Ergod. Th. & Dynam. Sys. (1983), **3**, 315–334
Printed in Great Britain

Periodic points of surface homeomorphisms with zero entropy

JOHN SMILLIE
Graduate Center – CUNY, 33 *West* 42 *Street, New York, N.Y.* 10036

(*Received* 16 *June* 1982)

Abstract. This paper deals with the question of which periods can occur as periods of periodic points of zero entropy surface homeomorphisms in a given isotopy class. We give new examples of isotopy classes for which there are non-trivial restrictions and describe how the possible periods can be computed. Certain phenomena occur only for surfaces of large genus. These results have applications to the periodic data question for Morse–Smale maps.

0. *Introduction*

Recent results in the field of dynamical systems establish interesting relationships between the periods of periodic points and the dynamical complexity in low dimensional dynamical systems. An important measure of complexity is the topological entropy. The entropy of a map f is a global measure of the stretching caused by f. In [**2**] it is shown that there are restrictions on the periods of periodic points for self maps of the interval with zero entropy. In [**1**] and [**6**] it is shown that similar restrictions apply for orientation reversing homeomorphisms of surfaces.

In this paper we discuss arbitrary homeomorphisms of surfaces. Our main proposition implies that a set of periods of periodic points of some homeomorphism f with zero entropy can be realized by an f' isotopic to f where f' has a very simple form described in § 2. We use this result to show that for almost all surfaces there are isotopy classes of homeomorphisms which do not contain, for example, zero entropy maps with fixed points. In theorem 3.1 we give an explicit algorithm for computing possible periods of periodic points of zero entropy homeomorphisms isotopic to isometries.

Another line of investigation in dynamical systems for which this paper has relevance is the study of structurally stable diffeomorphisms. The structurally stable diffeomorphisms of surfaces with zero entropy are exactly the Morse–Smale diffeomorphisms. A basic invariant of such a diffeomorphism is the periodic data set which describes the period and type of perodic points. The conditions on periods described in this paper give conditions on periodic data sets that must be satisfied in order that they be periodic data sets of Morse–Smale maps in a given isotopy class. These conditions are not implied by the Morse inequalities or homology zeta function condition. In particular we get additional conditions in the case of orientation preserving maps of oriented surfaces and maps isotopic to the identity on non-orientable surfaces (compare with [**9**]). The existence of such conditions was not previously known. In a forthcoming joint paper with Steve Batterson we will describe necessary and sufficient conditions on periodic data sets that they be realizable by Morse–Smale diffeomorphism in a given isotopy class in these two cases.

The case of orientation preserving homeomorphisms of oriented manifolds is particularly interesting. We give here two propositions which illustrate fairly subtle behaviour in this case.

PROPOSITION 3.2. Let *M be the oriented surface of genus* 145. *There exists an orientation preserving homeomorphism* $f: M \to M$ *with zero entropy so that for any homeomorphism* f' *isotopic to* f *with zero entropy either*

(1) *the periods of all periodic points under* f' *are multiples of* 3, 5 *or* 14,

or (2) *the periods of all periodic points under* f' *are multiples of* 3, 7 *or* 10.

Furthermore, there exist f_1 *and* f_2 *isotopic to* f *with zero entropy for which any multiple of* 3, 5 *or* 14 *is a period of some periodic point under* f_1 *and any multiple of* 3, 7 *and* 10 *is a period of some periodic point under* f_2.

Remark According to this proposition there are different zero entropy homeomorphisms isotopic to f with points of period 5 or 7 but any single f' isotopic to f with points of period 5 and 7 has positive entropy.

PROPOSITION 3.3. *Let M be an oriented surface of genus less than* 145. *Let* f_1 *and* f_2 *be isotopic orientation-preserving homeomorphisms of zero entropy. We can construct an* f_3 *with zero entropy so that any n which is a period of a periodic point of* f_1 *or* f_2 *is also a period of a periodic point of* f.

Remark. This proposition shows that the phenomenon exhibited in proposition 3.2 cannot occur on a surface of lower genus.

1. In this section we collect facts about simple closed curves and describe Thurston's canonical form for surface homeomorphisms. A general reference is [**5**].

Let M be a compact, not necessarily oriented, surface with a hyperbolic metric. If M has a boundary we assume that the geodesic curvature of the boundary is zero.

Definition. Let $\mathcal{S}^+$ be the set of unoriented homotopy classes of simple closed curves in M which do not bound disks or Moebius bands. Let $\mathcal{S}$ consist of those curves in $\mathcal{S}^+$ which are not homotopic to boundary components. (Note: This definition differs from the definition in [**10**].)

LEMMA 1.1. *Every class in* $\mathcal{S}^+$ *is represented by a unique simple closed geodesic and every simple closed geodesic represents a class in* $\mathcal{S}^+$.

LEMMA 1.2. *Let* γ_1 *and* γ_2 *be geodesics representing distinct elements* $[\gamma_1]$, $[\gamma_2]$ *in* $\mathcal{S}^+$ *and let* c_1 *and* c_2 *be arbitrary curves representing these elements. Then*

$$\operatorname{card}\{\gamma_1 \cap \gamma_2\} \leq \operatorname{card}\{c_1 \cap c_2\}.$$

Definition. Let $i([\gamma_1], [\gamma_2]) = \operatorname{card}\{\gamma_1 \cap \gamma_2\}$.

Lemma 1.1 implies that this is a well defined function on $\mathcal{S}^+ \times \mathcal{S}^+$.

LEMMA 1.3. *If c is an oriented curve which does not bound a disk or Moebius band, then c is represented by a unique oriented geodesic.*

All of these assertions can be proved by considering the geometry of the universal covering space of M.

LEMMA 1.4. *Let c represent an element of $\mathcal{S}$ which has an annular neighbourhood U. If f is a homeomorphism with $f(c)=c$ and f is isotopic to the identity then f preserves the local orientation on U.*

Proof. We can assume that c is a geodesic. Let $\tilde{M}$ be the orientation cover of M. c lifts to two geodesics $\tilde{c}_1$ and $\tilde{c}_2$ in $\tilde{M}$. f lifts to a homeomorphism $\tilde{f}:\tilde{M}\to\tilde{M}$ isotopic to the identity. If f reverses the orientation of c then $\tilde{f}(\tilde{c}_1)=\tilde{c}_2$. But this would imply that $\tilde{c}_1$ is homotopic to $\tilde{c}_2$ which contradicts lemma 1.1. □

COROLLARY. *f as above preserves the normal direction.*

Proof. It follows from lemma 1.3 that f preserves the S^1 direction on U. Since f preserves the total orientation on U it must preserve the normal direction. □

Definition. A curve c representing an element of $\mathcal{S}^+$ is *normally flipped* by a homeomorphism g if c has an oriented normal bundle, g preserves the S^1 direction and g reverses the normal direction.

Thurston's canonical form. Thurston's theory of surface homeomorphisms allows us to isotope an arbitrary homeomorphism f to a homeomorphism $\check{f}$ which can be decomposed by an invariant family of curves Γ into isometry pieces and pseudo-Anosov pieces. In each pseudo-Anosov piece there is a curve c so that the length of the shortest geodesic representing $\check{f}^l c$ (or $f^l c$) grows exponentially with l. Thus the existence of a pseudo-Anosov piece implies that the entropy of f is positive.

Now let f be a homeomorphism with zero entropy on a surface of negative Euler characteristic. Let U be a disjoint union of ε-neighbourhoods of the curves in Γ. We may assume that $\check{f}$ is an isometry on the complement of U and that Γ is minimal in the sense that given an orbit of geodesics σ and an ε-neighbourhood of these geodesics $U_0\subset U$, $\check{f}|U_0$ is not homotopic to an isometry relative to $M-U_0$.

The following corollary gives a canonical form for $\check{f}|U$.

LEMMA 1.5. *Let A be an annulus (respectively, Moebius band). Let $f:A\to A$ be a homeomorphism which preserves length on ∂A. Let $\tilde{A}$ be the universal cover of A. $\tilde{A}=[-1,1]\times\mathbb{R}$ where the covering translations are generated by $T(x,y)=(x,y+1)$ (respectively, $T(x,y)=(-x,y+1)$). Then f is homotopic rel ∂ to a map f_1 on A covered by $\tilde{f}_1$ where*

$$\tilde{f}_1(x,y)=(\pm x,\pm y+\lambda x+c).$$

If B is a Moebius band then $\lambda=0$.

Proof. Let $\tilde{f}$ be a lift of f. Given that $\tilde{f}$ preserves length on each component of the boundary and either preserves direction on both components or reverses direction

https://doi.org/10.1017/S014338570000198X Published online by Cambridge University Press

on both components we can find a map of the form

$$L(x, y) = (\pm x, \pm y + \lambda x + c)$$

which induces the same map on $\partial\tilde{A}$. The linear homotopy

$$H_t(x, y) = t\tilde{f}(x, y) + (1-t)L(x, y)$$

on $\tilde{A}$, induces on A a homotopy rel ∂A from f to the map on A induced by L. If A is a Moebius band then the fact that L either commutes or anti-commutes with T gives $\lambda = 0$. If we choose a lift which preserves boundary components we can take the coefficient of x to be 1. □

COROLLARY. *If $f: A \to A$ is a homeomorphism with finite order on ∂A which is not homotopic rel ∂A to a homeomorphism of finite order then A is an annulus and f preserves orientation.*

Proof. If A is an annulus and the orientation is reversed, then the eigenvalues of the linear part of $\tilde{f}_1$ are $+1$ and -1. The linear part of $\tilde{f}_1^2$ is the identity. Thus $\tilde{f}_1^2$ is a translation and the corollary follows. If A is a Moebius band, then it is immediate that $\tilde{f}_1^2$ is a translation. □

It is possible to assume (using equivariant Teichmuller theory) that the metrics on different components of $M - U$ match up along the boundary curves to give a metric defined on all of M so that the curves in Γ are geodesics and U is an ε-neighbourhood of these curves in this metric. This assumption is convenient for technical reasons but it is not essential to our proofs.

Let Σ be a finite f invariant subset of M. It is shown in [**6**] that one can construct a Thurston canonical form for f relative to Σ. We now describe this construction. We can add boundary circles to the ends of the manifold $M - \Sigma$. Denote the new manifold by $\overline{M - \Sigma}$. By isotoping $f|M - \Sigma$ in ε-disk neighbourhoods of the points in Σ we may assume that $f|M - \Sigma$ extends to a homeomorphism f_1 on $\overline{M - \Sigma}$. We can put f in Thurston canonical form which we denote $\check{f}_1$.

According to [**6**] there are no pseudo-Anosov pieces in this decomposition. (It is shown that modifying f in a neighbourhood of Σ cannot introduce pseudo-Anosov pieces.)

Now add disks to the boundary components of $\overline{M - \Sigma}$. $\check{f}_1$ extends to these disks and we can identify the centres of these disks with the points in Σ. We denote this canonical form for f by f_Σ. Let Γ_Σ denote the set of twist curves in this decomposition.

The following lemma concerning Thurston canonical form will be used in § 2,

LEMMA 1.6. *Let α be a geodesic representing $[\alpha] \in \mathcal{S}$. If $[\alpha]$ has a finite order under f then either $\alpha \in \Gamma$ or α is disjoint from all geodesics $\gamma \in \Gamma$.*

Proof. Let t_γ denote the twist along γ. We have

$$\check{f}^l = t_{\gamma_1}^{n_1} \circ t_{\gamma_2}^{n_2} \circ \cdots \circ t_{\gamma_k}^{n_k} \qquad \text{where } n_i \geq 1.$$

If α does not satisfy the conclusion of the lemma then $i([\gamma_j], [\alpha])$ is not zero for some j. Since the different twist maps t_{γ_i} have disjoint support they commute. Thus the order of the γ_i is immaterial and we may assume that $i([\gamma_1], [\alpha]) \neq 0$. Write

$$f^l = t_{\gamma_1}^{n_1} \circ h.$$

There exists a curve β_1 so that $i([\gamma_i], [\beta_1]) \neq 0$ and $i([\gamma_j], [\beta_1]) = 0$ for $j \neq 1$.

If γ and β are represented by disjoint curves, then for any curve α, $t_\gamma^n[\alpha]$ is represented by a curve which intersects β the same number of times as α. Thus

$$i(t_\gamma^n[\alpha], [\beta]) \leq i([\alpha], [\beta]).$$

If we consider this formula with $[\alpha]$ replaced by $t_\gamma^n[\alpha]$ and n replaced by $-n$ we get the opposite inequality. We conclude that

$$(1) \qquad i(t^n[\alpha], [\beta]) = i([\alpha], [\beta]).$$

We have the following formula from [**5**]:

$$(2) \qquad |i(t_\gamma^n[\alpha], [\beta]) - ni([\alpha], [\gamma]) \cdot i([\beta], [\gamma])| \leq i([\alpha], [\beta]).$$

Replacing $[\alpha]$ by $h^m[\alpha]$, $[\beta]$ by $[\beta_1]$, $[\gamma]$ by $[\gamma_1]$, and n by mn_1 gives

$$(3) \quad |i(f^{m\cdot l}[\alpha], [\beta_1]) - m \cdot n_1 \cdot i(h^m[\alpha], [\gamma_1]) \cdot i([\beta_1], [\gamma_1])| \leq i(h^m[\alpha], [\beta_1]).$$

Using (1) repeatedly we have $i(h^m[\alpha], [\gamma_1]) = i([\alpha], [\gamma_1])$ and $i(h^m[\alpha], [\beta_1]) = i([\alpha], [\beta_1])$. Substituting in (3) gives

$$(4) \qquad |i(f^{m\cdot l}[\alpha], [\beta_1]) - m \cdot n_1 \cdot i([\alpha], [\gamma_1]) \cdot i([\beta_1], [\gamma_1])| \leq i([\alpha], [\beta_1]).$$

By assumption $n \cdot i([\alpha], [\gamma_1]) \cdot i([\beta_1], [\gamma_1])$ is not zero. (4) implies that as m varies $i(f^{m\cdot l}[\alpha], [\beta_1])$ must assume infinitely many distinct values, therefore the orbit of $[\alpha]$ under the action of f is infinite. □

We end this section with a discussion of flipped periodic points and their relation to Thurston canonical form.

Definition. Let h be a homeomorphism. An h-periodic point x of period p is *flipped* if h^p reverses the local orientation at x.

If h is an orientation preserving homeomorphism of an oriented manifold then no periodic points are flipped. If h is an orientation reversing homeomorphism of an oriented manifold then a periodic point is flipped if and only if its period is odd. If the manifold is not oriented there are no *a priori* restrictions on the existence or periods of flipped or unflipped periodic points.

Let f_Σ be a homeomorphism in relative Thurston canonical form. Note that f-flipped periodic points in Σ correspond to f_Σ-flipped periodic points.

Let $M_0 \subset M$ be the region where f_Σ is an isometry. $M - M_0$ is a disjoint union of annular neighbourhoods of curves in Γ_Σ. Let $x \in M_0$ be a flipped periodic point of period p. The fixed point set of an isometry of a Riemannian manifold is a closed submanifold $Y \subset M_0$. The tangent space to Y at x is the set of tangent vectors fixed by the derivative of the isometry at x. In the case at hand, Df_Σ^p is an isometry of $\mathbb{R}^2$ which reverses orientation. The fixed point set of Df_Σ^p must therefore be a one-dimensional subspace. Let X be the component of Y containing x. By the above reasoning, X is a circle or interval with its endpoints contained in the boundary of M. If z is an endpoint then Df_Σ^p reverses the local orientation at z. On the other hand, f_Σ^p preserves an orientation on a neighbourhood of each curve in Γ_Σ. We conclude that X is a circle contained in M_0. Let G_r denote the set of circles of flipped periodic points of arbitrary period. Clearly, distinct circles in G_r are disjoint.

2. We will construct examples of smooth diffeomorphisms with zero entropy and many periodic points. To construct these examples we modify Thurston canonical form by putting rotational shear in neighbourhoods of simple closed curves. Assume we are given an f with zero entropy and a family Γ of curves homotopy invariant under f, containing the twist curves of f and such that $i(\gamma_1, \gamma_2) = 0$ for $\gamma_1, \gamma_2 \in \Gamma$ and $i(\gamma_1, \alpha) = 0$ when α is a reflective curve not contained in Γ. We can put f in Thurston canonical form and choose a hyperbolic metric on M which is preserved by f outside neighbourhoods of the twist curves. We can represent all the curves in Γ by geodesics. Because of the condition $i(\gamma_1, \gamma_2) = 0$ these geodesics will be disjoint. We can choose an ε so that ε-neighbourhoods of these curves are disjoint and do not meet reflective curves. We can assume that f is an isometry outside ε-neighbourhoods of the twist curves.

Let γ be a geodesic in Γ of setwise period n. Let A be the ε-neighbourhood of γ. $f^n(A) = A$. If we use the map f^{n-1} to identify $f^{n-1}(A)$ with A we can think of $f|f^{n-1}(A)$ as a map from A to itself. Let $f_0: A \to A$ denote this map. We will construct an $f_1: A \to A$ and then modify f by replacing f_0 by f_1 on $f^{n-1}(A)$. A point of period p under f_1 will correspond to a point of period np under f.

Case 1. A is an annulus. We can choose coordinates so that $\tilde{A} = [-1, 1] \times \mathbb{R}$ and the covering translations are generated by a map T such that $T(x, y) = (x, y + 1)$. Let $\tilde{f}_0$ be a lift of f_0 to $\tilde{A}$. We may also assume that f_0 acts on $\partial\tilde{A}$ by translation.

Subcase 1a. f_0 preserves orientation and boundary components. $\tilde{f}_0$ is homotopic rel $\partial\tilde{A}$ to a map of the form

$$\tilde{f}_0(x, y) = (x, y + \rho_0(x))$$

where ρ_0 is linear. Define

$$\tilde{f}_1(x, y) = (x, y + \rho_1(x))$$

where ρ_1 is a C^∞ function satisfying:

(a) for $x \in [-1, -1+\delta]$, $\rho_1(x) = \rho_0(-1)$;
for $x \in [1-\delta, 1]$, $\rho_1(x) = \rho_0(1)$.

(b) $\rho([-1, 1]) \subset \mathbb{R}$ has length at least 1.

Let f_1 be the function induced on A. The first condition guarantees that $f\rho_1$ will have a smooth extension to all of M and that f_1 is homotopic to f_0 rel ∂. The second condition guarantees that f_1 will have points of all periods: let c_{x_0} be the circle of points in A with first coordinate x_0; f_1 induces a rotation on c_{x_0} with rotation number $\rho_1(x_0)$ mod 1. Condition (b) guarantees that all rotation numbers will be represented and thus all periods.

Subcase 1b. f_0 preserves orientation but interchanges boundary components. We may assume that

$$\tilde{f}_0(x, y) = (-x, -y + \rho_0(x))$$

with ρ_0 linear. Let

$$\tilde{f}_1(x, y) = (-x, -y + \rho_1(x))$$

where ρ_1 is a C^∞ function satisfying condition (a) above and $\rho_1 = 0$ for x in a neighbourhood of zero. Since

$$\tilde{f}_1^2(x, y) = (x, y - \rho(x) + \rho(-x)),$$

$\tilde{f}_1$ has $(0, 0)$ as a fixed point and has order 2 in a neighbourhood of $(0, 0)$. We can add a twist in a small neighbourhood of the image of $(0, 0)$ (as in subcase 1a) to get points of all periods.

Subcase 1c. f_0 reverses orientation and interchanges the boundary components. We may assume that

$$\tilde{f}_0(x, y) = (-x, y + \rho_0(x))$$

where $\rho_0(x)$ is linear. Let

$$\tilde{f}_1(x, y) = (-x, y + \rho_1(x)),$$

where ρ_1 satisfies condition (a) and f_1 takes c_x to c_{-x}, thus points not on c_0 must have even period. Now

$$\tilde{f}_1^2(x, y) = (x, y + \rho(x) + \rho(-x)).$$

If the set $\{\rho(x) + \rho(-x) : x \in [-1, 1]\}$ has length greater than one we get points of every even period. We can get points of a single odd period n by requiring that $\rho(0) = 1/n$. These points will be flipped.

Note. The case in which f_0 preserves boundaries and reverses orientation need not be considered. In this case ∂A would cross a reflective curve.

Case 2. A is a Moebius band. We can choose coordinates so that $\tilde{A} = I \times \mathbb{R}$ and the covering translations are generated by a map T where $T(x, y) = (-x, y + 1)$.

Subcase 2*a*. f_0 takes a generator of $\pi_1(A)$ to itself. We may assume that

$$\tilde{f}_0(x, y) = (x, y + \rho_0(x))$$

where ρ_0 is constant. Let

$$\hat{f}_1(x, y) = (x, y + \rho_1(x))$$

where ρ_1 satisfies condition (a) and $\rho_1(x) = \rho_1(-x)$. This second condition is necessary for $\hat{f}_1$ to commute with T. For $x_0 \in [0, 1]$, let c_{x_0} be the circle in A corresponding to points with x coordinate $\pm x_0$. $f(c_{x_0}) = c_{x_0}$ and the induced map is a rotation with rotation number $\frac{1}{2}\rho(x_0)$. Assume that the image of ρ contains an interval of length 2 and that $\rho(0) = 1/n$. This gives a flipped point of period n and unflipped points of all periods.

Note. The case in which f_0 takes a generator of π_1 to its inverse need not be considered because again ∂A would cross a reflective curve.

Choose a finite set of unflipped points in the complement of the ε-neighbourhoods of γ which represent all periods for f restricted to this set. In a neighbourhood of these points we can twist around annuli as in subcase 1a to get periodic points of all periods.

The purpose of the main proposition which follows is to show that if a finite set of periods can be realized by an f with zero entropy then it can in fact be realized by an $f' \sim f$, where f' is one of the diffeomorphisms we have just constructed. Let Σ be a finite f-invariant set. Let f_Σ denote the relative canonical form for f. Let Γ_Σ be the set of twist curves for f_Σ.

Definition. Let Φ denote the set of 2-dimensional submanifolds $Y \subset M$ such that

(1) each boundary component of Y is an element of Γ_Σ;

(2) the Euler characteristic, $\chi(Y)$, is greater than or equal to zero.

Definition. Let Σ_A consist of those points in Σ which are either flipped or contained in some maximal annulus or Moebius band in Φ.

PROPOSITION 2.1. (MAIN PROPOSITION.) *Let M be a surface with $\chi(M) < 0$. Let $f: M \to M$ be a homeomorphism with zero entropy. Let $\check{f}$ be the Thurston canonical form for f. Then there is a function $\sigma: \Sigma_A \to \mathcal{S}$ such that*

(1) *the image of σ is represented by a set of disjoint curves;*

(2) *σ is f-equivariant, thus the f-period of an element in Σ_A is a multiple of the period of its image in $\mathcal{S}$;*

(3) *all flipped periodic points are assigned to geodesics which are either normally flipped or have Moebius band neighbourhoods and all flipped periodic points assigned to the same f-orbit in $\mathcal{S}$ have the same period;*

(4) *if $\sigma(x)$ is represented by a normally flipped geodesic, but x is not a flipped periodic point, then x has even order;*

(5) *the f-periods of points in $\Sigma - \Sigma_A$ are multiples of $\check{f}$-periods of points in M.*

COROLLARY 2.2. *The image of Σ_A in $\mathcal{S}$ is represented by a set of geodesics, each of which is disjoint from the $\check{f}$-twist curves or equal to an $\check{f}$-twist curve.*

Proof. By assertion (2), the image of Σ_A in $\mathcal{S}$ consists of $\check{f}$-periodic elements and these are represented by $\check{f}$-periodic geodesics. Lemma 1.3 yields the corollary. □

If M is oriented and f preserves orientation then (3) and (4) are vacuous. If f reverses orientation then (4) is true automatically and in assertion (3) the term 'flipped periodic point' may be replaced by the term 'periodic point of odd period'.

Proof of main proposition. Recall that G_r is the set of components of flipped periodic points of f_Σ.

LEMMA 2.3. *If an element of G_r bounds a disk, then $M = S^2$. If an element of G_r bounds a Moebius band, then $M = K^2$. If two elements of G_r bound an annulus then $M = T^2$ or K^2. If some element c of G_r is homotopic to a curve in Γ_Σ then c is contained in an element of Φ.*

Proof. Let $c \in G_r$ consist of flipped points of period p. Assume c bounds a manifold X, then $X \cup f_\Sigma^p X$ is a closed submanifold of M. It is therefore equal to M. If X is a disk or Moebius band then M is respectively a sphere or Klein bottle.

Let $c_0, c_1 \in G_r$ consist of flipped points of period p_0 and p_1. Assume that c_0 and c_1 together form the boundary of an annulus A. If $f_\Sigma^p c_1 = c_1$ then $A \cup f_\Sigma^{p_0} A$ is a torus or Klein bottle and we are done. If $f_\Sigma^{p_0} c_1 \neq c_1$ then the two curves are disjoint and $A \cup f_\Sigma^{p_0} A$ is an annulus, the boundary components of which consist of flipped points of period p_1. We can apply the previous argument to these curves and this annulus.

Assume that c is homotopic to a $\gamma \in \Gamma_\Sigma$. Thus c and γ form the boundary of an annulus A. $A \cup f_\Sigma^p A$ is an annulus with boundary equal to $\gamma \cup f_\Sigma^p \gamma$. Thus

$$c \subset A \cup f_\Sigma^p A \in \Phi.$$

This completes the proof of the lemma. □

Since $\chi(M) < 0$, the lemma implies that each element of G_r represents a distinct element of $\mathscr{S}$.

Definition. Let x be a flipped periodic point in Σ, then $\sigma(x)$ is the class in $\mathscr{S}$ represented by the element of G_r containing x.

There is a partial ordering on Φ given by inclusion. We are interested in elements of Φ which are maximal with respect to this ordering.

LEMMA 2.4. *Two maximal elements of Φ are either equal or disjoint.*

Proof. Let X and Y be elements of Φ which intersect. We will show that $X \cup Y$ is again an element of Φ. The boundary components of X and Y are either disjoint or coincident. If they coincide we may perturb them without changing $X \cup Y$ so that they are disjoint. After this perturbation $X \cap Y$ is a manifold with boundary. We have the formula

$$\chi(X \cup Y) = \chi(X) + \chi(Y) - \chi(X \cap Y).$$

If $\chi(X \cap Y)$ is non-positive then $\chi(X \cup Y)$ is non-negative and we are done. If $\chi(X \cap Y)$ is positive then $X \cap Y$ has a disk component. Let D be a disk component of $X \cap Y$. D is a component of the boundary of X or Y. Without loss of generality, assume $\partial D \subset \partial X$. Since $D \subset X$ as well, D must be open and closed in X, so $D = X$. Now $X \cup Y \subset Y$ and $Y \in \Phi$ so we are done. □

Let Y be a maximal element of Φ. Since $\chi(M) < 0$, Y is a proper submanifold of M. There are three cases: Y can be a disk, annulus, or Moebius band. Let $\mathscr{A}$ denote the set of maximal annuli and Moebius bands.

LEMMA 2.5. *There can be at most one component of G_r in each maximal annulus or Moebius band.*

Proof. In each annulus there is a unique unoriented non-trivial homotopy class represented by a simple closed curve. In a Moebius band there are two such classes but only one of these can be represented by a curve not bounding a Moebius band. □

In each maximal annulus or maximal Moebius band not containing a curve in G_r, choose a core curve. We may assume that f_Σ permutes these curves. Let G_c be the set of these curves.

LEMMA 2.6. *No curve in G_c bounds a disk or a Moebius band. No two curves in $G_c \cup G_r$ bound an annulus.*

Proof. Let c_1 be a curve in G_c. c_1 is contained in some A which is maximal in Φ. If c_1 bounds a manifold $X = D^2$ or $X =$ Moebius band, then $A \cup X \in \Phi$. By the maximality of A, $X \subset A$ so that c_1 bounds X in A. This is not possible.

If two curves in G_c, $c_1 \subset A_1$ and $c_2 \subset A_2$ bound an annulus A, then $A_1 \cup A \cup A_2$ is in Φ so $A_1 \subset A_2$ and $A_2 \subset A_1$. Thus $c_1 = c_2$. Now assume $c_1 \in G_c$ and $c_2 \in G_r \backslash G_c$, bound an annulus A. Say $c_1 \subset A_1$. If A_1 is a Moebius band then c_1 represents an element of $H_1(M, \mathbb{Z}_2)$ with non-trivial algebraic self-intersection. But the disjoint curve c_2 represents the same homology class. This is impossible. Thus A_1 is an annulus. $\partial A_1 \subset \Gamma_\Sigma$. Thus c_1 is homotopic to an element of Γ_Σ, hence c_2 is homotopic to an element of Γ_Σ. This contradicts lemma 2.3. □

The lemma implies that each element of G_c represents an element of $\mathcal{S}$.

Definition. Let x be an element of Σ contained in some maximal annulus or Moebius band in Φ. Then $\sigma(x)$ is the class in $\mathcal{S}$ represented by the core curve of the maximal element of Φ that contains x.

Assertion (1) of proposition 2.1 follows from lemma 1.2 and the fact that the curves in $G_c \cup G_r$ are disjoint.

We now prove assertion (2). f_Σ acts on the set of maximal annuli and Moebius bands. Clearly the function which assigns to each $x \in \Sigma$ the homotopy class of the maximal annulus or Moebius band that contains it is f_Σ-equivariant. The actions of f and f_Σ on Σ are the same by construction. The actions of f and f_Σ on $\mathcal{S}$ are the same because f_Σ and f are isotopic. Thus σ is f-equivariant.

We prove assertion (3). Let $Q \subset \Sigma_A$ consist of flipped points mapped by σ to a single f-orbit in $\mathcal{S}$. The orbit in $\mathcal{S}$ is represented by curves $c_1, \ldots, c_p$ in G_r. These curves cannot have oriented neighbourhoods whose orientations are preserved by f_Σ and they cannot be longitudinally flipped hence they either have Moebius band neighbourhoods or have normal orientations reversed by f_Σ. Both of these properties are homotopy invariant hence they hold for the corresponding curves under f. This proves the first statement. The points in Q lie on the curves $c_1, \ldots, c_p$ and these curves are cyclically permuted by f_Σ. It follows that all points in Q have the same period under f_Σ hence under f.

We prove assertion (4). If x is not a flipped periodic point it does not lie on a curve in G_r. If $\sigma(x)$ is represented by a normally flipped geodesic it must be the homotopy class of the core curve in a maximal annulus A containing x. If the core curve γ of A is homotopic to a normally flipped curve under f then γ must be normally flipped by f_Σ. This follows from the corollary to lemma 1.4. $A-\gamma$ has two components. If γ has order p these components are reversed by f_Σ^p. Since $x \in A-\gamma$, the order of x must be an even multiple of p.

In order to prove assertion (5) we need the following extension of Epstein's result on isotopies.

LEMMA 2.7. Let *G be a family of disjoint simple closed curves representing distinct elements of* $\mathcal{S}$. *Then there is a homeomorphism h of M isotopic to the identity so that* $h(G)$ *is a disjoint union of geodesics.*

Proof. Let $c_1 \cdots c_k$ be the curves in G. If there is an h_i which takes $c_1 \cdots c_i$ to geodesics, we will construct an h_{i+1} isotopic to h_i that takes $c_1 \cdots c_{i+1}$ to geodesics. Let

$$M_i = M - \bigcup_{j=1}^{i} h_i(c_j).$$

Let $\bar{M}_i$ be the metric closure of M_i. $\bar{M}_i$ is a compact manifold with geodesic boundary. Variational arguments show $h_i(c_{i+1})$ is homotopic to a geodesic γ_{i+1} inside $\bar{M}_i$. Since c_{i+1} is not homotopic to any $h(c_j)$, for $j = 1 \cdots i$, γ_{i+1} is not a boundary component of $\bar{M}_i$, so $\gamma_{i+1} \subset M_i$.

According to [**4**] there is a homeomorphism h' which takes $h_i(c_{i+1})$ to γ_{i+1} which is isotopic to the identity via a compactly supported isotopy. h' extends trivially to a homeomorphism $\bar{h}$ of M isotopic to the identity. Let $h_{i+1} = \bar{h} \circ h_i$. □

The remainder of this section is devoted to proving assertion (5).

Definition. For $X \subset M$, a two-dimensional submanifold, let X^0 denote the union of X and all disk components of $M - X$.

Definition. A submanifold $Y \subset M$ is said to be *substantial* if $Y = X^0$ where X is a component of $M - \Gamma_\Sigma$ not contained in an element of Φ.

If D is any maximal disk then $\partial D \subset \bar{X}$ where X is a component of $M - \Gamma_\Sigma$ and $D \subset X^0$. Thus every element of $\Sigma - \Sigma_A$ is contained in a substantial Y. The intersection of two distinct substantial pieces consists of boundary components.

Let G_s be the set of curves which are components of the intersection of two distinct substantial pieces. The curves in G_s do not bound disks or Moebius bands. Recall that G consists of core curves of maximal annuli and Moebius bands. No two curves in the set $G_s \cup G_c$ are homotopic. All these assertions are easily proved.

The components of $M - \{\bigcup G_s \cup \bigcup \mathcal{A}\}$ correspond to substantial pieces. f_Σ acts on this set of components. We wish to describe this action in terms of $\check{f}$.

Let Γ_c and Γ_s be sets of geodesics representing curves in the sets G_c and G_s. By the previous lemma there is a homeomorphism $h : M \to M$ isotopic to the identity which takes the curves in $G_c \cup G_s$ to curves in $\Gamma_c \cup \Gamma_s$. The action of f_Σ on $G_c \cup G_s$ must correspond under h to the action of $\check{f}$ on $\Gamma_c \cup \Gamma_s$: both are determined by the action on the corresponding classes in $\mathcal{S}$. Furthermore, the action of f_Σ on $M - \{\bigcup G_c \cup \bigcup G_s\}$ must correspond under h to the action of f on $M - \{\bigcup \Gamma_c \cup \bigcup \Gamma_s\}$. Let B be a boundary component of C. Let $g = f_\Sigma^{-1} h^{-1} \check{f} h$. By the previous argument, $g(B) = B$. We must show that $g(C) = C$. If not then g reverses the normal orientation of B, but according to the corollary of lemma 1.4 g, being homotopic to the identity, must preserve the normal orientation of B.

Now let C be a substantial piece with f_Σ setwise period p. Then $h(C)$ has $\check{f}$ setwise period p. $h \circ f_\Sigma^p$ and $\check{f}^p \circ h$ both take C to $h(C)$. These maps are homotopic as maps into M. C is incompressible because its complement contains no disk components. According to a lemma of [**7**], $h \circ f_\Sigma^p$ and $\check{f}^p \circ h$ are homotopic as maps into $h(C)$.

We modify $f_\Sigma^p|C$ so that it has finite order. C is equal to X^0 for some X and f_Σ^p has finite order on X. We can extend $f_\Sigma^p|X$ to the maximal disks by extending radially the isometry induced by f_Σ^p on the boundaries of the maximal disks. This modification of f_Σ^p changes the periods of points in Σ contained in maximal disks. If x is such a point contained in D and y is the centre of D then the f_Σ^p period of x is a multiple of the f_Σ^p setwise period of D which is equal to the modified f_Σ^p period of y. Thus is suffices to prove that the periods of points under the modified map are multiples of the periods of points under $\check{f}^p$.

The following lemma completes the proof of the main proposition.

LEMMA 2.8. *Let Y be a 2-manifold with negative Euler characteristic. Let $g_i : Y \to Y$ $(i = 1, 2)$ be homotopic homeomorphisms of finite order. Then the order of g_1 is equal to the order of g_2 and the periods of unflipped periodic points in Y are the same under g_1 and g_2.*

Proof. Let g be either g_1 or g_2. Let χ_m be the Euler characteristic of the set of points of period m under g. The set of unflipped points of period m is discrete if m is less than the order of g. The set of flipped periodic points has Euler characteristic zero. So χ_m is the number of unflipped points of period m. If m is the order of g then the set of points of period m is the complement of the set of periodic points of lower period, so χ_m is negative. The lemma will be proved if we show that the χ_m are invariants of the homotopy class of g. According to [**3**],

$$\eta_g(t) = \prod_{m=1}^{\infty} (1 - t^m)^{-\chi_m/m},$$

where η_g is the homology zeta function. By considering the orders of the poles of η_g at primitive roots of unity we see that η_g determines the χ_m. η_g is a homology invariant hence *a fortiori* a homotopy invariant. □

A more detailed analysis shows that homotopic maps of finite order are conjugate.

3. In this section we use the previous propositions to construct examples of isotopy classes in which the dynamic behaviour of zero entropy maps is restricted.

(1) We construct an orientation preserving homeomorphism f of the surface of genus 3 with the following property: if f' is a zero entropy homeomorphism isotopic to f then all periods of periodic points of f' are even. f is constructed by first rotating the pictured surface 180° around the axis A and then putting in a non-trivial twist in one of the pictured annuli. The period of each f-periodic point

is two and any geodesic contained in the isometry region has period 2. The asserted properties of zero entropy maps isotopic to f follow from the main propositon.

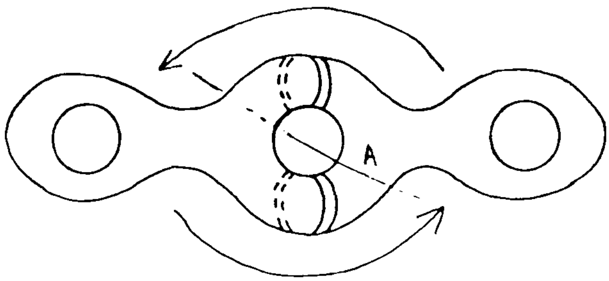

FIGURE 1

We can construct a similar example on any oriented surface of greater genus:

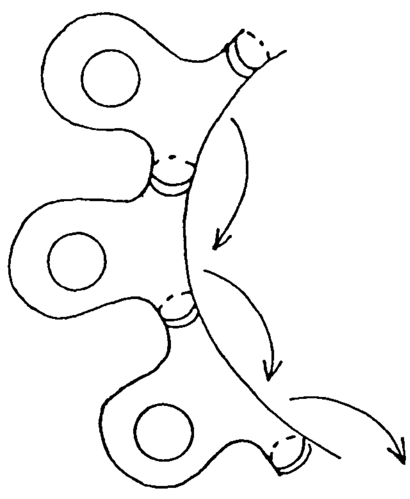

FIGURE 2

If g is the genus then the isometry region of M will have $g-1$ components which are cyclically permuted by f. If f' has zero entropy and f' is isotopic to f then all periods of periodic points of f' are divisible by $g-1$.

(2) We can construct a similar example on a non-oriented surface of genus greater than or equal to four by replacing the handles in figure 2 by cross caps.

(3) We construct a similar example which reverses orientation on an oriented surface. Let f be the composition of the following maps: a translation to the right, a reflection with the broken lines as fixed point set and a twist around one of the pictured annuli in each orbit. If n is the number of components of M-({twist curves} $\cup$ {reflective curves}) then n divides the period of any f'-periodic point. When n is even we get an example of an isotopy class in which no odd periods

can occur (cf. [**6**]). When n is odd there is at most one odd period which will be of the form np.

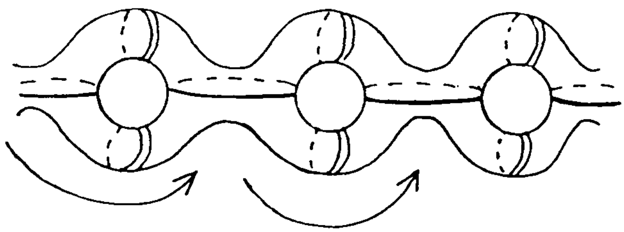

FIGURE 3. g = number of components + 1

These three families of examples prove:

PROPOSITION 3.1. *Let M be a surface, not necessarily oriented, with Euler characteristic less than -4. There exists a zero entropy homeomorphism $f: M \to M$ with the property that every f' homotopic to f with a fixed point has positive entropy. If M is oriented we may choose f to be either orientation preserving or orientation reversing.*

(4) We consider maps isotopic to the identity on non-orientable surfaces. There is an assignment of locally flipped periods to geodesics. There are no normally flipped geodesics. Flipped periods must be assigned to geodesics with Moebius band neighbourhoods. The maximal number of disjoint Moebius bands in a surface is its (non-oriented) genus. Thus the number of distinct periods of locally flipped periodic points is less than or equal to the genus.

(5) We construct an $f: N \to N$ where N is non-orientable and no zero entropy map isotopic to f can have any flipped periodic points. Let $N = K^2 \# M$, where M is oriented and $M \neq S^2$. Let γ be a two-sided curve in K^2. Let f be the identity outside a neighbourhood of γ and a non-trivial twist inside this neighbourhood. Since $N - \gamma$ is oriented and f preserves this orientation, there can be no Moebius bands or flipped geodesics of finite order.

The main proposition relates the dynamics of zero entropy homeomorphisms to the dynamics of maps in Thurston canonical form. If f is an isometry it is in Thurston canonical form.

In this section we discuss orientation-preserving finite order isometries f and the f-periods of points and geodesics. Let M be the oriented manifold on which f acts. Let n be the order of f. Let $B \subset M$ be the set of points with periods less than n. Let M_f be quotient of M by the action of f. Let $P: M \to M_f$ be the projection. The space M_f is a manifold and P is a branched covering map. The branching locus is B_f, the image of B under P. Let

$$P': M - B \to M_f - B_f$$

be the restriction of P. P' is an (unbranched) n-fold cyclic covering. Such a covering gives a surjective homomorphism from $\pi_1(M_f - B_f)$ to $\mathbb{Z}_n$ or equivalently a surjective homomorphism ϕ_f from $H_1(M_f - B_f, \mathbb{Z})$ to $\mathbb{Z}_n$.

Conversely, given a manifold N, a finite set $C \subset N$ and homomorphism ρ from $H_1(N-C)$ onto $\mathbb{Z}_n$, one can construct an isometry f on M of order n so that

$$M_f = N, \qquad B_f = C \quad \text{and} \quad \phi_f = \rho.$$

Let $\gamma \subset M - B$ be a simple closed curve disjoint from its images under f. $P(\gamma)$ will be a simple closed curve in $M_f - B_f$. Since we are concerned with the set of periods of geodesics and points we can ignore geodesics passing through points in B. The setwise period of such a geodesic is the period of the point in B that it contains. Choose an orientation on $P(\gamma)$ and let c be the corresponding element of $H_1(M_f - B_f)$. The setwise period of γ is then

$$n/(\text{order } \phi_f(c)).$$

Let $\mathscr{o}$ be an f-orbit. $P(\mathscr{o})$ is a point $x \in M_f$. Let e be the element of $H_1(M_f - B_f)$ represented by a small curve circling x once counterclockwise. Then the period of $\mathscr{o}$ is

$$n/(\text{order } \phi_f(e)).$$

According to [**8**], if the genus of M_f is not zero there exists an f-invariant simple closed curve. When such a curve exists there are no restrictions on possible subsets of the set of periods. We therefore restrict our attention to the case $M_f \approx S^2$.

$H(S^2 - B_f)$ is generated by elements e_i corresponding to points x_i in B_f with the sole relation $\Sigma e_i = 0$.

The genus of M is given by the Riemann Hurwitz formula

$$g(M) = 1 - n + \frac{n}{2} \sum_{i=1}^{k} \left(1 - \frac{1}{O_i}\right)$$

where O_i is the order of $\psi(e_i)$ in $\mathbb{Z}_n$.

We will give a combinatorial description of the possible periods of disjoint sets of simple closed curves and points under f. By the main proposition these sets of periods determine possible periods of periodic points of zero entropy homeomorphisms. We require the following definitions: a *tree* is a contractible graph; the *terminal* vertices are those vertices meeting only one edge. Let $\mathscr{C}$ be a disjoint set of simple closed curves not homotopically trivial in $M_f - B_f$. We may assume that $\mathscr{C}$ contains one small curve around each $x_i \in B_f$ since such curves can be added in the complement of the curves in $\mathscr{C}$. The setwise period of such a curve around x_i is the period of x_i.

For each $c \in \mathscr{C}$ choose a disjoint regular neighbourhood U_c and a product structure. Let

$$\lambda_c : U_c \to (-1, 1)$$

be the projection. Let G be the space obtained from M_f by collapsing each component of $M_f - \bigcup U_c$ to a point and collapsing each circle $\lambda_c^{-1}(r)$ to a point. G is a graph. Each point on an edge of G corresponds to a circle in M_f. Removing a circle from M_f $(= S^2)$ disconnects M_f. Thus removing a point from an edge of G disconnects G. G is therefore a tree. G has the additional property that each vertex meets at least three edges.

A terminal vertex in G corresponds to a disk component of $M_f - \bigcup U_c$ with a single boundary curve b. $b \subset U_c$ for some c. Since b is not trivial in $M_f - B_f$ this disk must contain an element x_i of B_f and c must be the small curve around x_i. In this way we label each terminal vertex with a unique element of B_f. Let $c \in \mathscr{C}$. Removing c divides G into two components. Let $a \in H_1(M_f - B_f)$ be the class which is the sum of classes associated to terminal vertices in one of these components. $[c]$ is homologous to $\pm a$. If c is the image of c' in M we can therefore compute the setwise period of c' from the labelled graph G. It is $n/(\text{order } a)$. We refer to this number as the period of the edge containing c'.

If we are given a labelled graph G we can construct a family of curves corresponding to the edges of the graph. We leave this construction to the reader.

Combining our main proposition with the preceding analysis proves the following:

THEOREM 3.1. *Let f be an isometry of order n so that M_f is a sphere. Let $\varepsilon_1, \ldots, \varepsilon_k$ be the associated elements of $\mathbb{Z}_n$. If f' has zero entropy and f' is isotopic to f then there is a tree T with k vertices labelled by $\varepsilon_1, \ldots, \varepsilon_k$ so that the f'-period of every f'-periodic point is a multiple of the period of some edge of T. Conversely, given any tree T with terminal vertices labelled $\varepsilon_1, \ldots, \varepsilon_k$ we can construct a zero entropy diffeomorphism f' isotopic to f so that every multiple of a period of an edge is the period of some f'-periodic point.*

We now give the proofs of the propositions stated in the introduction.

Proof of Proposition 3.2. We will construct an isometry f of order 210 having the required properties. We write $\mathbb{Z}_{210}$ as

$$\mathbb{Z}_2 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_5 \oplus \mathbb{Z}_7$$

and we specify elements of $\mathbb{Z}_{210}$ by giving their coordinates. B will consist of four elements. Let ε_i, $i = 1, 2, 3, 4$, be the corresponding elements of $\mathbb{Z}_{210}$.

TABLE 1

	$\mathbb{Z}_2$	$\mathbb{Z}_3$	$\mathbb{Z}_5$	$\mathbb{Z}_7$	Period $= n/\text{order } \varepsilon_i$
$\varepsilon_1 =$	(1,	0,	0,	0)	105
$\varepsilon_2 =$	(1,	0,	1,	1)	3
$\varepsilon_3 =$	(0,	1,	−1,	0)	14
$\varepsilon_4 =$	(0,	−1,	0,	−1)	10

Possible trees with four terminal vertices can have at most one edge not connected to a terminal vertex. The possible elements of $\mathbb{Z}_{210}$ that can be associated to this edge are (up to sign):

	Period
$\varepsilon_1 + \varepsilon_2 = (0, 0, 1, 1)$	6
$\varepsilon_1 + \varepsilon_3 = (1, 1, -1, 0)$	7
$\varepsilon_1 + \varepsilon_4 = (1, 2, 0, 6)$	5

The periods that can be obtained with a tree with an $\varepsilon_1 + \varepsilon_3$ edge are $\{3, 7, 10\}$ and their multiples. The periods that can be obtained with a tree with an $\varepsilon_1 + \varepsilon_4$ edge are $\{3, 5, 14\}$ and their multiples. The periods obtained with an $\varepsilon_1+\varepsilon_2$ edge are multiples of periods of points in B. Thus an $\varepsilon_1+\varepsilon_2$ edge contributes no new periods.

From the Riemann Hurwitz formula we calculate that the genus of M is 145. □

Proof of Proposition 3.3. We begin by showing that if $f: M\to M$ is a map in Thurston canonical form which is a counterexample to proposition 3.3 then there is an $f': M'\to M'$ which is also a counterexample where genus$(M')\le$ genus(M), f' is an isometry, $M_{f'}=S^2$ and $B_{f'}$ consists of four points.

We will use the following *ad hoc* terminology. A natural number p is *realized* (by f) if p is a multiple of the setwise f-period of some simple curve γ in M whose orbit consists of disjoint curves. A set of natural numbers p_i is *simultaneously realizable* if there is a set of curves γ_i where γ_i realizes p_i and the curves γ_i have disjoint orbits.

Assume that not all periods that can be realized by f can be realized simultaneously. We can choose a set of periods $q_1, \ldots, q_l$ so that each q_i can be realized, no q_i is a multiple of any other and every realizable p is a multiple of some q_i. There is a least number j such that $\{q_1, \ldots, q_j\}$ is not simultaneously realizable. We can choose curves $\gamma_1, \ldots, \gamma_j$ which realize $q_1, \ldots, q_j$ so that the orbits of $\gamma_1, \ldots, \gamma_{j-1}$ are disjoint. There exists a γ_k, $1\le k\le j-1$, that intersects γ_j so that no curve representing q_k disjoint from $\gamma_1, \ldots, \gamma_{j-1}$ is disjoint from γ_j. γ_j and γ_k are in the same component, N_0, of M-{twist curves}. Let l be the setwise period of N_0. By passing to the l'th iterate of f and dividing the q_i's by l we may assume that $f(N_0)=N_0$.

According to the previously quoted result of Meeks, the existence of some period which is not realizable implies that the genus of M_f is 0. Let $\bar\gamma_j$ and $\bar\gamma_k$ be the images of γ_j and γ_k in M_f. Let $B_f\subset M_f$ be the set of branch points. Let M_f' be M_f with a small open disk around each point in B_f removed. $\bar\gamma_j$ and $\bar\gamma_k$ each partition the boundary components of M_f' into two sets. The partition determines the homology class. We can find curves $\bar\gamma_j'$ and $\bar\gamma_k'$ homologous to $\bar\gamma_j$ and $\bar\gamma_k$ so that $M_f'-\bar\gamma_j'\cup\bar\gamma_k'$ has four components. Let b_r, $r=1\cdots 4$, be boundary components of a regular neighbourhood N_f of $\bar\gamma_j'\cup\bar\gamma_k'$. The periods represented by the curves b_r cannot divide the periods q_j and q_k: the period represented by b_i is a multiple of q_s for some $i\le s\le l$. If q_s divides q_j then $q_s=q_j=$ period of b_i. But b_i is disjoint from $\bar\gamma_k'$, contradicting the assumption that q_j and q_k are not simultaneously realizable.

Let N be a component of the inverse image of N_f in M^0. By passing to the appropriate power of f we may assume that it is connected. Add a disk to each boundary component of N. Replacing a submanifold of M with a disk cannot result in a manifold with lower Euler characteristic. f extends to these disks. The setwise periods of these disks are the periods represented by the curves b_r. With the resulting map and surface the periods of $\bar\gamma_j'$ and $\bar\gamma_k$ can be realized but not simultaneously.

Now assume that the counterexample f is an isometry, $M_f\approx S^2$, B_f consists of four points and genus $(M)\le 145$. Let n be the order of f and let

$$\varepsilon_i=\phi_f(e_i),\qquad 1\le i\le 4.$$

https://doi.org/10.1017/S014338570000198X Published online by Cambridge University Press

We write

$$n = p_1^{a_1} p_2^{a_2} \cdots p_l^{a_l}$$

where the exponents are not zero. Let ε_i^j be the $\mathbb{Z}_{p_j^{a_j}}$ coordinate of ε_i. ϕ is determined by the $4 \times l$ matrix $E = \varepsilon_i^j$.

According to [8] if l were less than 3 there would be a curve in M invariant under f and all periods could be realized. Now consider the case $l \geq 4$. We need some elementary observations about the matrix E which hold for any l.

Each column contains at least two generators: consider the column corresponding to $\mathbb{Z}_{p_1^{a_1}}$. The elements ε_i^1 generate $\mathbb{Z}_{p_1^{a_1}}$ so at least one of them is a generator. If exactly one of these elements were a generator then the sum of the ε_i's would be a generator, since the sum of a generator and non-generators is a generator. Their sum is zero however, so at least two of them are generators.

No row consists entirely of generators: if every coordinate of ε_1 were a generator then ε_1 would generate $\mathbb{Z}_n$. The point corresponding to ε_1 would be fixed under f and every period could be realized by a map homotopic to f. Similarly no $\varepsilon_i + \varepsilon_j$ can be a generator.

We cannot have two elements of order $p_i^{a_i}$. Say ε_1 and ε_2 have generators only in the first coordinate; then ε_3 and ε_4 must have generators in all coordinates but the first. But this would mean that $\varepsilon_1 + \varepsilon_3$ generates $\mathbb{Z}_n$.

Now if $l \geq 4$ some $p_i \geq 7$ so there must be two elements of order at least 7. The remaining elements cannot both have order 2 so one must have order greater than or equal to 3. If we arrange the orders of the ε_i in order of increasing magnitude we have $O_1 \geq 2$, $O_2 \geq 3$, $O_3 \geq 7$ and $O_4 \geq 7$. Applying the Riemann Hurwitz formula:

$$g(M) = 1 - n + \frac{n}{2} \sum_{i=1}^{4} 1 - \frac{1}{O_i},$$

$$g(M) \geq 1 - n + \frac{n}{2}\left(1 - \tfrac{1}{2} + 1 - \tfrac{1}{3} + 1 - \tfrac{1}{7} + 1 - \tfrac{1}{7}\right),$$

$$g(M) \geq 1 + \tfrac{37}{84} n.$$

If we combine this with the restriction $g(M) \leq 145$ we get $n < 328$. The only n which is a product of at least four primes and is less than 328 is 210. This is the value of n in our example. We leave to the reader the verification that when $n = 210$ the lowest value for the genus of M is 145.

Assume that exactly three distinct primes divide n.

Case 1. Assume that some ε_i, say ε_4, does not contain a generator in any coordinate. Since for each of the 3 coordinates there are two indices i such that ε_i has a generator in that coordinate there are at least six generators. No ε_i has generators in 3 coordinates so each ε_i, $i = 1, 2, 3$, must contain generators in two coordinates. We may assume that the situation is described by table 2, where a '$*$' in the i, j'th place implies that ε_i contains a generator in the i'th coordinate.

If ε_4 has a zero in the i'th place then the order of $\varepsilon_4 + \varepsilon_i$ is the same as the order of ε_i. For at least two values of i the order of $\varepsilon_i + \varepsilon_4$ must be distinct in order to

https://doi.org/10.1017/S014338570000198X Published online by Cambridge University Press

TABLE 2

	$p_1^{a_1}$	$p_2^{a_2}$	$p_3^{a_3}$
ε_1		*	*
ε_2	*		*
ε_3	*	*	
ε_4			

get non-simultaneously realizable periods. Thus the order of ε_4 is divisible by at least two primes and at least two exponents are greater than one. Assuming that $p_1^{a_1}$, $p_2^{a_2}$, $p_3^{a_3}$ are in increasing order we have $p_1^{a_1} \geq 2$, $p_2^{a_2} \geq 3$, $p_3^{a_3} \geq 5$ and hence $O_1 \geq 15$, $O_2 \geq 10$, $O_3 \geq 6$ and $O_4 \geq 6$. Arguing as before gives $n \leq 192$. There is only one solution $p_1^{a_1} p_2^{a_2} p_3^{a_3} \leq 192$ with at least two exponents greater than one; this is $n = 2^2 \cdot 3^2 \cdot 5 = 180$. With this n we get $O_1 \geq 36$, $O_2 \geq 20$, $O_3 \geq 45$, $O_4 \geq 6$. Plugging into our formula gives $n \leq 166$, a contradiction.

Case 2. Assume that each ε_i contains a generator in at least one coordinate and some ε_i, say ε_4, contains a generator in exactly one coordinate, say p_1. There are two indices $i = 1, 2, 3$ for which ε_i contains a generator in the second coordinate and two indices $i = 1, 2, 3$ for which ε_i contains a generator in the third coordinate. Some ε_{i_0} cannot contain a generator in the first coordinate. $\varepsilon_{i_0} + \varepsilon_4$ contains a generator in every coordinate, which is impossible.

Case 3. Assume that each ε_i contains a generator in two coordinates. There are two subcases. First assume that some column has exactly 3 generators in it. The sum of any two of these generators must be a generator otherwise a generator plus two non-generators would be zero. We have the situation displayed in table 3.

TABLE 3

	$p_1^{a_1}$	$p_2^{a_2}$	$p_3^{a_3}$
ε_1	*	*	•
ε_2	*	•	*
ε_3	*	?	?
ε_4	•	?	?

$\varepsilon_1 + \varepsilon_2$ has a generator in every coordinate.

If no column has exactly 3 generators, we have the situation displayed in table 4. The periods of curves corresponding to $\varepsilon_1 + \varepsilon_3$ and $\varepsilon_1 + \varepsilon_4$ are powers of p. Thus the period of one is a multiple of the period of the other. In order to get periods

TABLE 4

	$p_1^{a_1}$	$p_2^{a_2}$	$p_3^{a_3}$
ε_1	*	*	
ε_2	*	*	
ε_3	*		*
ε_4	*		*

which cannot be realized simultaneously, the curves corresponding to $\varepsilon_1+\varepsilon_2$ must realize new periods. The period of this curve is a multiple of $p_2 \cdot p_3$. The period of the point corresponding to ε_1 is a multiple of p_3. Since this does not divide the period corresponding to $\varepsilon_1+\varepsilon_2$, a_3 must be at least 2. Making the same argument with ε_3 we see that a_2 must be at least 2. Arranging $p_i^{a_i}$ in order of increasing size, $p_1^{a_1} \geq 2$, $p_2^{a_2} \geq 4$, $p_3^{a_3} \geq 9$. Thus $O_1 \geq 8$, $O_2 \geq 8$, $O_3 \geq 18$, and $O_4 \geq 18$. Applying our inequality gives $n \leq 175$, which is impossible, since when at least two exponents are greater than 2, $n \geq 180$. □

I would like to thank Steve Batterson and Michael Handel for helpful conversations. I would like to thank Steve Batterson for his constant encouragement. The calculations necessary for propositions 3.2 and 3.3 were first done by him. I would like to thank Michael Handel for sharing with me his intriguing work on surfaces.

REFERENCES

[1] P. Blanchard & J. Franks. The dynamical complexity of orientation reversing homeomorphisms of surfaces. *Inventiones Math.* **62** (1980), 333–339.
[2] R. Bowen & J. Franks. The periodic points of maps of the disk and interval. *Topology.* **15** (1976), 337–342.
[3] S. Batterson & J. Smillie. Filtrations and Periodic Data on Surfaces. *Preprint.*
[4] D. B. A. Epstein. Curves on 2-manifolds and isotopies. *Acta Math.* **69** (1956), 135–206.
[5] A. Fathi, F. Laudenbach & N. Poenaru. Travaux de Thurston sur les surfaces. *Astérisque* **66–67** (1979).
[6] M. Handel. The entropy of orientation-reversing homeomorphisms of surfaces. *Topology.* (To appear.)
[7] W. Jaco & P. Shalen. Surface homeomorphisms and periodicity. *Topology* **16** (1977), 347–367.
[8] W. Meeks. Circles invariant under diffeomorphisms of finite order. *J. Differential Geometry* **14** (1979), 377–383.
[9] C. Narasimhan. The periodic behavior of Morse-Smale diffeomorphisms on compact surfaces. *Trans. Amer. Math. Soc.* **248** (1979), 145–149.
[10] W. Thurston. On the geometry and dynamics of diffeomorphisms of surfaces. *Preprint.*