Subshifts on surfaces

DAVID FRIED

Department of Mathematics, University of California, Santa Cruz, California 95064

Abstract. We determine when a polynomial is the reduced zeta function of a basic set of a Smale diffeomorphism of a compact surface.

This paper is a continuation of the work of Blanchard and Franks on the problem of which subshifts of finite type Σ_A cannot arise as basic sets on surfaces [1, 2]. Our results essentially exhaust their technique in that we find precisely which polynomials over $\mathbb{Z}_2$ are reduced zeta functions for some 0-dimensional basic set of a structurally stable surface diffeomorphism (see theorems 2 and 3 below).

In § 3 of this paper we determine a lower bound for the dimension of the first mod 2 homology group of a connected surface S that supports a mixing Σ_A (theorem 4). This gives the first known examples of subshifts supported on some connected closed surfaces but not on all (proposition 3).

We thank Paul Blanchard and John Franks for their encouragement.

1. Reciprocality of reduced zeta functions

We call a non-zero rational function $R(t)$ over a field F, $R \in F(t)^*$, *reciprocal* if $R(t^{-1}) = t^k R(t)$ for some $k \in \mathbb{Z}$. One sees easily that reciprocal functions form a subgroup of $F(t)^*$. The relevance of this notion for dynamical systems stems from Poincaré duality as in the following theorem. All the properties of reduced zeta functions that we use are proved in [2].

PROPOSITION 1. *Suppose $f: S \to S$ is an embedding of a compact surface. Then the reduced Lefschetz zeta function $\bar{\zeta}_2(f) \in \mathbb{Z}_2(t)$ is reciprocal. The same holds for the relative reduced Lefschetz zeta function of a pair of surfaces.*

Proof. First suppose S has no boundary. Then f preserves the non-singular intersection pairing $\langle \, , \, \rangle : V \times V \to \mathbb{Z}_2$ where $V = H_1(S; \mathbb{Z}_2)$. Thus, choosing a basis for V, the matrices J representing $\langle \, , \, \rangle$ and the matrix α representing $f_{*1}: V \to V$ satisfy $\alpha' J \alpha = J$.

Let $q(t) = \det(I - \alpha t)$. Since J and α are invertible,

$$q(t) = \det(I - \alpha' t) = \det(I - J\alpha^{-1}J^{-1}t) = \det(I - \alpha^{-1}t)$$

and

$$q(t^{-1}) = \det(I - \alpha t^{-1}) = \det(\alpha t^{-1}) \det(I - \alpha^{-1}t) = t^k q(t).$$

Thus $q(t)$ is reciprocal.

The reduced Lefschetz zeta function is expressed in terms of the transformations

$$f_{*i}: H_i(S; \mathbb{Z}_2) \rightleftarrows$$

by

$$\tilde{\zeta}_2(t) = \frac{\det(I - tf_{*1})}{\det(I - tf_{*0})\det(I - tf_{*2})}.$$

The denominator is a product of reciprocal terms of form $1 - t^n$ and the numerator is $q(t)$. So the theorem is shown in this case.   —

Second, suppose only that f is homotopic to a diffeomorphism g. Let S' be the surface obtained from S by coning off each circle in ∂S. Then g extends to a diffeomorphism $h: S' \rightarrow S'$. The ratio $\tilde{\zeta}_2(h)\tilde{\zeta}_2(g)^{-1}$ is the relative zeta function for the map of pairs $h: (S', S) \rightleftarrows$, hence it is a product of terms $(1 - t^n)^{-1}$. As $\tilde{\zeta}_2(h)$ is reciprocal by the case considered above and $\tilde{\zeta}_2(f) = \tilde{\zeta}_2(g)$ we see that $\tilde{\zeta}_2(f)$ is reciprocal.

Finally, let f be any self-embedding of S. We may approximate f by an embedding

$$f_0: S \rightarrow \mathrm{int}(S).$$

We now argue as in [1: lemma 2]. We define a larger surface S' by coning off those boundary components $C \subset \partial S$ for which $f_0 C$ bounds a disk in the subsurface $\overline{S - f_0 S}$. Then f_0 extends to an embedding

$$g: S' \rightarrow \mathrm{int}(S')$$

such that no component B of the subsurface $\overline{S' - gS'}$ is a disk. By considering the Euler characteristic one shows that such B are annuli with one boundary component in $\partial S'$ and the other in $g(\partial S')$ [1]. Thus g is homotopic to a diffeomorphism.

Now the previously proved case of our theorem shows $\tilde{\zeta}_2(g)$ is reciprocal. The ratio $\tilde{\zeta}_2(g)\tilde{\zeta}_2(f_0)^{-1}$ is the relative zeta function of the map of pairs $g: (S', S) \rightleftarrows$ but this is clearly one. Thus

$$\tilde{\zeta}_2(f) = \tilde{\zeta}_2(f_0) = \tilde{\zeta}_2(g)$$

is reciprocal, proving our theorem in the absolute case.

For a pair of compact surfaces (M, N), the relative zeta function is the ratio of two absolute zeta functions.   □

We will now extend the above result to the reduced Artin–Mazur zeta function in the Axiom A case. Let S be a compact surface, $f: S \rightarrow S$ an Axiom A embedding and

$$(M, N), \quad N \subset \mathrm{int}\, M,$$

a pair of invariant compact subsurfaces with

$$f(M) \subset \mathrm{int}(M), \quad f(N) \subset \mathrm{int}\, N$$

(such pairs arise naturally when f has no cycles). Let Λ_i be the basic sets in $M - N$ and $\Lambda = \bigcup \Lambda_i$. Let $\zeta(t) \in \mathbb{Q}(t)$ be the Artin–Mazur zeta function for Λ,

$$\zeta(t) = \zeta(t, f | \Lambda),$$

and let $\zeta_2(t) \in \mathbb{Z}_2(t)$ be its mod 2 reduction.

THEOREM 1. *$\zeta_2(t)$ is reciprocal.*

Proof. $\zeta_2(t)^{-1}$ is the product of terms $(1 - t^n)^{\pm 1}$ arising from sources and sinks in Λ and the relative reduced Lefschetz zeta function for (M, N) by [2]. The latter is reciprocal by proposition 1. □

2. Surfaces supporting subshifts of finite type

Let A be a 0–1 matrix and Σ_A the associated subshift of finite type. We say that a compact surface S *supports* Σ_A if there is an embedding $g: S \to S$ and compact subsurfaces M_1 and $M_2 \subset \text{int } M_1$ such that

(1) $gM_i \subset \text{int } M_i, i = 1, 2$;
(2) the invariant set

$$\Lambda = \bigcap_{n>0} (g^n M_1) - \bigcup_{n>0} (g^{-n} M_2)$$

is topologically conjugate to Σ_A and is a basic set of unstable dimension 1;
(3) Λ meets every component of S.

By taking M_1 smaller if necessary we can arrange that every component of the region $R = \overline{M_1 - M_2}$ meets Λ. We shall assume from now on that this holds.

The following immediate corollary of theorem 1 generalizes the main result of [1].

THEOREM 2. *If S supports Σ_A and $p_A(t)$ is the* mod 2 *reduction of*

$$\det (I - tA), \qquad p_A(t) \in \mathbb{Z}_2[t],$$

then $p_A(t)$ is reciprocal.

Blanchard and Franks proved that $p_A(t)$ cannot have odd degree unless $p_A(1) = 0$, and this is clear from theorem 2.

We now prove the converse of theorem 2 and show that there are no further non-trivial restrictions on these polynomials $p_A(t)$. The following proposition tells when certain periodic data is realizable in a given mapping class.

PROPOSITION 2. *Let S be a closed connected surface. Let f be a representative for a mapping class on S and let*

$$\alpha = f_{*1} : H_1(M; \mathbb{Z}) \supset.$$

Suppose $\{m_i\}_{i \in I}, \{n_j\}_{j \in J}$ are finite non-empty collections of positive integers and that

$$p(t) \in \mathbb{Z}_2[t].$$

If one has

$$\frac{\det (I - t\alpha)}{(1 - t)^2} = \frac{p(t)}{\prod (1 - t^{m_i}) \prod (1 - t^{n_j})} \quad in \ \mathbb{Z}_2(t) \qquad (*)$$

then there is a fitted diffeomorphism F isotopic to f such that the m_i are the periods of the periodic sources of F, the n_j are the periods of the periodic sinks and the rest of the non-wandering set of F consists of one mixing basic set Σ_A with $p_A(t) = p(t)$.

Note. If F satisfies the conclusions of the theorem then (*) is the zeta function form of the Lefschetz formula mod 2 [2], so (*) is necessary as well as sufficient.

Proof of proposition 2. It is well known that one can produce a fitted diffeomorphism F_0 isotopic to f with $\{m_i\}_{i \in I}$ as periods of sources and the $\{n_j\}_{j \in J}$ as periods of sinks: one chooses a handle decomposition for S with $(\#J)$ 0-handles and $(\#I)$ 2-handles and then one fits f appropriately [4]. Let

$$\varnothing \subset M_0 \subset M_1 \subset M_2 = S_g$$

be the filtration associated to this handle decomposition. Applying the Lefschetz formula to F_0 and using (*), one sees that $p(t)$ is the relative reduced Lefschetz zeta function for

$$F_0: (M_1, M_0) \leftrightarrows$$

[2]. As the basic sets $\Lambda_k \subset M_1 - M_0$ have unstable dimension 1, $p(t)^{-1}$ is the reduced zeta function

$$\zeta_2(t, F_0 | \bigcup \Lambda_k)$$

by [2]. Consequently if there is only one Λ_k and it is mixing then we are done [2].

This can be arranged by isotoping F_0 to a different fitted diffeomorphism F with the same sources and sinks and an appropriate geometric intersection matrix $G(F)$ on $M_1 - M_0$. We need

LEMMA 1. *Given a handle decomposition of S_g, there is a fitted diffeomorphism $h: S_g \leftrightarrows$ isotopic to the identity with fixed sinks and sources that has positive geometric intersection matrix $G(h)$ in dimension one.*

Proof of lemma 1. Let h be the composition in any order of fitted diffeomorphisms F_{ij}, $i \neq j$, that wrap the ith 1-handle through the jth, $i \neq j$, and leave everything else unchanged. The geometric intersection matrix for F_{ij} is

$$G(F_{ij}) = I + 2E_{ij},$$

E_{ij} an elementary matrix, and the product of these matrices (in any order) is positive. Choosing the same order to define h, one obtains the geometric intersection matrix $G(h)$, proving the lemma.   □

To finish the proof of proposition 2, take $F = hF_0$. Then every non-zero column in the geometric intersection matrix

$$G(F) = G(h)G(F_0)$$

is positive. It follows that there is only one basic set in $M_1 - M_0$ and it is mixing. □

We now state our converse to theorem 2.

THEOREM 3. *Given a reciprocal polynomial $p(t) \in \mathbb{Z}_2[t]$ with constant term 1, there is a connected oriented surface S and a mixing subshift of finite type Σ_A supported on S with $p_A(t) = p(t)$.*

Proof. By proposition 2, it suffices to find S, a mapping class on S, and sets $\{m_i\}$, $\{n_j\}$ such that (∗) holds.

If $p(t)$ has odd degree $2g+1$, then by reciprocality $p(1) = 0$. Hence we have

$$p(t) = q(t)(t-1), \qquad q(t) \in \mathbb{Z}_2[t].$$

Then we choose $m_1 = 1$, $n_1 = 2$ and we take no other m_i, n_j.

If $p(t)$ has even degree $2g$, let $q(t) = p(t)$. Then we choose $m_1 = 1$, $n_1 = 1$ and we take no other m_i, n_j.

In either case, we need only solve

$$\det (I - \alpha t) = q(t) \qquad\qquad (\ast\ast)$$

for α, where $q(t)$ is a given reciprocal polynomial over $\mathbb{Z}_2$ of degree $2g$ and α is the map induced on the first $\mathbb{Z}_2$-homology V of the closed oriented surface S_g of genus g by some homeomorphism f.

Choose a standard basis for V so that the intersection pairing is represented by the block matrix

$$\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = J.$$

If β is symmetric then

$$\alpha = \begin{pmatrix} 0 & I \\ -I & \beta \end{pmatrix}$$

satisfies the condition $\alpha^t J \alpha = J$. It follows that α is induced by some homeomorphism of S_g [3: p. 178]. Also one easily checks that

$$\det (I - \alpha t) = t^g \cdot \det (\beta - (t + t^{-1})I).$$

So to solve (∗∗) we need only prove

LEMMA 2. *Given a reciprocal polynomial* $q(t) \in \mathbb{Z}_2[t]$ *of degree* $2g$ *with constant term* 1 *there is a symmetric* $g \times g$ *matrix* β *over* $\mathbb{Z}_2$ *with*

$$\det (\beta - (t + t^{-1})I) = q(t) \cdot t^{-g}.$$

Proof. One sees easily that $q(t) \cdot t^{-g}$ is uniquely expressible as $r(t + t^{-1})$ where $r(t) \in \mathbb{Z}_2[t]$ has degree g. Thus it suffices to solve $\det (\beta - t) = r(t)$ for β, given a degree g polynomial $r(t) \in \mathbb{Z}_2[t]$.

Using direct sums, one may assume without loss of generality that $r(t)$ is irreducible. Let F be the extension field of $\mathbb{Z}_2$ obtained by adjoining a root ξ of r. As F is finite, this extension is separable and so the $\mathbb{Z}_2$-valued symmetric bilinear form

$$\langle x, y \rangle = \mathrm{Tr}\,(xy), \qquad x, y \in F,$$

is non-singular. For this bilinear form, the linear transformation $\beta(x) = \xi x$ of F is symmetric and has characteristic polynomial $r(t)$. As $\langle \ , \ \rangle$ is a non-singular symmetric bilinear form over $\mathbb{Z}_2$ and is not alternating, there is a basis for F over $\mathbb{Z}_2$ for which $\langle \ , \ \rangle$ is represented by the identity matrix. Then β is given by a symmetric matrix with characteristic polynomial $r(t)$. This completes the proof of theorem 3. $\square$

We recently received an example from John Franks that shows that the characteristic polynomial of an integral symplectic matrix can be any monic reciprocal polynomial over $\mathbb{Z}$ of even degree. This gives another solution to (**) in the preceding proposition.

3. Bounds on the first Betti number

The results of § 1 and § 2, and those of earlier investigators, do not restrict the topological type of a surface S that supports a given Σ_A. To obtain such restrictions we will keep track of first Betti numbers as we simplify the surface by the methods in [1].

Since S supports Σ_A iff the associated closed surface supports Σ_A, we do not lose generality by assuming $\partial S = \varnothing$. For convenience we will only consider mixing subshifts Σ_A, so we may suppose S is connected.

We introduce some algebraic terminology to describe our restrictions on S. Let $\mathscr{C} \subset \mathbb{Z}_2[t]$ be the collection of all finite products

$$\prod (1 - t^i)^{a_i}, \qquad a_i \geq 0.$$

For $p(t) \in \mathbb{Z}_2[t]$ we write dist $(p(t), \mathscr{C})$ (the *distance* from $p(t)$ to the *cyclic* polynomials) for the least value of deg $(q(t))$ over all solutions $q(t) \in \mathbb{Z}_2[t]$ to

$$p(t) = q(t)c(t), \quad c(t) \in \mathscr{C}.$$

As previously we write $p_A(t)$ for the $\mathbb{Z}_2$ reduction of det $(I - tA)$. Then we find the following constraint on a surface supporting a given subshift.

THEOREM 4. *Let Σ_A be a mixing subshift of finite type. If S is a closed connected surface that supports Σ_A then the first mod 2 Betti number $\beta_1(S)$ and $p_A(t)$ satisfy*

$$\beta_1(S) \geq \text{dist} ((1-t)^2 p_A(t), \mathscr{C}).$$

Remark. The right-hand side is easily computed for given A.

Proof of theorem 4. Let (M, N) be a filtration pair for $\Sigma_A \subset S$. As Σ_A is mixing, we may suppose $M - N$ is connected. Let T be the connected surface obtained from the region $R = \overline{M - N}$ by capping off the boundary components of R. We can compare the Euler characteristics of S and T.

LEMMA 3. $\chi(T) \geq \chi(S)$.

Proof. Let $k(\cdot)$ denote the number of boundary components of a surface. As disks have Euler characteristic one, we have

(a) $\chi(T) = \chi(R) + k(R)$.

Let $P = \overline{S - R}$. Then P is a compact surface and every component B of P meets R. This gives $\chi(B) \leq k(B)$ with equality iff B is a disk. Summing over all B's gives

(b) $\chi(P) \leq k(P) = k(R)$.

Also

(c) $\chi(S) = \chi(R) + \chi(P)$.

Combining (a), (b) and (c) gives

$$\chi(S) = \chi(R) + \chi(P) \leq \chi(R) + k(R) = \chi(T),$$

proving the lemma.   □

In [1] it was shown that T carries an Axiom A-No Cycles diffeomorphism $f: T \supseteq$ whose only basic set of unstable dimension 1 was Σ_A. Then, as in the remark following proposition 2, the Lefschetz formula gives

$$\frac{\det (I - \alpha t)}{(1-t)^2} = \frac{p_A(t)}{\prod (1-t^{m_i}) \prod (1-t^{n_j})} \qquad (*)$$

where α is the action of f on $H_1(T; \mathbb{Z}_2)$, the m_i are the periods of the sources of f and the n_j are the periods of the sinks. Multiplying this out gives

$$\text{dist} (p_A(t)(1-t)^2, \mathscr{C}) \leq \deg (\det (I - \alpha t)) = \beta_1(T).$$

But by lemma 3,

$$\beta_1(T) = 2 - \chi(T) \leq 2 - \chi(S) = \beta_1(S).$$

Combining the last two inequalities proves the theorem. $\qquad \square$

Based on theorem 4, we can give an example of a subshift of finite type Σ_A that is supported on T^2 but not S^2. Surely analogous examples can be found for higher genus surfaces by the same techniques.

We have

PROPOSITION 3. *Let $\alpha \in Gl(2, \mathbb{Z})$ be a hyperbolic toral automorphism with an odd number of fixed points and at least two distinct orbits of period 3. Let $f: T^2 \supseteq$ be the Axiom A-No Cycles diffeomorphism obtained by double DA-ing α along two period 3 orbits. Then f has a unique basic set Λ of unstable dimension 1, $\Lambda \cong \Sigma_A$ for some A and Σ_A is not supported on S^2 or $\mathbb{R}P^2$.*

Proof. Except for the last statement, our conclusions follow from general theory. We now compute $p_A(t)$. Over $\mathbb{Z}_2$, we have

$$\frac{p_A(t)}{(1-t^3)^2} = \frac{\det (I - t\alpha)}{(1-t)^2}$$

since f has a unique source and sink orbit, each of period 3. Our assumption that α has an odd number of fixed points gives that $\text{Tr } \alpha$ is odd so

$$\det (I - t\alpha) = 1 + t + t^2 \quad (\text{mod } 2).$$

We obtain $p_A(t) = (1 + t + t^2)^3$.

The distance from $(1-t)^2 p_A(t) = (1-t^3)^2(1+t+t^2)$ to $\mathscr{C}$ is easily seen to be 2. So theorem 4 implies $\beta_1(S) \geq 2$ for any closed connected surface S supporting Σ_A and thus

$$S \neq S^2, \quad S \neq \mathbb{R}P^2. \qquad \square$$

REFERENCES

[1] P. Blanchard & J. Franks. An obstruction to the existence of certain dynamics in surface diffeomorphisms. *Ergod. Th. and Dynam. Sys.* **1** (1981), 255–260.
[2] J. Franks. A reduced zeta function for diffeomorphisms, *Amer. J. Math.* **100** (1978), 217–243.
[3] W. Magnus, A. Karass & D. Solitar. *Combinatorial Group Theory.* Dover: New York, 1976.
[4] M. Shub & D. Sullivan. Homology theory and dynamical systems. *Topology* **14** (1975), 109–132.