^{p(\gamma)})^{-1} = \zeta(t). \qquad \square$$

The following technical lemma produces a filtration which is essential to the development of the paper. The filtration allows us to decompose a Smale diffeomorphism into several Smale diffeomorphisms.

FILTRATION LEMMA. *If f is a Smale diffeomorphism of an oriented surface M with basic sets $\Sigma_1 \ldots \Sigma_n$ then there exists a collection of submanifolds $\phi = M_0 \subset M_1 \subset \cdots \subset M_n = M$ such that for $1 \le i \le n$*

(1) $f(M_i) \subset \text{int } M_i$;

(2) $\Sigma_i = \bigcap_{j=-\infty}^\infty f^j(M_i \setminus M_{i-1})$;

(3) $\text{cl}(M_i \setminus f(M_i))$ *is a union of disjoint annuli.*

This lemma was inspired by [B-F] and differs from the standard filtration lemma owing to condition (3).

Proof. We may assume that we are given a filtration of M that satisfies conditions (1) and (2) of the lemma. In order to prove (3) we claim it suffices to show that for each j, $M_j \setminus f(M_j)$ contains no disks. To see this observe that since $\chi(M_j) = \chi(f(M_j))$ we have $\chi(M_j \setminus f(M_j)) = 0$. Each component of the closure of the difference is an oriented manifold with non-trivial boundary. If there are no disk components then the Euler characteristic of each component is non-positive and hence is zero. Thus each component is an annulus.

The proof proceeds as follows. Assume that we are given a filtration for which there is a disk component in $\text{cl}(M_j \setminus f(M_j))$. We will construct a new filtration with the same indexing set and the same intersections with the basic sets but with a smaller total number of boundary components of filtration pieces. Thus after a finite number of alterations we arrive at a filtration with no disks in $\text{cl}(M_j \setminus f(M_j))$.

Assume that $\operatorname{cl}(M_j \setminus f(M_j))$ contains a disk D. There are two cases.

Case 1. $\operatorname{bdy}(D) \subset \operatorname{bdy}(f(M_j))$.

Let $D' = f^{-1}(D)$ and define

$$M'_k = \begin{cases} M_k & \text{if } k < j, \\ M_k \cup D' & \text{if } k \geq j. \end{cases}$$

Clearly $M'_k \subset M'_{k+1}$ and $f(M'_k) \subset \operatorname{int}(M'_k)$ for $k < j$. If $k \geq j$ then

$$f(M'_k) = f(M_k \cup D') = f(M_k) \cup D \subset \operatorname{int}(M_k) \subset \operatorname{int}(M'_k).$$

Since $\operatorname{bdy}(D') \subset \operatorname{bdy}(M_j)$ it follows that $\operatorname{bdy}(D' \cup M_j)$ is strictly contained in $\operatorname{bdy}(M_j)$. If $k > j$ then $\operatorname{bdy}(M'_k) \subset \operatorname{bdy}(M_k)$ because $\operatorname{bdy}(D') \subset \operatorname{int}(M_k)$.

Case 2. $\operatorname{bdy}(D) \subset \operatorname{bdy}(M_j)$.

Define

$$M'_k = \begin{cases} M_k & \text{if } k > j, \\ M_k - D & \text{if } k \leq j. \end{cases}$$

Clearly $M'_k \subset M'_{k+1}$ and $f(M'_k) \subset \operatorname{int}(M'_k)$ for $k > j$. If $k \leq j$ then D is disjoint from $f(M_j)$ and hence from $f(M_k)$ so $f(M_k) \subset \operatorname{int}(M_k) \setminus D$,

$$f(M'_k) = f(M_k) \setminus f(D) \subset \operatorname{int}(M_k) \setminus D = \operatorname{int}(M'_k).$$

Since $\operatorname{bdy}(D) \subset \operatorname{bdy}(M_j)$ it follows that $\operatorname{bdy}(M_j \setminus D)$ is strictly contained in $\operatorname{bdy}(M_j)$. If $k < j$ then $\operatorname{bdy}(M'_k) \subset \operatorname{bdy}(M_k)$ because $\operatorname{bdy}(D) \cap \operatorname{bdy}(M_k) = \varnothing$ □

Given a filtration from the lemma we will analyze the dynamics of the Smale map via the pieces $N_i = \operatorname{cl}(M_i \setminus M_{i-1})$.

PROPOSITION. *If f is a Smale diffeomorphism of an oriented surface, Σ_i is an index-one basic set, and N_i is as above then*

(1) *Each non-annular component of N_i meets Σ_i.*

(2) *If N_i has a non-annular component then the period of Σ_i is divisible by the number of non-annular components.*

Before proving the proposition we state a lemma which follows from the Morse–Smale inequalities (see [**F**]).

LEMMA. *Let f be a Smale diffeomorphism on a connected surface M and suppose f has no basic sets of index one. Then $M = S^2$ and f has exactly one source and one sink.*

Proof of proposition. Construct a new manifold $\hat{N}_i$ from N_i by adding disks to all boundary components. We can define a Smale map $\hat{f}$ on $\hat{N}_i$ which agrees with f on $N_i \cap f^{-1}(N_i)$ and has one periodic source or sink in each attached disk. Choose k so that $\hat{f}^k$ fixes all components of $\hat{N}_i$. If Σ_i does not meet some component $\hat{C}$ of $\hat{N}_i$ then $\hat{f}^k | \hat{C}$ has no index-one basic sets. By the previous lemma $\hat{C}$ must be an annulus to which two disks have been attached. This proves part one of the proposition.

If N_i contains some non-annular component then each such component must meet Σ_i. Since f acts topologically transitively on Σ_i these components must be cyclically permuted.   □

We shall assume that all filtrations are obtained from the filtration lemma. The *atom associated to a basic set* Σ_i is defined to be the union of these components of cl $(M_i \setminus M_{i-1})$, which meet Σ_i. We close this section with a corollary to the filtration lemma and a result from [J-S],

COROLLARY. *If f is a Smale diffeomorphism of an oriented surface M, then f is isotopic relative to the chain recurrent set to a map, which is invariant on* bdy (M_i) *for each i.*

PROPOSITION [J-S]. *Let Q be a surface, $Q_1 \subset Q$ a compact, incompressible subsurface with $\pi_1(Q_1)$, non-cyclic, and $f: Q_1 \to Q_1$ a map. If f is homotopic in Q to the inclusion map of Q_1 into Q, then it is homotopic in Q_1 to the identity map of Q_1.*

2. Basic sets in the isotopy class of the identity

Given a basic set Σ_i of a Smale diffeomorphism let z_i (or z) denote the mod 2 reduction of the homology zeta function. In [F] it is shown that z_i is equal to the mod 2 reduction of $(\zeta_i)^{(-1)^u}$ where u is the unstable index of Σ_i. If $u = 1$ then $z_i = \det (I - At)$ where A is the matrix representing the subshift. In this section we apply the filtration lemma to show that in the isotopy class of the identity there are strong limitations on the polynomials z_i. Let $\mathscr{C} \subset \mathbb{Z}_2[t]$ be the collection of all finite products $\prod (1 - t^i)^{a_i}$, $a_i \geq 0$.

THEOREM 1. *Let M be an oriented surface of genus g and suppose f is a Smale diffeomorphism of M which is isotopic to the identity. If Σ is a basic set of index one with reduced homology zeta function $z(t)$ then either Σ can be realized on the disk (in which case $z(t) = c(t)/(1 - t)$, for some $c(t) \in \mathscr{C}$) or*
  (1) *if $g = 0$ then $z(t) = c(t)/(1 - t)^2$, for some $c(t) \in \mathscr{C}$, and Σ is mixing;*
  (2) *if $g = 1$ then $z(t) \in \mathscr{C}$;*
  (3) *if $g \geq 2$ then $z(t) = (1 - t)c(t)$, for some $c(t) \in \mathscr{C}$, and Σ is mixing.*
Conversely there exist subshifts with each such reduced homology zeta function which can be realized by a Smale diffeomorphism isotopic to the identity on the surface.

Remark. In particular there is a subshift with $z(t) = (1 + t + t^2)^2$ which can occur as a basic set of a Smale diffeomorphism of S^2 but cannot be realized by a Smale diffeomorphism isotopic to the identity on any surface of higher genus or on the disk. We do not know if there is a subshift which occurs in the isotopy class of the identity on a surface of positive genus but does not occur on the disk.

Proof. Let Σ be a basic set of index one and let N be the atom associated to Σ. Modify f so that it is invariant on N. Let $\hat{N}$ consist of N together with all disks in the complement of N. Modify f on $\hat{N}$ to obtain a map $\hat{f}$ for which there is precisely one sink or source in each of the added disks. Let $m_1, \ldots, m_k$ be the periods of these disks under $\hat{f}$.

Case 1. Each component of N has genus 0.
If bdy $(\hat{N}) = \varnothing$ then $M = \hat{N} = S^2$ and N has one component. If $p(t)$ is the reduced homology zeta function of Σ then

$$\frac{1}{(1 - t)^2} = \frac{p(t)}{\prod (1 - t^{m_i})}.$$

If $m_i = 1$, for some i, then Σ can be realized on a disk. If $m_i \neq 1$ for each i then by the Lefschetz theorem there must be a fixed point. The fixed point can only occur in Σ which is consequently mixing.

Now suppose bdy $(\hat{N}) \neq \varnothing$ and let q be the number of components of $\hat{N}$. Consider $\hat{f}^q$ restricted to a component $\hat{N}_1$. If there is a boundary which is mapped to itself under $\hat{f}^q$ then Σ can be realized on a map of q disks and thus can also be realized on a map of the disk. If $\hat{N}_1$ is not an annulus then $\hat{f}^q | \hat{N}_1$ is homotopic to the identity from the result of [J–S]. In particular the boundaries are invariant.

Suppose $\hat{N}_1$ is an annulus and $\hat{f}^q$ switches the boundaries. Since $\hat{f}^q$ is orientation preserving the generator of $\pi_1(\hat{N}_1)$ must be flipped. However this generator is also a non-trivial element of $\pi_1(M)$. We conclude that the boundaries of the annulus are invariant.

Case 2. Each component of N has positive genus h.

We will show that N is connected. Let N_1 be a component of N and produce two curves, α and β, in N_1 with intersection number 1. Since $f(\alpha)$ is isotopic to α, $f(\alpha) \cap \beta \neq \varnothing$. Thus $f(N_1) \cap N_1 \neq \varnothing$ and $f(N_1) = N_1$. Since f acts transitively on the components of N and each component is invariant there must be exactly one component. Applying [J–S] again to $\hat{N}$ it follows that each boundary component is invariant. Produce f' and N' by attaching a disk with a fixed sink or source to each of these boundary components. Let l be the number of boundary components and $p(t)$ the reduced homology zeta function of Σ. Then

$$\frac{(1-t)^{2h}}{(1-t)^2} = \frac{p(t)}{(1-t)^l \prod (1-t^{m_i})}$$

so

$$p(t) = (1-t)^{2h+l-2} \prod (1-t^{m_i}).$$

Note that $2h + l - 2 \geq 0$. If $2h + l - 2 > 0$ then the form of (2) and (3) is immediate. If $2h + l - 2 = 0$ then $h = 1$, $l = 0$ and it follows that $g = 1$.

If $g > 1$ we will show that Σ is mixing by again showing that it has a fixed point. Since f' is isotopic to the identity on N' the Lefschetz number of f' is the Euler characteristic of N'. Note that every f'-fixed point in $N' \setminus N$ has Lefschetz index 1. If $h > 1$ then $\chi(N') < 0$ and f' has a fixed point of index -1 which consequently must occur in Σ. If $h = 1$ and $g > 1$ then $N' \setminus \hat{N} \neq \varnothing$ and f' has a fixed point of index 1. Since $\chi(N') \leq 0$ there must again be a fixed point of index -1 in Σ.

The converse of this result in the case $g = 0$ or $g = 1$ follows from [Fr]. If $g \geq 2$ then we can realize the zeta function on a disk and hence on any surface. $\qquad \square$

3. *Compatible basic sets*

We now consider the question of which combinations of subshifts can occur in a given isotopy class. The following definitions lead to a formal statement of the problem.

Definition. The *subshift data* of a Smale diffeomorphism of a surface is

$$\langle \{m_i\}_{i \in I}, \{q_j\}_{j \in J}, \{A_k\}_{k \in K} \rangle$$

where the first two sequences give the periods of the sinks and sources respectively.

Each A_k is a matrix for which the corresponding subshift of finite type is topologically conjugate to the restriction of f to a basic set of index one. (Note that A_k is not uniquely determined – see [W]. The *reduced data* of f is the subshift data with each A_k replaced by the corresponding reduced homology zeta function.

Problem. Given an isotopy class of a compact oriented surface characterize the subshift data and reduced data of a Smale diffeomorphism in that class.

The solution of the reduced data problem for the case where K consists of one element is available in [Fr]. Theorem 2 provides a combinatorial solution of the reduced data problem for the isotopy class of the identity on S^2. Additionally it provides information on the periods of the basic sets.

THEOREM 2. *Let f be a Smale diffeomorphism which is isotopic to the identity map on S^2. If $\langle \{m_i\}_{i \in I}, \{n_j\}_{j \in J}, \{A_k\}_{k \in K} \rangle$ is the subshift data and p_k is the reduced homology zeta function of A_k then there exists a tree $\mathcal{T}$ and automorphism ψ of $\mathcal{T}$ such that*

(a) *For each sink and source orbit of f there is associated an orbit of vertices of $\mathcal{T}$ with the same ψ-period such that each vertex has valence one.*

(b) *For each index-one basic set there is associated an orbit of vertices of $\mathcal{T}$ such that the ψ-period divides the period of the basic set and each vertex has two or more incident edge orbits.*

(c) *For $k \in K$, $p_k(t) = \prod (1 + t^{r_l})/(1 + t^{2q})$ where q is the period of the associated orbit of vertices and r_l is the period of an incident edge orbit.*

Moreover given any collection of reduced data for which there is a tree and automorphism satisfying (a), (b) and (c) there is a Smale diffeomorphism isotopic to 1_{S^2} with that reduced data.

Note. A Smale map on a disk admits a similar analysis. Such a map corresponds to a tree with a distinguished vertex fixed by ψ.

Theorem 2 is an abstraction of reasoning used in the following example. The reader should verify that the conclusions are consistent.

Example.

ASSERTION A. *There exists a Smale diffeomorphism isotopic to 1_{S^2} with $\mathcal{R}$ consisting of*

  (i) *a period-6 sink;*
 (ii) *a fixed source;*
(iii) *a fixed saddle;*
(iv) *one index-one basic set.*

ASSERTION B. *There does not exist a Smale diffeomorphism isotopic to 1_{S^2} with $\mathcal{R}$ consisting of*

  (i) *two period 3 sinks;*
 (ii) *a fixed source;*
(iii) *a fixed saddle;*
(iv) *one index one basic set.*

If $p(t)$ is the reduced homology zeta function of the index one basic set in a Smale diffeomorphism with the dynamics specified by A then

$$\frac{(1+t)p(t)}{(1+t^6)(1+t)} = \frac{1}{(1+t)^2}$$

so

$$p(t) = \frac{1+t^6}{(1+t)^2} = \frac{(1+t^3)(1+t^3)}{(1+t)^2} = (1+t+t^2)^2 = 1+t^2+t^4.$$

Note that the index one basic set for B must have the same reduced zeta function.

Proof of assertion A. Proposition 2 of [Fr] provides a Smale diffeomorphism of S^2 with $\mathcal{R}$, consisting of a period-3 sink, a fixed source, and an index one basic set with reduced homology zeta function $1+t+t^2$. Using this map construct a Smale diffeomorphism which permutes two copies of S^2 and has a chain recurrent set consisting of a period-six sink, a period-two source, and an index-one basic set with reduced homology zeta function $1+t^2+t^4 = p$. It is easy to construct a Morse–Smale diffeomorphism of S^2 with a fixed source, a fixed saddle, and a period two sink. Removing the disks containing the period two orbits and joining provides the desired map.
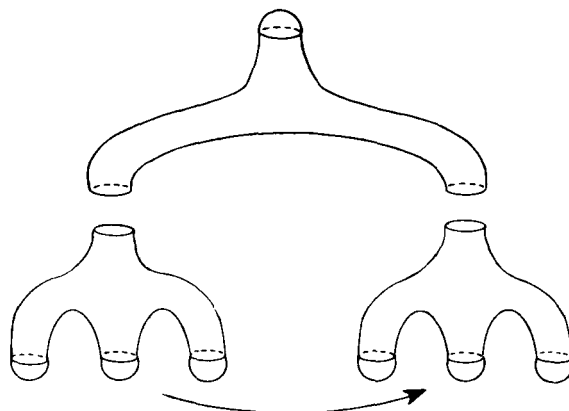

FIGURE 1

Proof of assertion B. Suppose f is a Smale diffeomorphism with the desired dynamics. Take a special filtration and for each j collapse to circles all annular components of cl $(M_j \setminus f(M_j))$ which do not meet $\mathcal{R}$. Consider the atom N_i containing the basic set Σ with reduced homology zeta function $p(t)$. Since $p(t) = 1+t^2+t^4$ the period of Σ is either one or two. Consequently N_i has either one or two components. We will outline the argument for both of these cases.

Suppose N_i has two components. We now examine the boundary components. Fill each boundary component with a disk containing exactly one sink or source so as to produce a Smale diffeomorphism which permutes two copies of S^2. The reduced homology zeta function of this map is $1/(1+t^2)^2$. If the periods of the

boundary components are given by r_i's, then each r_i is even and we have

$$\frac{1}{(1+t^2)^2} = \frac{(1+t+t^2)^2}{\prod(1+t^{r_i})}$$

It follows that the product consists of the two elements $1+t^2$ and $1+t^6$. Thus there is a boundary component of period 6. Since genus$(S^2) = 0$ there must be another atom with 6 components. This is impossible.

Suppose N_i is connected. Adding disks again to the boundary yields the following relationship with the boundary periods represented by r_i's

$$\frac{1}{(1+t)^2} = \frac{(1+t+t^2)^2}{\prod(1+t^{r_i})} = \frac{(1+t+t^2)^2}{(1+t^3)^2}.$$

It follows that N_i has two boundary orbits of period three. From the illustration we see that there must be two atoms, each with three components, to share these boundaries. However the only such atoms correspond to the period three sink orbits. This would produce a Smale diffeomorphism of S^2 with no sources.
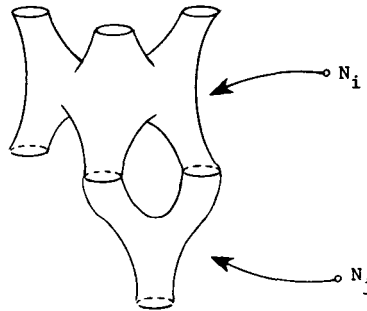

FIGURE 2

Proof of theorem 2. Let f be a Smale diffeomorphism of S^2. For each j collapse all annular components of cl $(M_j \setminus f(M_j))$ which do not meet $\mathcal{R}$. Let $\mathcal{S}$ be the set of boundary curves of the atoms. Construct a graph whose vertices correspond to components of $S^2 \setminus \mathcal{S}$ and whose edges correspond to common boundaries between components. Since S^2 is simply connected the graph is a tree. The map f induces an automorphism of this tree. We associate each basic set to the vertex orbit obtained from its atom. Franks formula for the reduced zeta function yields (c).

Conversely assume we are given a collection of data, a graph, and an automorphism satisfying (a), (b), and (c). Use lemma 5.5 of [B–S] to direct the edges of the graph. Let v be a vertex with two or more incident edge orbits and let q be the period of v. Suppose $r_1, \ldots, r_n$ are the periods of the edges directed into v and $s_1, \ldots, s_m$ are the periods of the edges directed out of v. Let

$$p(t) = \frac{\prod(1+t^{r_i})\prod(1+t^{s_j})}{1+t^{2q}}.$$

Since each r_i and s_j is a multiple of q there is a polynomial p' such that $p(t) = p'(t^q)$.

From [**Fr**] there is a Smale diffeomorphism, f', of S^2 with a single mixing basic set, Σ', of index one and reduced data

$$\left\langle \left\{ \frac{s_1}{q}, \ldots, \frac{s_m}{q} \right\}, \left\{ \frac{r_1}{q}, \ldots, \frac{r_n}{q} \right\}, \{p'\} \right\rangle.$$

Let f be the map on the disjoint union of q spheres with the property that f permutes the sphere components and f^q is conjugate to f' on each sphere. The index-one basic set of f has period q and reduced homology zeta function $p(t)$. Repeat this construction on each vertex of valence greater than one. For the remaining vertex orbits construct northpole-southpole maps which permute q copies of S^2 where q is the period of the vertex. Now for each edge in the graph remove the corresponding sink on one sphere and source on the other. Since the graph is a tree gluing the pieces together yields a sphere and a Smale map which realizes the data. □

We now consider surfaces of arbitrary genus. Given such a surface the following theorem establishes the existence of a subshift whose realization in a Smale diffeomorphism severely restricts the dynamics of the other basic sets.

THEOREM 3. *Let M_g be an oriented surface of genus g and suppose both p_1 and p_2 are odd prime numbers with $p_1 > p_2 > 2g + 1$. Then there is a subshift Σ such that if f is a Smale diffeomorphism on M_g and Σ is a basic set of f then every other basic set of f*

(1) *has a period which is divisible by p_1 or p_2; and*

(2) *can be realized as a basic set of a Smale diffeomorphism of the disk.*

Furthermore the hypothesis is non-vacuous in that there exist Smale diffeomorphisms of M_g which contain Σ as a basic set.

Proof. Pick $g \geq 0$ and odd primes p_1, p_2 with $p_1 > p_2 > 2g + 1$. Let $Q(t)$ be a reciprocal $\mathbb{Z}_2$-polynomial of degree $2g$ with $Q(1) = 1 = Q(0)$. Define

$$P(t) = \frac{(1 - t^{p_1})(1 - t^{p_2})Q(t)}{(1 - t)^2}.$$

Then

$$\frac{Q(t)}{(1 - t)^2} = \frac{P(t)}{(1 - t^{p_1})(1 - t^{p_2})}.$$

From [**Fr**] there is a homeomorphism h of M_g such that $Q(t) = \det(I - \alpha t)$ where α is the map induced on first $\mathbb{Z}_2$-homology by h and there is a Smale diffeomorphism which is isotopic to h with reduced data $\langle \{p_1\}, \{p_2\}, \{P(t)\}\rangle$. The basic set corresponding to $P(t)$ is mixing and will be denoted by Σ.

Now suppose f is a Smale diffeomorphism of M_g which contains Σ as one of its basic sets. Let N be the atom associated to Σ. Since Σ is mixing, N is connected. Let $m_1, \ldots, m_j$ be the periods of the boundary orbits of N. Attach a disk to each boundary and extend f so that there is precisely one sink or source in each disk. Let β be the map on first $\mathbb{Z}_2$-homology and define $R(t) = \det(I - \beta t)$. Note that

$\deg(R(t)) \leq 2g$. Then

$$\frac{R(t)}{(1-t^2)} = \frac{P(t)}{\prod(1-t^{m_i})}.$$

So

$$\prod(1-t^{m_i})R(t) = (1-t^{p_1})(1-t^{p_2})Q(t). \qquad (*)$$

We analyze $(*)$ in the algebraic closure of $\mathbb{Z}_2$. The right side has $p_1 - 1$ primitive p_1-roots of unity and $\deg(R) < p_1 - 1$. Consequently for some i_0, $1 - t^{m_{i_0}}$ has a primitive p_1-root of unity. It follows that p_1 divides m_{i_0} for some index i_0.

If $m_{i_0} \neq p_1$ produce an odd number r and integer d such that $m_{i_0} = 2^d r p_1$. Then $(1 - t^{m_{i_0}}) = (1 - t^{rp_1})^{2^d}$. If $r > 1$ then the degree of the left side of $(*)$ is greater than the degree of the right side of $(*)$. If $r = 1$ and $d \geq 1$ divide a factor of $(1 - t^{p_1})$ from each side. Then there are at least $p_1 - 1$ primitive p_1-roots of unity on the left side but $\deg(Q) < p_1 - 1$. We have established that $m_{i_0} = p_1$. A similar argument yields $p_2 = m_{i_1}$ for some index i_1.

Since $Q(1) = 1$ the right side of $(*)$ has a root of order two at one. However one is a root of each factor in the product on the left. Hence there are precisely two factors in the product (i.e. $(1 - t^{p_1})$ and $(1 - t^{p_2})$).

Consequently $R(t) = Q(t)$ and the genus of N is equal to the genus of M_g. The set $M_g - N$ must be a disjoint union of disks each of which is part of an orbit of period either p_1 or p_2. □

4. Pseudo-Anosov isotopy classes

In this section we analyze the basic sets of a Smale diffeomorphism in a pseudo-Anosov isotopy class on an oriented surface. It will be shown that any such map possesses a mixing, index-one basic set which is distinguished from the other basic sets by the following property.

Definition. A basic set Σ *possesses distinct Nielsen classes of periodic points* provided there exist $k \in \mathbb{Z}^+$ and two points in the same mixing component of Σ such that the points are fixed under f^k and in distinct Nielsen classes.

THEOREM 4. *If f is a Smale diffeomorphism in a pseudo-Anosov class on an oriented surface then*:

(1) *There is exactly one basic set Σ_0 which possesses distinct Nielsen classes of periodic points.*

(2) *Σ_0 is mixing.*

(3) *All other basic sets can be realized on a disk.*

(4) *f is isotopic to a Smale diffeomorphism for which Σ_0 is the only index-one basic set.*

Proof. Consider a filtration of f from §1. The set of homotopy classes represented by boundary curves of this filtration is finite and f-invariant. Since f is in a pseudo-Anosov class any finite order curve must be trivial. Thus every boundary curve bounds a disk. Since $M \neq S^2$ this disk is unique. Let $\mathcal{D}$ be the set of disks bounded by boundary curves. $\mathcal{D}$ has a finite number of maximal elements and these

are disjoint. Let $D_1, \ldots, D_m$ denote the maximal disks. $M \setminus \bigcup_{i=1}^{m} D_i$ is an atom containing a single basic set Σ_0.

Property (3) is now immediate. To establish (4) modify f on the disks $D_1, \ldots, D_m$ so that each disk contains exactly one sink or source.

For basic sets other than Σ_0 each mixing component is contained in a disk and thus all fixed points of f^k are in the same Nielsen class. However the total number of distinct Nielsen classes of fixed points in M of f^k grows exponentially with k. Hence for sufficiently large k there are two distinct Nielsen classes in the same component of Σ_0.

It remains to show that Σ_0 is mixing. Assume that Σ_0 has i mixing components and consider f^i. This map is in a pseudo-Anosov class and has i basic sets which each possess distinct Nielsen classes of periodic points. By the uniqueness of (1), $i = 1$. $\qquad\square$

REFERENCES

[B-F] P. Blanchard & J. Franks. An obstruction to the existence of certain dynamics in surface diffeomorphisms. *Ergod. Th. & Dynam. Sys.* **1** (1981), 255–260.
[B-S] S. Batterson & J. Smillie. Filtrations and periodic data on surfaces. To appear in *Amer. Jour. Math.*
[F] J. Franks. *Homology and Dynamical Systems. CBMS Regional Conf. Series*, **49** (1982).
[Fr] D. Fried. Subshifts on surfaces. *Ergod. Th. & Dynam. Sys.* **2** (1982), 15–21.
[J-S] W. Jaco & P. Shalen. Surface homeomorphisms and periodicity. *Topology* **16** (1977), 347–367.
[W] R. Williams. Classification of subshifts of finite type. *Ann. of Math.* **98** (1973), 120–153; Errata **99** (1979), 370–381.