ISSN 1364-0380 (on line) 1465-3060 (printed) **713**

Geometry & Topology
Volume 7 (2003) 713–756
Published: 30 October 2003

Periodic points of Hamiltonian surface diffeomorphisms

John Franks
Michael Handel

Department of Mathematics, Northwestern University
Evanston, IL 60208-2730, USA
and
Department of Mathematics, CUNY, Lehman College
Bronx, NY 10468, USA

Email: `john@math.northwestern.edu`, `handel@g230.lehman.cuny.edu`

Abstract

The main result of this paper is that every non-trivial Hamiltonian diffeomorphism of a closed oriented surface of genus at least one has periodic points of arbitrarily high period. The same result is true for S^2 provided the diffeomorphism has at least three fixed points. In addition we show that up to isotopy relative to its fixed point set, every orientation preserving diffeomorphism $F\colon\ S \to S$ of a closed orientable surface has a normal form. If the fixed point set is finite this is just the Thurston normal form.

AMS Classification numbers Primary: 37J10

Secondary: 37E30

Keywords: Hamiltonian diffeomorphism, periodic points, geodesic lamination

Proposed: Benson Farb Received: 28 March 2003
Seconded: Leonid Polterovich, Joan Birman Revised: 26 October 2003

1 Introduction

The main result of this paper is that every non-trivial Hamiltonian diffeomorphism of a closed oriented surface of genus at least one has periodic points of arbitrarily high period. The same result is true for S^2 provided the diffeomorphism has at least three fixed points. It was previously known (see [12], [13]) that any non-trivial area preserving homeomorphism of S^2 with at least three fixed points has infinitely many periodic points. The existence of unbounded periods is a substantially stronger conclusion that is not only interesting in its own right but has applications [16] to the algebraic properties of the group of area preserving diffeomorphisms of closed oriented surfaces.

Theorem 1.1 *Suppose $F\colon S \to S$ is a non-trivial, Hamiltonian diffeomorphism of a closed oriented surface S and that if $S = S^2$ then $\mathrm{Fix}(F)$ contains at least three points. Then F has periodic points of arbitrarily high period.*

Our second result is closely related to, and depends on, a celebrated result of Thurston [29] which provides a canonical form, up to isotopy relative to a finite invariant set, for any homeomorphism of a compact surface. With the added hypothesis that the homeomorphism is a diffeomorphism, we obtain an analogous result for isotopy relative to the fixed point set of the map, even if that fixed point set is infinite.

Theorem 1.2 *Every orientation preserving diffeomorphism $F\colon S \to S$ of a closed orientable surface has a normal form up to isotopy relative to its fixed point set.*

More details, including the definition of *normal form* is provided below in §6.

The main technical work in the paper (sections 8 through 11) is an analysis of the dynamics of an element $f \in \mathcal{P}(S)$, the set of diffeomorphism of the surface S that are isotopic to the identity relative to the set of periodic points of f. (The standard example of an element of $\mathcal{P}(S)$ is the time one map of a flow whose periodic points are all fixed.) Preservation of area plays no role in this analysis. But rather striking conclusions follow when the area preserving hypothesis is added. Indeed we ultimately prove, for example, that the identity is the only area preserving element of $\mathcal{P}(S^2)$ that fixes at least three points. One can construct non-trivial area preserving examples in $\mathcal{P}(T^2)$ by stopping an irrational flow at a closed set of points. These examples have non-trivial mean rotation vector and so are not Hamiltonian. (See Corollary 2.12.) All elements of $\mathcal{P}(S)$ have zero entropy [23]. While the converse is not true, $\mathcal{P}(S)$ is a rich family of zero entropy diffeomorphisms and the techniques that we develop here should be useful in understanding the extent to which arbitrary zero entropy diffeomorphisms behave like the time one maps of flows. See also [19].

As a consequence of Theorem 1.1 and the techniques used to prove it we will establish the following result which is of some independent interest.

Theorem 1.3 *Suppose $F\colon S \to S$ is a non-trivial, Hamiltonian diffeomorphism of a closed oriented surface S and that if $S = S^2$ then $\mathrm{Fix}(F)$ contains at least three points. Then there exist $n > 0$ and a finite set $P \subset \mathrm{Fix}(F^n)$ such that F^n is not isotopic to the identity relative to P.*

We will show that a consequence of this is the following slight strengthening of Theorem 1.1.

Theorem 1.4 *Suppose $F\colon S \to S$ is a non-trivial, Hamiltonian diffeomorphism of a closed oriented surface S and that if $S = S^2$ then $\mathrm{Fix}(F)$ contains at least three points. Then there exist $n > 0$ and $p > 0$ such that F^n has a point of period k for every $k \geq p$.*

2 The flux homomorphism and rotation vectors

Suppose S is a closed oriented surface and ω is a smooth volume form. We will generally assume a fixed choice of ω and refer to a diffeomorphism $F\colon S \to S$ which preserves ω as an *area preserving diffeomorphism.* We denote the group of diffeomorphisms preserving ω by $\mathrm{Diff}_\omega(S)$ and its identity component by $\mathrm{Diff}_\omega(S)_0$. Similarly, the identity component of $\mathrm{Diff}(S)$, is denoted $\mathrm{Diff}(S)_0$.

Remark 2.1 The inclusion of $\mathrm{Diff}_\omega(S)_0$ into $\mathrm{Diff}(S)_0$ is a homotopy equivalence for all S (see Corollary 1.5.4 of [2]). It follows that an ω preserving diffeomorphism F_1 that is isotopic to the identity is isotopic to the identity through ω preserving diffeomorphisms. Moreover, any isotopy F_t of F_1 to the identity is homotopic relative to its endpoints to an isotopy of F_1 to the identity through ω preserving diffeomorphisms.

Definition 2.2 Consider the set K of paths F_t in $\mathrm{Diff}_\omega(S)_0$, with $t \in [0,1]$, such that $F_0 = id$. We let $\mathcal{G}_\omega(S)$ denote the set of equivalence classes in K

where two paths are equivalent if they are homotopic relative to their endpoints. There is a group multiplication on $\mathcal{G}_\omega(S)$ defined by composition, i.e., $[G_t][F_t] = [H_t]$ where H_t is defined by $H_t(x) = G_t(F_t(x))$ for $x \in S$.

Remark 2.3 If S has genus greater than 1 then $\mathrm{Diff}(S)_0$ is simply connected (see Theorem 1 of [6] or [18]) so $[F_t]$ depends only on F_1. This provides a canonical identification of $\mathcal{G}_\omega(S)$ with $\mathrm{Diff}_\omega(S)_0$ by identifying $[F_t]$ with F_1.

Remark 2.4 Each $[F_t] \in \mathcal{G}_\omega(S)$ determines a preferred lift $\tilde{F}_1 \colon \tilde{S} \to \tilde{S}$ of F_1 to the universal cover of S. This is the lift of F_1 obtained by lifting the isotopy F_t starting at the identity of $\tilde{S}$. It is clearly independent of the choice of representative in the homotopy class $[F_t]$. We are primarily concerned here with the case that the genus is 1. In this case (see Theorem 1 of [6] or [17]) the lifting provides a canonical identification of $\mathcal{G}_\omega(T^2)$ with the group of diffeomorphisms of $\mathbb{R}^2$ which are lifts of area preserving diffeomorphisms of T^2 isotopic to the identity. In general this is the viewpoint we will most often take on $\mathcal{G}_\omega(T^2)$.

We want now to define the *flux homomorphism* from $\mathcal{G}_\omega(S)$ to $H^1(S,R)$. We will largely follow Chapter 3 of [2], where it is shown that for any smooth $\sigma \colon S^1 \to S$ and any path G_t in $\mathrm{Diff}_\omega(S)_0$ starting at the identity, the map $G_\sigma \colon S^1 \times I \to S$ defined by $G_\sigma(x,t) = G_t(\sigma(x))$ has the property that the value of

$$\int_{S^1 \times I} G_\sigma^*(\omega)$$

depends only on the equivalence class $[G_t]$ of G_t in $\mathcal{G}_\omega(S)$ and the homology class of the singular cycle $[\sigma]$. Moreover, the function assigning the value to the homology class $[\sigma]$ is a homomorphism. These properties are summarized in the following result:

Theorem 2.5 (See [2]) *There is a unique homomorphism*

$$\mathcal{F} \colon \mathcal{G}_\omega(S) \to H^1(S,R)$$

with the property that

$$\mathcal{F}([G_t])([\sigma]) = \int_{S^1 \times I} G_\sigma^*(\omega)$$

for any $[G_t] \in \mathcal{G}_\omega(S)$ and any smooth singular cycle $\sigma \colon S^1 \to S$. The homomorphism $\mathcal{F}$ is called the flux homomorphism.

Remark 2.6 The flux is well defined when G_t is an isotopy from the identity to $G_1 \in \mathrm{Diff}_\omega S$ even if G_t is not area preserving for $t \in (0,1)$. Indeed, Remark 2.1 implies that the path G_t in $\mathrm{Diff}(S)_0$ is homotopic relative to its endpoints to a path $H_t \in \mathrm{Diff}_\omega S \subset \mathrm{Diff}(S)_0$ and Stokes' theorem implies that the integrals defining $\mathcal{F}(H_t)$ and $\mathcal{F}(G_t)$ have the same value.

Remark 2.7 The value of the integral defining $\mathcal{F}([G_t])([\sigma])$ is the same as the value of a similar integral in a covering space of S and we will want to make use of this. More precisely, suppose $p\colon \bar{S} \to S$ is a covering space, $\sigma\colon S^1 \to S$ is a singular cycle and that $\bar{\sigma}$ is a lift of σ to $\bar{S}$. Let $\bar{\omega} = p^*(\omega)$, and let $\bar{G}_t$ be the lift of G_t with $\bar{G}_0 = id$. We observe that

$$\mathcal{F}([G_t])([\sigma]) = \int_{S^1 \times I} \bar{G}^*_{\bar{\sigma}}(\bar{\omega}).$$

This is because $p \circ \bar{G}_{\bar{\sigma}} = G_\sigma$ and hence $G^*_\sigma(\omega) = \bar{G}^*_{\bar{\sigma}}(p^*(\omega)) = \bar{G}^*_{\bar{\sigma}}(\bar{\omega})$.

Closely related to the flux homomorphism is the *mean rotation vector* associated to elements of $\mathcal{G}_\omega(S)$ which we now define following [14]. We want to define a "homological displacement" of a point $x \in S$ by $[G_t]$. We will do this by considering the path $G_t(x)$ from x to $G_1(x)$ and closing it to form a loop whose homology class will define this displacement. There are choices of how to close it. We consider S as a convex polygon with edges identified. We choose a subset X of this polygon consisting of its interior and a subset of its boundary chosen so the identification map to S is a bijection. Then each path in S can be closed by taking the points of X corresponding to its endpoints and joining them to a base point in the interior of the polygon (see [14] for more details).

It is easy to see that if the subset of the boundary of the polygon which is in X is a Borel measurable set then the function $\delta\colon S \to H_1(S,R)$, which assigns to x the homology class of the loop formed by closing the path $G_t(x)$ in this way, is a bounded Borel measurable function. We will call δ a *homological displacement function* associated to $[G_t]$.

Concatenating the paths $G_t(G_1^i(x))$ for $0 \le i < n$ gives a path from x to $G_1^n(x)$ and the homology class of the loop formed by closing this path is $\sum_{i=0}^{n-1} \delta(G_1^i(x))$, (See [14] for more details).

Definition 2.8 Suppose that $\delta\colon S \to H_1(S,R)$ is a homological displacement function associated to an isotopy $[G_t]$. We define the *rotation vector* of a point $x \in S$ for the isotopy G_t to be

$$\rho(x, G_t) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \delta(G_1^i(x)) \in H_1(S,R)$$

if this limit exists.

If we are discussing the case when S is an annulus A we will often identify G_t with a lift $\tilde{G}$ of G_1 to the universal cover of A and identify $H_1(A, \mathbb{R})$ with $\mathbb{R}$. We will then refer to the resulting value of $\rho(x, \tilde{G})$ as the *rotation number* of x.

Even for a general S, it follows from the ergodic theorem that the limit $\rho(x, G_t)$ exists for almost all x with respect to any invariant Borel measure and defines an integrable function. It is also clear that $\rho(x, G_t)$ is independent of the choice of the displacement function δ used to define it. This is because the two homology classes formed by closing the path from x to $G_1^n(x)$ in two different ways would differ by an amount bounded independently of n and hence this difference disappears in the limit defining $\rho(x, G_t)$.

We also observe that the ergodic theorem implies that for any invariant Borel measure ν

$$\int_S \rho(x, G_t)\ d\nu = \int_S \delta(x)\ d\nu.$$

This quantity is a key element of our discussion, in case the measure ν is the probability measure μ determined by ω. More precisely, we let μ be the measure on S determined by the volume form

$$\frac{1}{\text{area}(S)}\ \omega,$$

i.e., normalized so $\mu(S) = 1$.

Definition 2.9 We define the *mean rotation vector* of the isotopy G_t to be

$$\rho_\mu(G_t) = \int_S \rho(x, G_t)\ d\mu = \int_S \delta\ d\mu \in H_1(S, R).$$

In [14] it is shown that the mean rotation vector $\rho_\mu\colon \mathcal{G}_\omega(S) \to H_1(S, R)$ is a homomorphism. If S is an annulus A, as with the pointwise rotation vector, we will often identify G_t with a lift $\tilde{G}$ of G_1 to the universal cover of A and identify $H_1(A, \mathbb{R})$ with $\mathbb{R}$. We will refer to the resulting value of $\rho_\mu(\tilde{G})$ as the *mean rotation number.*

Definition 2.10 An area preserving diffeomorphism $G\colon S \to S$ is called *Hamiltonian* provided there is an isotopy G_t of S with $G_1 = G$ and $G_0 = id$ and $\mathcal{F}([G_t]) = 0 \in H^1(M, R)$. Equivalently (see Proposition 2.11 below), G is Hamiltonian if there is an isotopy G_t with mean rotation vector $\rho_\mu(G_t) = 0 \in H_1(S, R)$. We will denote the group of Hamiltonian diffeomorphisms by $\mathcal{H}(S)$ or simply $\mathcal{H}$.

Note that because of Remark 2.6 it makes no difference in this definition whether the diffeomorphisms G_t are area preserving for $t \in (0,1)$. Also note that any orientation preserving area preserving diffeomorphism $G\colon S^2 \to S^2$ is Hamiltonian since any diffeomorphism is isotopic to the identity and $H^1(S^2, R) = 0$.

The mean rotation vector and the flux homomorphism are closely related. In fact their values on an isotopy G_t are essentially Poincaré duals.

Proposition 2.11 *Suppose $G_t \in \mathcal{G}_\omega(S)$ and $\rho_\mu(G_t) \in H_1(S,R)$ is its mean rotation vector. Then if $u \in H_1(S,R)$,*

$$\rho_\mu(G_t) \wedge u = \frac{\mathcal{F}(G_t)(u)}{\operatorname{area}(S)}$$

where $\wedge$ denotes the intersection pairing in $H_1(S,R)$.

Proof It suffices to prove the result for $u = [\alpha]$ where α is any simple closed curve in S which is not null homologous, since these classes generate $H_1(S,\mathbb{R})$. In particular we may assume the genus of S is at least one. Hence we fix such a simple closed curve α and let $u = [\alpha]$. Since the intersection pairing with u is a linear function, it is clear from the definition of rotation vector that if δ is a homological displacement function associated with $[G_t]$ then

$$\rho_\mu(G_t) \wedge u = \int \rho(x, G_t) \wedge u \; d\mu = \int \delta \wedge u \; d\mu. \tag{1}$$

The intersection pairing with u determines a map $H_1(S,\mathbb{Z}) \to \mathbb{Z}$. This in turn defines an infinite cyclic cover $p\colon \bar{S} \to S$ with fundamental group the kernel of the homomorphism $\pi_1(S) \to H_1(S,\mathbb{Z}) \to \mathbb{Z}$.

Geometrically, this cover is obtained by cutting S along α to form a surface S' with two boundary components (each a copy of α) and letting $\bar{S} = S' \times \mathbb{Z}/\sim$, where $\sim$ is the relation identifying points of the "positive" copy of α in (S', n) with the "negative" copy of α in $(S', n+1)$. The generator of the covering translations is $T\colon \bar{S} \to \bar{S}$ given by $T(x,n) = (x, n+1)$. Let S_n denote the closure in $\bar{S}$ of the set $(S', n)/\sim$. Then each S_n is diffeomorphic to S' and $T^i(\bar{S}_n) = S_{n+i}$. Let $\bar{G}_t\colon \bar{S} \to \bar{S}$ denote the lift of G_t with $\bar{G}_0 = id$.

We will let $\alpha_0 = S_{-1} \cap S_0$ and $\alpha_i = T^i(\alpha_0)$ so $\partial S_n = \alpha_n \cup \alpha_{n+1}$. We would like to have the property that $\bar{G}_1(S_0) \subset \operatorname{Int}(S_{-1} \cup S_0 \cup S_1)$. This may not be the case, but it will be true if S is replaced by a finite cover, specifically by $S(k) = \bar{S}/T^k$ for some large k, since the infinite cyclic cover $\bar{S}(k)$ is just $\bar{S}$ with covering translations generated by T^k.

But both sides of the equality in the statement of the proposition change by a factor of $1/k$ if we change from S to $S(k)$. This is because the flux homomorphism applied to a lift of $[\alpha]$ will not change (see Remark 2.7) while $\text{area}(S(k)) = k\,\text{area}(S)$. For the other side of the equality, the intersection number $\rho(x, G_t) \wedge [\alpha]$ will decrease by a factor of $1/k$ for a lift of α in the cover $S(k)$ so $\rho_\mu(G_t) \wedge [\alpha]$ will decrease by a factor of $1/k$. Hence the result we want for S is true if and only if it is true for $S(k)$. This means we may assume without loss of generality that $\bar{G}_1(S_0) \subset \text{Int}(S_{-1} \cup S_0 \cup S_1)$ and hence that $\bar{G}_1(S_n) \subset \text{Int}(S_{n-1} \cup S_n \cup S_{n+1})$ for all n. Note in particular that $\bar{G}_1(\alpha_0) \cap \alpha_i = \emptyset$ for all $i \neq 0$.

To define the homological displacement function for G_t and hence the rotation vector we considered S as a polygon with edge identifications. We wish now to choose this data in such a way that the closed curve α is the image of two identified edges of this polygon. Then the corresponding homological displacement function δ has a lift $\bar{\delta}\colon \bar{S} \to H_1(S, R)$ which satisfies

$$\bar{\delta}(x) \wedge u = \bar{\delta}(x) \wedge [\alpha] = \begin{cases} -1, & \text{if } x \in S_n \text{ and } G_1(x) \in S_{n-1}; \\ 0, & \text{if } x \in S_n \text{ and } G_1(x) \in S_n; \\ +1, & \text{if } x \in S_n \text{ and } G_1(x) \in S_{n+1}. \end{cases}$$

It follows that if $D_n^+ = \bar{G}_1(S_n) \cap S_{n+1}$ and $D_n^- = \bar{G}_1(S_n) \cap S_{n-1}$, then

$$\begin{aligned} \rho_\mu(G_t) \wedge u &= \int_S \delta \wedge u \; d\mu = \int_{S_0} \bar{\delta} \wedge u \; d\mu \\ &= \mu(D_0^+) - \mu(D_0^-) = \mu(D_0^+) - \mu(D_1^-) \\ &= \frac{\text{area}(D_0^+) - \text{area}(D_1^-)}{\text{area}(S)}. \end{aligned}$$

Let V denote the compact submanifold of $\bar{S}$ bounded by α_0 and $\bar{G}_1(\alpha_1)$. Equivalently $V = (\bar{G}_1(S_0) \cap S_1) \cup (S_0 \setminus \text{Int}(G_1(S_1)))$. Then

$$\int_V \bar{\omega} - \int_{S_0} \bar{\omega} = \text{area}(D_0^+) - \text{area}(D_1^-) = \text{area}(S)\big(\rho_\mu(G_t) \wedge u\big). \tag{2}$$

Since $\bar{S}$ has no de Rham cohomology in dimension 2, we know that there is a one-form θ on $\bar{S}$ such that $d\theta = \bar{\omega}$ where $\bar{\omega}$ is the lift of ω. We can apply Stokes' theorem to

$$\int_V \bar{\omega} - \int_{S_0} \bar{\omega} = \int_{\partial V} \theta \; - \int_{\partial S_0} \theta = \int_{\bar{G}_1(\alpha_1)} \theta \; - \int_{\alpha_1} \theta. \tag{3}$$

Now let $G\colon \alpha \times [0,1] \to S$ be defined by $G(x,t) = G_t(x)$ so that by definition

$$\mathcal{F}(G_t)(u) = \int_{\alpha \times [0,1]} G^*(\omega).$$

If $\bar{G}\colon \alpha \times [0,1] \to \bar{S}$ is the lift of G with $\bar{G}(\alpha \times \{0\}) = \alpha_1$ then again applying Stokes' theorem we have

$$\int_{\alpha\times[0,1]} G^*(\omega) = \int_{\alpha\times[0,1]} \bar{G}^*(\bar{\omega}) = \int_{\alpha\times\{1\}} \bar{G}^*(\theta) - \int_{\alpha\times\{0\}} \bar{G}^*(\theta) = \int_{\bar{G}_1(\alpha_1)} \theta - \int_{\alpha_1} \theta.$$

Combining this with equations (3) and (2) we conclude that

$$\frac{\mathcal{F}(G_t)(u)}{\mathrm{area}(S)} = \rho_\mu(G_t) \wedge u. \qquad \square$$

As an immediate corollary of Proposition 2.11 we have the following.

Corollary 2.12 *If $G_t \in \mathcal{G}_\omega(S)$ and $\rho_\mu(G_t) \in H_1(S,R)$ is its mean rotation vector, then $\mathcal{F}([G_t]) = 0$ if and only if $\rho_\mu(G_t) = 0$.*

The following result is due to Conley and Zehnder [5].

Theorem 2.13 *If $F\colon T^2 \to T^2$ is a Hamiltonian diffeomorphism then F has at least three fixed points. Moreover three points $\{x_1, x_2, x_3\} \subset \mathrm{Fix}(F)$ may be chosen so that F is isotopic to the identity rel$\{x_1, x_2, x_3\}$.*

3 Chain recurrence

We will make use of several facts about area preserving surface homeomorphisms.

Theorem 3.1 *An orientation preserving homeomorphism h of the disk D^2 which leaves invariant a measure whose support intersects the interior of the disk must have an interior fixed point. In particular if h is area preserving then it has an interior fixed point.*

Proof The Poincaré recurrence theorem says that there is a recurrent point in the interior of D^2. The Brouwer plane translation theorem then asserts that the restriction of h to the interior of the disk must have a fixed point. $\square$

Suppose that $h\colon A \to A$ is a homeomorphism of the closed annulus that is isotopic to the identity. Recall that a sequence of points $x_1, \ldots, x_n$ in A is an *ε-chain from x_1 to x_n* if the distance from $h(x_i)$ to x_{i+1} is less than ε for $1 \le i \le n-1$. A point x is *chain recurrent* if there is an ε-chain from x to itself for all $\varepsilon > 0$. The set of chain recurrent points is denoted $\mathcal{R}(h)$. We say that $x, y \in \mathcal{R}(h)$ are in the same *chain transitive component* if for every $\varepsilon > 0$ there is an ε-chain from x to y and another from y to x. If $\mathcal{R}(h) = A$ and if all points in A belong to the same chain transitive component then h *is chain recurrent.*

Let $\tilde{h}\colon \tilde{A} \to \tilde{A}$ be a lift to the universal covering space. Identify $\tilde{A}$ with $\mathbb{R} \times [0,1]$ and denote projection onto the first coordinate by p_1. The *rotation number* of x under $\tilde{h}$ was defined above to be the number obtained from the rotation vector by identifying $H_1(A, \mathbb{R})$ with $\mathbb{R}$. Equivalently it is

$$\rho(x, \tilde{h}) = \lim_{n \to \infty} \frac{p_1(\tilde{F}^n(x)) - p_1(x)}{n},$$

if this limit exists.

The following result is an immediate consequence of Theorem 2.2 of [10].

Theorem 3.2 *Suppose $h\colon A \to A$ is a chain recurrent homeomorphism of the closed annulus that preserves orientation and boundary components. Let $\tilde{h}\colon \tilde{A} \to \tilde{A}$ be a lift to the universal covering space. If $w_0, w_1 \in \tilde{A}$ and $p/q \in \mathbb{Q}$ satisfy*

$$\rho(w_0, \tilde{h}) \leq p/q \leq \rho(w_1, \tilde{h})$$

then h has a periodic point with rotation number p/q. If p and q are relatively prime this point has period q.

We will use the following method for proving that homeomorphisms are chain recurrent.

Definition 3.3 A homeomorphism $h\colon A \to A$ of a closed annulus satisfies the *essential circle intersection property* provided that every essential simple closed curve $\beta \subset S$ satisfies $h(\beta) \cap \beta \neq \emptyset$.

Lemma 3.4 *Suppose that A is a closed annulus and $h\colon A \to A$ is a homeomorphism. Suppose there does not exist a compact disk $D \subset \mathrm{Int}(A)$ and an $n \neq 0$ such that $h^n(D) \subset \mathrm{Int}(D)$. If h satisfies the essential circle intersection property and if ∂A is in the chain recurrent set of h then the chain recurrent set of h is all of A.*

Proof The homeomorphism h possesses a *complete Lyapunov function* ϕ: $A \to \mathbb{R}$ (see [9]). Recall this is a smooth function with $\phi(\mathcal{R}(h))$ compact and nowhere dense and satisfying $\phi(h(x)) < \phi(x)$ for $x \notin \mathcal{R}(h)$. In particular ϕ is constant if and only if all of A is chain recurrent. If ϕ is not constant then there is a regular value $c \in \phi(A)$ with $c \notin \phi(\mathcal{R}(h))$. Hence $\phi^{-1}(c)$ is a finite set of circles in $\mathrm{Int}(D)$, disjoint from their h-image. By hypothesis, these circles must be inessential and so must bound disks $\{D_j\}$ in $\mathrm{Int}(A)$. There is no loss in assuming that these disks are components of $\phi^{-1}(-\infty, c]$. But then $h(\cup D_j)$ is a proper subset of $\cup D_j$ and hence for some $n > 0$ one of the D's is mapped into its interior by h^n, a contradiction. $\square$

4 Hyperbolic structures

Some of our proofs rely on mapping class group techniques that use hyperbolic geometry. In this section we establish notation and recall standard results about hyperbolic structures on surfaces. Let S be a closed orientable surface. We will say that a connected open subset M of S has *negative Euler characteristic* if $H_1(M, \mathbb{R})$ is infinite dimensional or if M is of finite type and the usual definition of Euler characteristic has a negative value. Suppose that $F\colon S \to S$ is an orientation preserving homeomorphism, that $K \subset S$ is a closed F-invariant zero dimensional set and that $S \setminus K$ has negative Euler characteristic. Let $M = S \setminus K$ and $f = F|_M\colon M \to M$.

If K is infinite, the surface M can be written as an increasing union of finitely punctured compact connected subsurfaces M_i whose boundary components determine essential non-peripheral homotopy classes in M. We may assume that boundary curves in M_{i+1} are not parallel to boundary curves in M_i. It is straightforward (see [3]) to put compatible hyperbolic structures on the M_i's whose union defines a complete hyperbolic structure on M. Of course, when K is finite, M also has a complete hyperbolic structure. All hyperbolic structures in this paper are assumed to be complete.

We use the Poincaré disk model for the hyperbolic plane H. In this model, H is identified with the interior of the unit disk and geodesics are segments of Euclidean circles and straight lines that meet the boundary in right angles. A choice of hyperbolic structure on M provides an identification of the universal cover $\tilde{M}$ of M with H. Under this identification covering translations become isometries of H and geodesics in M lift to geodesics in H. The compactification of the interior of the unit disk by the unit circle induces a compactification of H by the 'circle at infinity' S_∞. Geodesics in H have unique endpoints on S_∞. Conversely, any pair of distinct points on S_∞ are the endpoints of a unique geodesic.

Each covering translation T extends to a homeomorphism (also called) $T\colon H \cup S_\infty \to H \cup S_\infty$. The fixed point set of a non-trivial T is either one or two points in S_∞. We denote these point(s) by T^+ and T^-, allowing the possibility that $T^+ = T^-$. If $T^+ = T^-$, then T is said to be *parabolic.* If T^+ and T^- are distinct, then T is said to be *hyperbolic* and we may assume that T^+ is a sink and T^- is a source.

We use the identification of H with $\tilde{M}$ and write $\tilde{f}\colon H \to H$ for lifts of $f\colon M \to M$ to the universal cover. A fundamental result of Nielsen theory is that every lift $\tilde{f}\colon H \to H$ extends uniquely to a homeomorphism (also called) $\tilde{f}\colon H \cup S_\infty \to H \cup S_\infty$. (A proof of this fact appears in Proposition 3.1 of [20]). If $f\colon M \to M$ is isotopic to the identity then there is a unique lift $\tilde{f}$, called the *identity lift,* that commutes with all covering translations and whose extension over S_∞ is the identity.

For any extended lift $\tilde{f}\colon H \cup S_\infty \to H \cup S_\infty$ there is an *associated action $\tilde{f}_\#$ on geodesics in H* defined by sending the geodesic with endpoints P and Q to the geodesic with endpoints $\tilde{f}(P)$ and $\tilde{f}(Q)$. The action $\tilde{f}_\#$ projects to an *action $f_\#$ on geodesics in M.*

We occasionally allow ∂S to be non-empty. There is still a hyperbolic structure on M but now the universal cover $\tilde{M}$ is naturally identified with the intersection of H with the interior of the convex hull of a Cantor set $C \subset S_\infty$. The frontier of $\tilde{M}$ in $H \cup S_\infty$ is the union of C with the full pre-image of ∂S. If f is isotopic to the identity, then the identity lift commutes with all covering translations and extends to a homeomorphism of the frontier of $\tilde{M}$ that fixes C.

5 Lifting to annuli

Throughout this section $F\colon S \to S$ is a diffeomorphism of a closed surface S. An isolated end of an open set $U \subset S$ has neighborhoods of the form $N(E) = S^1 \times [0,1)$. The set $fr(E) = cl_S(N(E)) \setminus N(E)$, called the *frontier of* E, is independent of the choice of $N(E)$.

Lemma 5.1 *Suppose that $M \subset S$ is an open F-invariant set and that E is an isolated end of M whose frontier is contained in Fix(F). Then E can be compactified by a circle C and $f = F|_M$ can be extended continuously over C. If the frontier of E is not a single point then $f|_C =$ identity.*

Proof If $fr(E)$ is a single point of $\mathrm{Fix}(F)$ then we compactify by blowing up that point using the differentiability of f. We may therefore assume that $fr(E)$ is a non-trivial compact connected subset of $\mathrm{Fix}(F)$. The existence of

C and a continuous extension of f over C is a consequence of the theory of prime ends (see [25] for a good modern exposition). Moreover, a dense set $P \subset C$ corresponds to accessible points in $fr(E)$, i.e. for any $p \in P$ there is an embedded arc $\alpha\colon [0,1] \to M \cup C$ with $\alpha(0) = p$ and $\alpha((0,1]) \subset M$ and such that the restriction of α to $(0,1]$ extends to a continuous map $\hat{\alpha}\colon [0,1] \to S$ with $\hat{\alpha}(0) \in fr(E)$. A result of Mather (Theorem 18 of [25]) asserts that two such extensions, $\hat{\alpha}_1$ and $\hat{\alpha}_2$ correspond to the same point of C if and only if $\hat{\alpha}_1(0) = \hat{\alpha}_2(0)$ and $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are isotopic by an isotopy with support in M. A consequence of this is that the homeomorphism $f|_C$ depends only on the isotopy class of F relative to $fr(E)$. By Lemma 4.1 of [21] F is isotopic rel $fr(E)$ to a homeomorphism that is the identity on a neighborhood of $fr(E)$. Thus $f|_C$ is the identity as desired. □

Lemma 5.2 *Suppose that M is a connected F-invariant open subset of S whose frontier, $fr(M) = cl(M) \setminus M$, is contained in $\mathrm{Fix}(F)$. If $S = T^2$, assume that $fr(M) \neq \emptyset$. Denote $F|_M$ by f. If $\beta \subset M$ is a simple closed curve that is essential in M and if $f(\beta)$ is homotopic to β in M then there is a closed annulus $\mathbb{A}$, a homeomorphism $\hat{f}\colon \mathbb{A} \to \mathbb{A}$ that is isotopic to the identity and a covering map $\pi\colon \mathrm{Int}(\mathbb{A}) \to M$ such that*

1. *$\hat{f}|_{\mathrm{Int}\,\mathbb{A}}$ is a lift of f.*
2. *$\pi^{-1}(\beta)$ contains an essential simple closed curve $\hat{\beta}$.*
3. *if β is non-peripheral then $\hat{f}|_{\partial\mathbb{A}}$ depends only on the isotopy class of f. In particular, if β is non-peripheral and $f\colon M \to M$ is isotopic to the identity then $\hat{f}|_{\partial\mathbb{A}}$ is the identity.*
4. *if $f\colon M \to M$ is isotopic to the identity, and either $S \neq S^2$ or $S \setminus M$ contains at least three points then $\hat{f}|_{\partial\mathbb{A}}$ has fixed points*

We will call π the canonical cyclic cover *associated to β and $\hat{f}$ the* canonical cyclic lift *of f.*

Proof If M is an open annulus then $\mathrm{Int}(\mathbb{A}) = M$, $\hat{f}|_{\mathrm{Int}(\mathbb{A})} = f$ and it suffices to compactify the ends of M by circles and extend f over these circles. Lemma 5.1 therefore completes the proof in this case. We may now assume that M has negative Euler characteristic.

Suppose at first that β is not peripheral. Let H be the universal cover of M, let $\tilde{\beta}$ be a lift of β and let $T\colon H \cup S_\infty \to H \cup S_\infty$ be the indivisible covering translation that fixes the endpoints of $\tilde{\beta}$. Since $f(\beta)$ is homotopic to β, there is an extended lift $\tilde{f}\colon H \cup S_\infty \to H \cup S_\infty$ that commutes with T. The choice of $\tilde{f}$ is well defined up to composition with iterates of T because these are the only covering translations that commute with T. Let $\mathbb{A}$ denote the quotient space obtained by dividing $(H \cup S_\infty) \setminus T^\pm$ by the action of T. Denote by $\hat{f}\colon \mathbb{A} \to \mathbb{A}$ the homeomorphism induced by $\tilde{f}$. It is independent of the choice of lift $\tilde{f}$, so $\hat{f}$ is unique. The extension $\tilde{f}|_{S_\infty}$ depends only on the isotopy class of f and the choice of lift $\tilde{f}$. Since $\hat{f}$ is independent of the choice of lift, $\hat{f}|_{\partial\mathbb{A}}$ depends only on the isotopy class of f. If f is isotopic to the identity then we may take $\tilde{f}$ to be the identity lift. Thus $\tilde{f}|_{S_\infty}$ and $\hat{f}|_{\partial\mathbb{A}}$ are the identity.

If β is peripheral, compactify the isolated end corresponding to β using Lemma 5.1 and repeat the previous argument. If f is isotopic to the identity then the identity lift fixes a Cantor set in S_∞ so $\hat{f}$ fixes a Cantor set in $\partial\mathbb{A}$. $\square$

The following lemma relates periodic points of $\hat{f}$ to periodic points of f.

Lemma 5.3 *Suppose that $\hat{f}\colon \mathbb{A} \to \mathbb{A}$ is the canonical cyclic lift of $f\colon M \to M$ associated to a simple closed curve β and that $\hat{x}$ is a periodic point of $\hat{f}$ with period q. If $\rho(\hat{x}, \hat{f})$ is non-zero and does not occur as $\rho(\hat{z}, \hat{f})$ for some $\hat{z} \in \partial\mathbb{A}$, then the image $x \in M$ of $\hat{x}$ is periodic for f with period q.*

Proof We use the notation of the proof of Lemma 5.2. If M is an open annulus then the lemma is obvious so assume that M has negative Euler characteristic. There is a lift $\tilde{x} \in H$ of $\hat{x}$ such that $\tilde{f}^q(\tilde{x}) = T^p(\tilde{x})$ where $\rho(\tilde{x}, \tilde{f}) = p/q$. Since $f^q(x) = x$, it suffices to show that $f^s(x) \neq x$ for $1 \leq s \leq q-1$. Suppose to the contrary that there exists $1 \leq s \leq q-1$ and a covering translation $S\colon H \to H$ such that $\tilde{f}^s(\tilde{x}) = S(\tilde{x})$. Recall that the covering translation $\tilde{f}^q S \tilde{f}^{-q}$ is denoted $\tilde{f}^q_\#(S)$. We have

$$\begin{aligned}\tilde{f}^q_\#(S)(T^p(\tilde{x})) &= \tilde{f}^q_\#(S)(\tilde{f}^q(\tilde{x})) = \tilde{f}^q(S(\tilde{x})) = \tilde{f}^{q+s}(\tilde{x}) \\ &= \tilde{f}^s\tilde{f}^q(\tilde{x}) = \tilde{f}^s T^p(\tilde{x}) = T^p(\tilde{f}^s(\tilde{x})) = T^p(S(\tilde{x})).\end{aligned}$$

Covering translation that agree on a point are equal so $\tilde{f}^q_\#(S) = T^p S T^{-p}$. Let $z \in S_\infty$ be the unique attracting fixed point for the action of S. Then $T^p(z)$ and $\tilde{f}^q(z)$ are both the unique attracting fixed point for the action of $\tilde{f}^q_\#(S) = T^p S T^{-p}$. Thus $\tilde{f}^q(z) = T^p(z)$. We are assuming that $\hat{f}^s(\hat{x}) \neq \hat{x}$ and hence that $S \neq T^j$ for any j. In particular, $z \neq T^\pm$ and z projects to a point $\hat{z} \in \partial\mathbb{A}$ such that $\rho(\hat{z}, \hat{f}) = p/q$. This contradiction completes the proof. $\square$

We use the following lemma in several places to prove the existence of periodic points of arbitrarily high period.

Lemma 5.4 *Suppose that $F \in \mathcal{H}(S)$ and that M is a connected F-invariant open subset of S whose frontier, $fr(M) = cl(M) \setminus M$, is contained in $Fix(F)$. If $S = T^2$, assume that $fr(M) \neq \emptyset$. Denote $F|_M$ by f. Suppose also that $\beta \subset M$ is a simple closed curve whose homotopy class in M is non-trivial and fixed by f and that the lift $\hat{f}\colon \mathbb{A} \to \mathbb{A}$ to the canonical cyclic cover associated to β has points with two distinct rotation numbers. Then there are periodic points of f of every sufficiently high period.*

Proof Let $\pi\colon \operatorname{int}(\mathbb{A}) \to M$ be the covering map and let $\hat{\omega} = \pi^*(\omega)$. The smooth measure on the interior of $\mathbb{A}$ determined by $\hat{\omega}$ is $\hat{f}$-invariant.

We want first to show that $\hat{f}\colon \mathbb{A} \to \mathbb{A}$ has the essential circle intersection property. Suppose that $\hat{\alpha}$ is an essential simple closed curve in $\mathbb{A}$ and that $\hat{f}(\hat{\alpha}) \cap \hat{\alpha} = \emptyset$. There is no loss in assuming that $\hat{\alpha}$ is smooth and contained in the interior of $\mathbb{A}$.

The region bounded by $\hat{\alpha}$ and $\hat{f}(\hat{\alpha})$ is an annulus that we denote $\hat{A}_0$. Choose a smooth parametrization $\hat{h}\colon S^1 \times [0,1] \to \hat{A}_0$ such that $\hat{h}(s,1) = \hat{f}(\hat{h}(s,0))$ for each $s \in S^1$ and let $h\colon S^1 \times [0,1] \to M$ be $h = \pi \circ \hat{h}$. Then

$$\int_{S^1 \times [0,1]} h^*(\omega) = \int_{S^1 \times [0,1]} \hat{h}^*(\hat{\omega}) = \int_{\hat{A}_0} \hat{\omega} \neq 0. \tag{4}$$

On the other hand since $F \in \mathcal{H}(S)$ there is an isotopy F_t with $F_0 = id$ and $F_1 = F$. In the case $S = T^2$ or S^2 the frontier of M is non-empty and consists of fixed points. Hence we may assume there is a point $p \in fr(M)$ such that $F_t(p) = p$ for all t, since any diffeomorphism of S^2 or T^2, which is isotopic to the identity and has a fixed point p, is isotopic to the identity relative to p.

Let $\alpha = \pi \circ \hat{\alpha}\colon S^1 \to M$. If $F_\alpha\colon S^1 \times [0,1] \to S$ is given by $F_\alpha(s,t) = F_t(\alpha(s))$ then

$$\int_{S^1 \times [0,1]} F_\alpha^*(\omega) = \mathcal{F}(F_t)([\alpha]) = 0 \tag{5}$$

by hypothesis.

The two maps $h, F_\alpha\colon S^1 \times [0,1] \to S$ agree on their ends, so we may glue them together to form a map on the torus, $G\colon T^2 \to S$. The map G must have degree zero because any map of a torus into a surface of negative Euler characteristic has degree zero and in the cases that $S = S^2$ or $S = T^2$ we have the fact that the image of G lies in $S \setminus \{p\}$. Consequently

$$0 = \int_{T^2} G^*(\omega) = \int_{S^1 \times [0,1]} F_\alpha^*(\omega) - \int_{S^1 \times [0,1]} h^*(\omega)$$

which contradicts equations (4) and (5). Thus we have shown that $\hat{f}$ has the essential circle intersection property.

We would like to show that $\hat{f}$ is chain recurrent. If we knew that $\partial\mathbb{A} \subset \mathcal{R}(\hat{f})$ then this would follow from Lemma 3.4. It turns out we can reduce to the case $\partial\mathbb{A} \subset \mathcal{R}(\hat{f})$ by extending $\hat{f}$ to a slightly larger annulus without affecting the periods of periodic points. If the rotation number of the restriction of $\hat{f}$ to a boundary component is irrational then that component is in $\mathcal{R}(\hat{f})$ since this circle homeomorphism must be conjugate to either an irrational rotation or a Denjoy type example, both of which have all points chain recurrent. If the rotation number on a boundary circle is rational there is a periodic point p_0 on that circle. In this case there is an isotopy h_t on the circle, with h_0 the given homeomorphism and h_1 a finite order homeomorphism, and which has the property that $\mathrm{Per}(h_t) = \mathrm{Per}(h_0)$ for all $t \in [0,1)$. We now attach a collar neighborhood $S^1 \times [0,1]$ to the boundary of A and extend $\hat{f}$ to the union by having it preserve concentric circles of the collar and act on them in the way prescribed by the isotopy h_t. We can do this to both boundary components if necessary. Call the enlarged annulus $\mathbb{B}$ and the extended map $\hat{g}\colon \mathbb{B} \to \mathbb{B}$. Then $\partial\mathbb{B} \subset \mathcal{R}(\hat{g})$ and $\hat{g}$ has a point of period q with rotation number $\frac{p}{q}$ if and only if there is a point of period q with rotation number $\frac{p}{q}$ for $\hat{f}$.

It is also the case that $\hat{g}$ satisfies the essential circle intersection property. This is because if γ is an essential circle with $\hat{g}(\gamma) \cap \gamma = \emptyset$ and $x \in \gamma$ then points $z \in \alpha(x)$ and $w \in \omega(x)$ lie in $\mathcal{R}(\hat{g})$ but in different components of $\mathcal{R}(\hat{g})$ since γ separates them. On the other hand $\mathcal{R}(\hat{g}) \cap (\mathbb{B} \setminus \mathbb{A})$ consists of all periodic points in $\mathbb{B}\setminus\mathbb{A}$ and has two components, one containing each of the components of $\partial\mathbb{B}$. It follows that γ cannot intersect $\mathbb{B} \setminus \mathbb{A}$ since any point of $\mathbb{B} \setminus \mathbb{A}$ has both its alpha and omega limit sets in the same component of $\mathcal{R}(\hat{g}) \cap (\mathbb{B} \setminus \mathbb{A})$. Thus $\gamma \subset A$ and since $\hat{f}$ has the essential circle intersection property, so does $\hat{g}$. Also if D is a closed disk and $\hat{g}^n(D) \subset \mathrm{Int}(D)$ then $D \cap (\mathbb{B}\setminus\mathbb{A}) = \emptyset$ because the omega limit set of any point in $\mathbb{B} \setminus \mathbb{A}$ lies in the component of $\mathrm{Per}(h_0)$ containing $\partial\mathbb{B}$.

By Lemma 5.3 it suffices to show that every sufficiently large q occurs as the period of a periodic point for $\hat{f}$ with non-zero rotation number not realized on $\partial\mathbb{A}$. This property holds for $\hat{f}$ if and only if holds for $\hat{g}$ so Lemma 3.4 and Theorem 3.2 complete the proof. □

6 Normal form

In this section we introduce a normal form for certain mapping classes of infinitely punctured surfaces.

Definition 6.1 An orientation preserving homeomorphism $F\colon S \to S$ of an orientable closed surface *has a normal form relative to its fixed point set* if there is a finite set R of simple closed curves in $M = S \setminus \mathrm{Fix}(F)$ and a homeomorphism ϕ isotopic to F rel $\mathrm{Fix}(F)$ such that:

(1) ϕ permutes disjoint open annulus neighborhoods $A_j \subset M$ of the elements $\gamma_j \in R$.

Let $\{S_i\}$ be the components of $S \setminus \cup A_j$, let $X_i = \mathrm{Fix}(F) \cap S_i$, let $M_i = S_i \setminus X_i$ and let r_i be the smallest positive integer such that $\phi^{r_i}(M_i) = M_i$. Note that $r_i = 1$ if $X_i \neq \emptyset$.

(2) If X_i is infinite then $\phi|_{S_i} =$ identity.

(3) If X_i is finite then M_i has negative Euler characteristic and $\phi^{r_i}|_{M_i}$ is either pseudo-Anosov or periodic. In the periodic case, $\phi^{r_i}|_{M_i}$ is an isometry of a hyperbolic structure on M_i.

We say that ϕ is a *normal form for* F and that R is *the set of reducing curves for* ϕ.

If R has minimal cardinality among all sets of reducing curves for all normal forms for F, then we say that R *is a minimal set of reducing curves.* The following lemma asserts that, up to isotopy rel $\mathrm{Fix}(F)$, there is a unique minimal set $R(F)$ of reducing curves.

Lemma 6.2 *If F has a canonical form relative to its fixed point set then the minimal reducing set $R(F)$ is well defined up to isotopy rel* $\mathrm{Fix}(F)$.

Proof Suppose that R is a minimal reducing set for the normal form ϕ. Let $Y(F)$ be a finite subset of $\mathrm{Fix}(F)$ that contains each finite X_i and at least two points from each infinite X_i. Then ϕ is a Thurston normal form [29] for the isotopy class of F relative to $Y(F)$ and R is a minimal set of reducing curves for this relative isotopy class. It is shown in section 2 of [22] that R is well defined up to isotopy rel $Y(F)$. Since this holds for all $Y(F)$, R is well defined up to isotopy rel $\mathrm{Fix}(F)$. □

Theorem 1.2 *Every orientation preserving diffeomorphism $F\colon S \to S$ of a closed orientable surface has a normal form up to isotopy relative to its fixed point set.*

Proof By Lemma 4.1 of [21] there is a neighborhood W of the accumulation set $A(F)$ of $\mathrm{Fix}(F)$ such that:

- W and $S \setminus W$ have finitely many components.
- $F|_W\colon W \to S$ is isotopic to the inclusion relative to $\mathrm{Fix}(F) \cap W$.
- if $x \in W \setminus \mathrm{Fix}(F)$ then the path from $F(x)$ to x determined by this isotopy is contained in $S \setminus \mathrm{Fix}(F)$.

We may assume without loss that each component of W intersects $\mathrm{Fix}(F)$ in an infinite set. The isotopy extension theorem (Theorem 5.8 of [26]) and the second and third item imply that F is isotopic rel $\mathrm{Fix}(F)$ to $\phi_1\colon S \to S$ satisfying $\phi_1|_W = \text{identity}$. By [4] each component of $S \setminus W$ is invariant under the action of ϕ_1. If some component of $S \setminus W$ is a disk that contains at most one element of $\mathrm{Fix}(F)$ then add it to W. After a further isotopy rel $\mathrm{Fix}(F)$, we may assume that $\phi_1|_W$ is still the identity. Suppose that some component of $S \setminus W$ is an open annulus A that is disjoint from $\mathrm{Fix}(F)$. If $\phi_1|_{(W \cup A)}$ is isotopic to the identity rel $\mathrm{Fix}(F)$, then add A to W. After a further isotopy rel $\mathrm{Fix}(F)$, we may assume that $\phi_1|_W$ is still the identity. If $\phi_1|_{(W \cup A)}$ is not isotopic to the identity rel $\mathrm{Fix}(F)$ then after a further isotopy rel $\mathrm{Fix}(F)$ we may assume that $\phi_1|_A$ is a non-trivial Dehn twist.

Denote $S \setminus \mathrm{Fix}(F)$ by M. At this point $M \setminus (W \cap M)$ consists of a finite number of unpunctured annuli on which ϕ_1 acts by Dehn twists and finitely many, finitely punctured subsurfaces of negative Euler characteristic. Apply the Thurston decomposition theorem to each of the non-annular subsurfaces to produce $\phi\colon S \to S$. Let R_1 be the union of the core curves of the Dehn twist annuli and Let R be the union of R_1 with the boundary curves of W that are not boundary curves of a Dehn twist annuli. (Those curves are already accounted for in R_1.) After modifying ϕ on annular neighborhoods of the reducing curves we may assume that (1) is satisfied. Properties (2) and (3) are immediate from the construction. □

We now restrict our attention to the case that F is isotopic to the identity.

Lemma 6.3 *Assume that ϕ, R, M_i and r_i are as in Definition 6.1 and that F is isotopic to the identity. Then:*

- *each* $\gamma \in R$ *is* ϕ*-invariant.*
- *each* $r_i = 1$.
- *if* $\phi|_{M_i}$ *is periodic then it is the identity.*

Proof If γ bounds a disk D in S then $D \cap \mathrm{Fix}(F)$ is non-trivial. The disk $F(D)$, and hence the disk D' bounded by $\phi(\gamma)$, has the same intersection with $\mathrm{Fix}(F)$ as does D. Since $\phi(\gamma)$ and γ are either equal or disjoint and not parallel in M, they must be equal. This completes the proof of the first item if γ is inessential in S. Since D contains fixed points, ϕ does not interchange the components of $S \setminus \gamma$. In particular, if M_i is adjacent to the annulus A_j containing γ, then M_i is ϕ-invariant. This proves the second item if some component of ∂M_i is inessential in S.

Suppose that M_i is ϕ-invariant and that X and Y are distinct components of ∂M_i that are essential in S. If $Y = \phi(X)$ as sets then X and Y must be parallel in S. Up to isotopy in S, ϕ reverses the orientation on the simple closed curve X, in contradiction to the assumption that F is isotopic to the identity. We have now shown that if M_i is ϕ-invariant then every component of ∂M_i is ϕ-invariant.

The lemma is obvious if $\mathrm{Fix}(F) = \emptyset$ so assume that M has at least one puncture. Each puncture in M is contained in exactly one M_i which must therefore be ϕ-invariant. If M_i is ϕ-invariant then each annulus A_j adjacent to M_i is ϕ-invariant and so each M_k adjacent to A_j is ϕ-invariant. The first two items follow by induction.

For the third item we may assume that M_i is finitely punctured. For this argument it is convenient to replace the punctures in M_i with boundary components to form a compact surface $\hat{M}_i$. If $\phi|_{M_i}$ is periodic, then ϕ extends to an isometry of a hyperbolic structure on $\hat{M}_i$ that setwise fixes each boundary component. Since F induces the identity on the first homology of S, $\phi|_{\hat{M}_i}$ induces the identity on first homology of $\hat{M}_i$. Any isolated fixed point of $\phi|_{\hat{M}_i}$ has positive fixed point index so the Lefschetz theorem implies that not all fixed points are isolated. Since $\phi|_{\hat{M}_i}$ is an orientation preserving isometry, it must be the identity. □

Remark 6.4 If each $r_i = 1$ and if each $\phi|_{M_i}$ is the identity, then we may assume that $\phi|_{A_j}$ is an iterate of a Dehn twist about γ_j.

7 A special case

In this section we prove a special case of Theorem 1.1.

Proposition 7.1 *Suppose $F\colon S \to S$ is a non-trivial, Hamiltonian diffeomorphism of a closed oriented surface S and that if $S = S^2$ then $\mathrm{Fix}(F)$ contains at least three points. If F is not isotopic to the identity relative to $\mathrm{Fix}(F)$, then there exist $p > 0$ such that F has a periodic point of period k for every $k \geq p$.*

Proof Let ϕ be a canonical form for F with minimal reducing set $R(F)$. Let Y be a finite subset of $\mathrm{Fix}(F)$ that contains all finite X_i and at least two points from each infinite X_i. Then ϕ is a Thurston normal form for F relative to Y. If some $\phi|_{M_i}$ is pseudo-Anosov then ([29], Expose 10 of [7]) there exist $p > 0$ such that F has a periodic point of period k for every $k \geq p$ and we are done.

By Lemma 6.3 and Remark 6.4, ϕ is a composition of non-trivial Dehn twists along the elements of $R(F)$. Thus it suffices to show that if $R(F)$ is non-empty, then there exist $p > 0$ such that F has a periodic point of period k for every $k \geq p$. So suppose $\beta \in R(F)$. We work with $M_Y = S \setminus Y$, $f_Y = F|_{M_Y}$ and $\phi_Y = \phi|_{M_Y}$. Let $\hat{f}_Y, \hat{\phi}_Y\colon \mathbb{A} \to \mathbb{A}$ be the canonical cyclic lifts corresponding to $\beta \subset M_Y$.

We wish to apply Lemma 5.4 to $\hat{f}_Y$ which will give points of every sufficiently large period for f_Y. It suffices to show that the rotation numbers of lifts of boundary points for any lift of $\hat{f}_Y\colon \mathbb{A} \to \mathbb{A}$ to its universal covering space are different for the two components of the boundary. By Lemma 5.2(2), $\hat{f}_{Y|\partial\mathbb{A}} = \hat{\phi}_{Y|\partial\mathbb{A}}$ so we can calculate the rotation number using ϕ in place of F. Let A_j be the annulus containing β and let β_0 and β_1 be the boundary components of A_j. Let M_0 and M_1 be the (possibly equal) pieces in the Thurston decomposition whose boundaries contain β_0 and β_1 respectively.

If $\alpha_0 \neq \beta_0$ is a component of ∂M_0 then we can join α_0 and β_0 by an arc on which ϕ is the identity. We can lift this arc and α_0 and β_0 to H in such a way that the lift $\tilde{\beta}_0$ has the same endpoints as the lift $\tilde{\beta}$ of β. There is a lift of ϕ which is the identity on $\tilde{\beta}_0$, $\tilde{\alpha}_0$ and the lifted arc joining them. Hence it is clear that the endpoints of the lift $\tilde{\alpha}_0$ which lie in S_∞ will have the same rotation number as any point of $\tilde{\beta}_0$ for this lift of ϕ and thus for any lift of ϕ. If there are no components of the boundary of M_0 other than β_0 then we may choose an arc in H joining a point of $\tilde{\beta}_0$ to a point in S_∞ which crosses no lift of a reducing curve. Again, one lift of ϕ will be the identity on this arc and on

$\tilde\beta_0$. So again points of the boundary of S_∞ which lie on the same side of $\tilde\beta$ as $\tilde\beta_0$ all have the same rotation number as points of $\tilde\beta_0$ for any lift of ϕ.

The same argument shows points in S_∞ on the other side of $\tilde\beta$ have the same rotation number as points of $\tilde\beta_1$. It follows that points in the two boundary components of the universal covering of $\mathbb{A}$ have rotation numbers related like the rotation numbers of opposite sides of an annulus with a non-trivial Dehn twist. i.e. they differ by a non-zero integer. The proof now follows from Theorem 3.2 and Lemma 5.4. □

8 Translation and Reeb classes

We will denote the orbit, backward orbit and forward orbit of x by $\mathcal{O}(x)$, $\mathcal{O}^-(x)$ and $\mathcal{O}^+(x)$ respectively. Unless stated otherwise, all maps are assumed to be orientation preserving homeomorphisms.

In this section we recall and enhance some results from [20].

Definition 8.1 Let $Tr\colon \mathbb{R}^2 \to \mathbb{R}^2$ be translation to the right by one unit. If $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ is an orientation preserving homeomorphism and if $X = \cup_{i=1}^r \mathcal{O}(x_i)$ for $x_i \in \mathbb{R}^2$ with distinct orbits, then we say that f *is a translation class relative to* X if the isotopy class of f relative to X is conjugate to the isotopy class of Tr relative to the union of some (and hence any - see Lemma 2.1 of [20]) r distinct orbits of Tr.

Definition 8.2 Let $R\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a homeomorphism that preserves $\mathbb{R} \times \{0,1\}$ and that agrees with translation to the left [respectively right] by one unit on $\mathbb{R} \times \{1\}$ [respectively $\mathbb{R} \times \{0\}$]. If $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ is an orientation preserving homeomorphism and if $X = \mathcal{O}(x) \cup \mathcal{O}(y)$ for some $x, y \in \mathbb{R}^2$ with distinct f orbits then we say that f *is a Reeb class relative to* X if the isotopy class of f relative to X is conjugate to the isotopy class of R relative to the union of the R-orbits of $(0,0)$ and $(0,1)$.

Parts (1) and (2) of the following theorem are Corollary 2.2 and Theorem 2.6 of [20] respectively.

Theorem 8.3 *If $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ is fixed point free then*

(1) *f is a translation class relative to any one of its orbits.*

(2) *f is either a translation class or a Reeb class relative to any two of its orbits.*

Several of our arguments proceed by contradiction and so require a method for proving that the conclusions of Theorem 8.3 fail. Lemma 8.7 below provides this method.

Definition 8.4 Suppose that $S = \mathbb{R}^2$ or $S = \mathrm{int}(\mathbb{A})$ and that $f\colon S \to S$ is an orientation preserving homeomorphism. An arc $\beta' \subset S$ connecting x to $f(x)$ is called a *translation arc for* x if $f(\beta') \cap \beta' = f(x)$. For our applications there will be a closed infinite discrete set $X \subset S$ that is the union of orbits and that intersects β' exactly in x and $f(x)$. Equip $M = S \setminus X$ with a complete hyperbolic structure. We say that the unique geodesic $\beta \subset M$ in the homotopy class of β' is a *translation arc geodesic* for x relative to X.

Remark 8.5 In [20] a more general object called a homotopy translation arc is considered but for our present purposes this simpler definition will suffice.

Definition 8.6 An open disk U is called a *free disk* for the homeomorphism f if $f(U) \cap U = \emptyset$.

Lemma 8.7 *Assume that $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ is a fixed point free orientation preserving homeomorphism. Then:*

(1) *If X is a single orbit and f is a translation class relative to X then translation arc geodesics for elements of X are unique.*

(2) *If $X = \mathcal{O}(x_1) \cup \mathcal{O}(x_2)$ and f is a Reeb class relative to X then translation arc geodesics for $x \in X$ are unique. Moreover:*

- *no free disk contains two elements of X.*
- *a translation arc geodesic for $f^i(x_1)$ can not intersect a translation arc geodesic for $f^j(x_2)$.*

(3) *If $X = \mathcal{O}(x_1) \cup \mathcal{O}(x_2)$ and f is a translation class relative to X then there is a compact set K with the following property: if β_1 is a translation arc geodesic for an element of $\mathcal{O}^-(x_1)$ and β_2 is a translation arc geodesic for an element of $\mathcal{O}^+(x_2)$ and if both β_1 and β_2 are disjoint from K then $\beta_1 \cap \beta_2 = \emptyset$.*

Proof of Lemma 8.7 (1) is Theorem 2.2 of [20].

To prove (2) we will prove the a priori stronger statement that there is a unique geodesic arc $\alpha \subset M$ connecting x to $f(x)$ with the property that $f_\#(\alpha) \cap \alpha = f(x)$. (It is immediate from the definitions that this property depends only on the conjugacy class of the relative isotopy class of f which is not the case if we insist that α is a translation arc geodesic.) It suffices to assume that $f = R$, $x_1 = (0,0)$ and $x_2 = (0,1)$. We will give the argument for x_2, all other cases being completely analogous. We may assume that $\mathbb{R} \times \{0,1\}$ is a union of geodesics. Suppose that $\alpha \neq [0,1] \times \{1\}$. After an isotopy that preserves $\mathbb{R}\times\{0,1\}$, we may assume (c.f. Lemma 3.5-(3) in [20]) that $R(\alpha)$ is a geodesic. In particular $R(\alpha)$ and α intersect only in their common endpoint. Let α_0 be the maximal initial segment of α whose interior is disjoint from $\mathbb{R} \times \{0,1\}$ and let w be the terminal endpoint of α_0. Since regions bounded by a pair of geodesic arcs must contain punctures, $\alpha_0 \neq \alpha$. In particular $R(\alpha_0)$ is disjoint from α_0. There are three possibilities: $w \in (-1,1) \times \{1\}$; $w \in (k, k+1) \times \{1\}$ for some $k \neq -1, 0$; and $w \in \mathbb{R} \times \{0\}$. The first option is impossible since the region bounded by a pair of geodesic arcs must contain punctures. The latter two are impossible because in those cases $R(\alpha_0)$ could not be disjoint from α_0. This proves the main statement in (2). The first bulleted item follows from the standard construction of translation arcs (see, for example, Lemma 4.1 in [20]) since the existence of a free disk containing two elements of X would give rise to two translation arcs not isotopic relative to X. The second bulleted item follows from the fact that $\mathbb{R} \times \{0\}$ and $\mathbb{R} \times \{1\}$ are disjointly embedded lines.

As above, the proof of (3) will use the fact that $f_{\#}(\beta_i) \cap \beta_i$ is exactly the terminal endpoint of β_i but not the fact that β_i is isotopic to a translation arc. It suffices to assume that $f = Tr$, $x_1 = (0,0)$ and $x_2 = (\frac{1}{2}, 0)$. We may assume that $\mathbb{R} \times \{0\}$ is a union of geodesics. Let K be any disk that contains $[0,1] \times \{0\}$ and that has geodesic boundary. We may assume that $Tr(\beta_1)$ is geodesic and in particular that $Tr(\beta_1)$ and β_1 have no interior intersections. In particular, if b is a subpath of β_1 that intersects $\mathbb{R} \times \{0\}$ exactly in its endpoints then the endpoints of b span an interval on $\mathbb{R}\times\{0\}$ with length ≤ 1. It follows that the union of all $[k, k+1] \times \{0\}$'s that intersect β_1 is connected. Since $\beta_1 \cap [0,1] \times \{0\} = \emptyset$, β_1 is disjoint from $[0,\infty) \times \{0\}$ and so can be isotoped into the left half plane. The analogous argument shows that β_2 can be isotoped into the right half plane. As geodesics have minimal intersections in their isotopy classes, this completes the proof. □

We only use the following result for $k = 2$ but the proof is the same for general k.

Definition 8.8 For all $k \geq 1$ let $R_k \colon \mathbb{R}^2 \to \mathbb{R}^2$ be a homeomorphism that preserves $\mathbb{R} \times \{0, 1, \ldots, k\}$ and that agrees with translation to the left by one unit on $\mathbb{R}\times\{j\}$ for odd $1 \leq j \leq k$ and with translation to the right by one unit on $\mathbb{R} \times \{j\}$ for even $0 \leq j \leq k$. If $f \colon \mathbb{R}^2 \to \mathbb{R}^2$ is an orientation preserving homeomorphism and if $X = \cup_{i=0}^{k} \mathcal{O}(x_i)$ for $x_i \in \mathbb{R}^2$ with distinct orbits then we say that that f *is a multiple Reeb class relative to* X if the isotopy class of f relative to X is conjugate to the isotopy class of R_k relative to $\cup_{j=0}^{k}\mathcal{O}((0,j))$.

Disjoint subsets Y and Z of a circle C *alternate around* C if any interval in C with distinct endpoints in Y [respectively Z] contains an element of Z [respectively Y].

Lemma 8.9 *Suppose that $f\colon D^2 \to D^2$ is an orientation preserving homeomorphism of the disk whose restriction to $\mathrm{Int}(D^2)$ is fixed point free. Suppose also that $\{x_0, \ldots x_k\} \subset \mathrm{Int}(D^2)$, $k \geq 1$, satisfy:*

(1) *the α and ω limit sets are single points $\alpha(x_i), \omega(x_i) \in \partial D^2$. The $\alpha(x_i)$'s and $\omega(x_i)$'s are all distinct.*

(2) *the elements of $\{\alpha(x_i)\}$ and $\{\omega(x_i)\}$ alternate around ∂D^2.*

Then $f|_{\mathrm{Int}(D^2)}$ is a multiple Reeb class relative to $X = \cup_{i=0}^{k}\mathcal{O}(x_i)$.

Proof of Lemma 8.9 We will prove there are disjoint properly embedded lines $B_i \subset \mathrm{int}D^2$ that contain $\mathcal{O}(x_i)$ and that are f-invariant up to isotopy relative to X. The lemma then follows from (2).

By Theorem 2.7 of [20] there is a line $\lambda \subset \mathrm{Int}(D^2) \setminus X$ such that:

- λ is properly embedded in $\mathrm{Int}(D^2)$.
- $f(\lambda) \simeq \lambda$ relative to X.
- both components C_1 and C_2 of $\mathrm{Int}(D^2) \setminus \lambda$ intersect X non-trivially.

By (2), we may assume (see the last paragraph of the proof of Theorem 2.3 on page 258 of [20]) that the ends of λ converge to single points in $\partial D^2 \setminus (\{\alpha(x_i)\} \cup \{\omega(x_i)\}$. After an isotopy relative to X we may assume that $f(\lambda) = \lambda$.

By induction on the number of orbits in $X \cap C_j$, there are disjointly embedded lines $B_i \subset \mathrm{int}(D^2)$ that contain $\mathcal{O}(x_i)$ and that are f-invariant up to isotopy relative to $X \cap C_{j(i)}$ where $\mathcal{O}(x_i) \subset C_{j(i)}$. There is no loss in assuming that $B_i \subset C_{j(i)}$. Since B_i and $f(B_i)$ are contained in $C_{j(i)}$, B_i is invariant up to isotopy relative to X as desired. □

The following lemma is the analogue of Lemma 8.7(2). It's proof is virtually the same as that of Lemma 8.7(2) and is left to the reader.

Lemma 8.10 *If f is a multiple Reeb class relative to X then no free disk contains two elements of X.*

9 Limit points on S_∞

Suppose that $f\colon M \to M$ is an orientation preserving homeomorphism of a hyperbolic surface and that f is fixed point free and isotopic to the identity. We denote the identity lift by $\tilde{f}\colon H \to H$.

The following proposition (which is essentially Proposition 5.1 of [19]) assigns two, not necessarily distinct, points in S_∞ to each $\tilde{f}$ orbit.

Proposition 9.1 *For any $\tilde{x} \in H$ the α and ω limit sets for $\tilde{f}$ are single points $\alpha(\tilde{x})$ and $\omega(\tilde{x})$ in S_∞.*

Proof We give the argument for $\omega(\tilde{x})$. The Brouwer translation theorem implies that $\omega(\tilde{x}) \subset S_\infty$. Since $\tilde{f}|_{S_\infty}$ is the identity, $\omega(\tilde{x})$ is connected. If the lemma fails then $\omega(\tilde{x})$ contains an interval $I \subset S_\infty$.

Choose a geodesic $\tilde{\gamma} \subset H$ that is the axis of a covering translation and that has endpoints in the interior of I; let U be the component of $H \setminus \tilde{\gamma}$ whose closure intersects S_∞ in a subinterval J of I.

We claim that there is a point $\tilde{y} \in H$ whose backward orbit is contained in the closure $\bar{U}$ of U. To construct $\tilde{y}$ choose positive integers $m_i < n_i$ such that:

- $\tilde{f}^{m_i}(\tilde{x}), \tilde{f}^{n_i}(\tilde{x}) \notin U$
- $\tilde{f}^j(\tilde{x}) \in U$ for $m_i < j < n_i$
- $n_i - m_i \to \infty$

Also choose a fundamental domain W for the action of the covering translation T corresponding to $\tilde{\gamma}$ and let l_i satisfy $\tilde{z}_i = T^{l_i}(\tilde{f}^{n_i-1}(\tilde{x})) \in W$. Then $\tilde{z}_i \in U \cap W$ and $\tilde{f}(\tilde{z}_i) \notin U$. It follows that $\tilde{z}_i$ lies in a compact neighborhood of the intersection of $\tilde{\gamma}$ with W and hence that the accumulation set of the $\tilde{z}_i$'s in H is non-empty. Let $\tilde{y}$ be any such accumulation point. The second and third item imply that $\mathcal{O}^-(\tilde{y}) \subset \bar{U}$.

For the rest of the argument translation arc geodesics for $\tilde{f}\colon H \to H$ are defined relative to $X = \mathcal{O}(\tilde{x}) \cup \mathcal{O}(\tilde{y})$.

There are translation arc geodesics α_i for $\tilde{f}^i(\tilde{x})$, $i > 0$, whose union is an immersed line with accumulation set equal to I. (For a proof that translation arcs can be locally constructed see, for example, Lemma 4.1 of [20]). Similarly, there are translation arc geodesics β_j for $\tilde{f}^{-j}(\tilde{y})$, $j > 0$, whose union is an immersed line with accumulation set contained in J. Since J is contained in the interior of I there are arbitrarily large i and j such that $\alpha_i \cap \beta_j \neq \emptyset$. This contradicts Theorem 8.3 and Lemma 8.7. □

Definition 9.2 If $\alpha(\tilde{x}) \neq \omega(\tilde{x})$ then define $\tilde{\gamma}(\tilde{x}) \subset H$ to be the oriented geodesic connecting $\alpha(\tilde{x})$ to $\omega(\tilde{x})$ and define $\gamma(x) \subset S$ to be its projected image. We say that $\tilde{\gamma}(\tilde{x})$ *[respectively* $\gamma(x)$*] is the geodesic associated to* $\mathcal{O}(\tilde{x})$ *[respectively* $\mathcal{O}(x)$*]*. Denote the *union of all oriented geodesics in* M *that occur as some* $\gamma(x_i)$ *by* $\Gamma(f)$.

10 Annuli

As in the previous section, $\tilde{f}\colon H \to H$ is the identity lift of a fixed point free homeomorphism $f\colon M \to M$ of a connected surface with negative Euler characteristic. We state both of our main results, Lemma 10.2 and Lemma 10.5, before turning to their proofs.

Definition 10.1 A *geodesic is birecurrent* if it is contained in the accumulation set in the unit tangent bundle of its forward end and the accumulation set in the unit tangent bundle of its backward end.

Lemma 10.2 *If* $\mathrm{Per}(f) = \emptyset$ *then:*

(1) *Each* $\gamma \in \Gamma(f)$ *is simple.*

(2) *Birecurrent elements of* $\Gamma(f)$ *have no transverse intersections.*

Definition 10.3 Suppose that U is a free disk for f that contains both x and $f^m(x)$ for some $x \in M$ and $m > 0$. Choose lifts $\tilde{x} \in \tilde{U} \subset H$. The covering translation $T\colon H \to H$ satisfying $\tilde{f}^m(\tilde{x}) \in T(\tilde{U})$ is called a *near cycle* for $\tilde{x}$ of *period* m. If the period is not relevant to the discussion then we will simply say that T is a near cycle for $\tilde{x}$.

Remark 10.4 This definition of near cycle is less general than the one in [19] but is sufficient for our purposes.

Lemma 10.5 *If* $\mathrm{Per}(f) = \emptyset$ *and* $T\colon H \to H$ *is a near cycle for* $\tilde{x}$*, then* T^+ *and* T^- *can not both lie in the same component of* $S_\infty \setminus (\{\alpha(\tilde{x}), \omega(\tilde{x})\})$.

The proofs are technical so we begin with some motivation for our approach. Consider the situation of Lemma 10.2-(1). We must show that if the geodesic γ associated to some birecurrent point $x \in M$ has self intersections then f has periodic points. We have no control over the topology of $\mathcal{O}(x)$, it may even be dense, and so can not hope to analyze the isotopy class of f relative to (the closure) of $\mathcal{O}(x)$. Lifting to covers simplifies the relative isotopy class. If we lift all the way to the universal cover $\tilde{f}\colon H \to H$, then $\tilde{f}$ is a translation class relative to $\mathcal{O}(\tilde{x})$, so we completely understand the relative isotopy class, but the geodesic $\tilde{\gamma}$ associated to $\tilde{x}$ no longer has self intersections; we have oversimplified.

The right cover to look is a cyclic one as in Lemma 5.2 with underlying space an open annulus $\mathbb{A}^\circ$. If γ has self intersections then there is a lift $\hat{f}\colon \mathbb{A}^\circ \to \mathbb{A}^\circ$ of $f\colon M \to M$ and a lift $\hat{x}$ of x such that the geodesic $\hat{\gamma}$ associated to $\mathcal{O}(\hat{x})$ is properly immersed with a finite but non-zero number of self intersections. Moreover the restriction of $\hat{f}$ to a neighborhood of the ends of $\mathcal{O}(\hat{x})$ is conjugate to the restriction of $\tilde{f}$ to a neighborhood of the ends of $\mathcal{O}(\tilde{x})$. Thus, up to relative isotopy, $\hat{f}$ acts on the ends of $\mathcal{O}(\hat{x})$ as a translation. This is made precise in Lemma 10.6.

We are now in a situation just like the one we encountered in the proof of Theorem 1.2. In both cases there is subsurface that contains all but finitely many of the punctures and on which we completely understand the relative isotopy class. In the case of Theorem 1.2, the infinitely punctured subsurface was fully invariant and we simply applied the Thurston classification theorem to the finitely punctured complement. In our present case the infinitely punctured subsurface is not fully invariant, so Thurston's theorem does not apply to the finitely punctured complement. The 'fitted family' techniques of [20] were developed for just this sort of application and it is these techniques that we employ in the proofs of Lemmas 10.2 and 10.5.

In [20] the underlying space is the interior of a disk and here the underlying space is the interior of an annulus. In addition to recalling definitions from section 5 of [20] we will use results that are proven there but not explicitly stated. These appear as Lemmas 10.7, 10.8 and 10.9 in this paper. Rather than repeat the proofs from [20] we give detailed references.

We now turn to the proof of Lemmas 10.2 and 10.5.

For $T\colon H \to H$ a non-trivial covering translation with (not necessarily distinct) endpoints $T^{\pm}$, let $\mathbb{A}^\circ$ be the quotient space of H by the action of T and let $\hat{f}\colon \mathbb{A}^\circ \to \mathbb{A}^\circ$ be the homeomorphism induced from $\tilde{f}$. We suppress the dependence of $\mathbb{A}^\circ$ and $\hat{f}$ on T to simplify notation. Denote the quotient space of $S_\infty \setminus \{T^+, T^-\}$ under the action of T by $\partial_h \mathbb{A}^\circ$. If T is hyperbolic, then $\partial_h \mathbb{A}^\circ$ compactifies $\mathbb{A}^\circ$ to a closed annulus as in Lemma 5.2. If T is parabolic then $\partial_h \mathbb{A}^\circ$ compactifies one end of $\mathbb{A}^\circ$. In either case, $\hat{f}$ extends over $\partial_h \mathbb{A}^\circ$ by the identity. Assuming that neither $\alpha(\tilde{x})$ nor $\omega(\tilde{x})$ is T^+ or T^-, denote their images in $\partial_h \mathbb{A}^\circ$ by $\alpha(\hat{x})$ and $\omega(\hat{x})$ respectively.

We work relative to a finite union $\hat{X} = \cup_{i=1}^r \mathcal{O}(\hat{x}_i) \subset \mathbb{A}^\circ$ of $\hat{f}$-orbits. Fix once and for all a complete hyperbolic structure on $\mathbb{A}^\circ \setminus \hat{X}$.

The next lemma asserts that up to isotopy relative to $\hat{X}$, $\hat{f}$ acts on the ends of the orbits of $\hat{X}$ by translation. More precisely, for each $\hat{x}_i$ there is a ray B_i^+ [respectively B_i^-] that is forward [backward] $\hat{f}$ invariant up to isotopy relative to $\hat{X}$ and that contains the forward [backward] end of $\mathcal{O}(\hat{x}_i)$. As usual, the relative isotopy class of $\hat{f}$ is reflected in the map $\hat{f}_\#$ that $\hat{f}$ induces on geodesics.

Lemma 10.6 *Assume that* $\mathrm{Per}(\hat{f}|_{\mathbb{A}^\circ}) = \emptyset$ *and that* $\hat{X} = \cup_{i=1}^r \mathcal{O}(\hat{x}_i)$ *where the* $\alpha(\hat{x}_i)$*'s are distinct and the* $\omega(\hat{x}_i)$*'s are distinct. Define translation arc geodesics with respect to* $\hat{X}$. *Then there exist translation arc geodesics* $\hat{\beta}_i^\pm$ *for elements of* $\mathcal{O}(\hat{x}_i)$ *such that:*

(1) $\hat{B}_i^+ = \cup_{n=0}^\infty \hat{f}_\#^n(\hat{\beta}_i^+)$ *is an embedded ray that converges to* $\omega(\hat{x}_i)$.

(2) $\hat{B}_i^- = \cup_{n=0}^\infty \hat{f}_\#^{-n}(\hat{\beta}_i^-)$ *is an embedded ray that converges to* $\alpha(\hat{x}_i)$.

(3) *the* $\hat{B}_i^\pm$*'s are all disjoint.*

Proof of Lemma 10.6 Choose lifts $\tilde{x}_i \in H$ of $\hat{x}_i$. The argument for (1) takes place in a disk neighborhood $U_i \subset H \cup S_\infty$ of $\omega(\tilde{x}_i)$ such that $T^k(U_i) \cap U_i = \emptyset$ for all $k \neq 0$. An isotopy relative to $\tilde{X} = \cup_{k=-\infty}^\infty \cup_{j=0}^r \mathcal{O}(T^k(\tilde{x}_j))$ with support in U_i projects to an isotopy relative to $\hat{X}$ with support in the image $\hat{U}_i \subset \mathbb{A}^\circ$ of $U_i \cap H$.

There are two cases. In the first we assume that $T^k(\alpha(\tilde{x}_j)) \neq \omega(\tilde{x}_i)$ for all j and k. In this case we may assume that $U_i \cap \tilde{X} \subset \mathcal{O}(\tilde{x}_i)$. There are translation arc geodesics $\tilde{\beta}(n, \tilde{x}_i)$ for $\tilde{f}^n(\tilde{x}_i)$ that converge to $\omega(\tilde{x}_i)$ as $n \to \infty$. For all sufficiently large n, say $n \geq N$, $\tilde{f}(\tilde{\beta}(n, \tilde{x}_i))$ and $\tilde{\beta}(n+1, \tilde{x}_i)$ are contained in U_i. Lemma 8.7-(1) implies that $\tilde{f}(\tilde{\beta}(n, \tilde{x}_i))$ and $\tilde{\beta}(n+1, \tilde{x}_i)$ are isotopic in H relative to $U_i \cap \tilde{X}$ and hence isotopic in U_i relative to $\tilde{X}$. Thus $\tilde{f}_\#(\tilde{\beta}(n, \tilde{x}_i)) = \tilde{\beta}(n+1, \tilde{x}_i)$. It follows that $\tilde{B}_i^+ = \cup_{n=0}^\infty \tilde{f}_\#^n(\tilde{\beta}(N, \tilde{x}_i) \subset U_i$ is an embedded ray that converges to $\omega(\tilde{x})$. Let $\hat{\beta}_i^+$ be the image of $\tilde{\beta}(N, \tilde{x}_i)$ and $\hat{B}_i^+$ be the image of $\tilde{B}_i^+$.

In the second case there are j and k such that $T^k(\alpha(\tilde{x}_j)) = \omega(\tilde{x}_i)$. We may assume that $U_i \cap \tilde{X} \subset \mathcal{O}^+(\tilde{x}_i) \cup \mathcal{O}^-(T^k(\tilde{x}_j))$. As in the previous case there are translation arc geodesics $\tilde{\beta}(n, \tilde{x}_i)$ for $\tilde{f}^n(\tilde{x}_i)$ that converge to $\omega(\tilde{x}_i)$ as $n \to \infty$. If $\tilde{f}(\tilde{\beta}(n, \tilde{x}_i))$ and $\beta(n+1, \tilde{x}_i)$ are contained in U_i but are not isotopic relative to $U_i \cap \tilde{X}$ then for some $m < 0$, $\beta(n+1, \tilde{x}_i) \cup \tilde{f}_\#(\beta(n, \tilde{x}_i))$ separates $\tilde{f}^m(T^k(\tilde{x}_j))$ from its $\tilde{f}$-image. In that case every translation arc geodesic for $\tilde{f}^m(T^k(\tilde{x}_j))$

intersects either $\tilde{\beta}(n+1, \tilde{x}_i)$ or $\tilde{f}_\#(\beta(n, \tilde{x}_i))$. Theorem 8.3 and Lemma 8.7 imply that this can not happen for sufficiently large n, say $n > N$. We may therefore define $\hat{\beta}_i^+$ and $\hat{B}_i^+$ as in the previous case.

This completes the proof of (1). Condition (2) follows by replacing f with f^{-1}. It is clear from the construction that (3) will be satisfied if N is sufficiently large. $\square$

Lemma 10.7 *Suppose that $\hat{f}\colon \mathbb{A}^\circ \to \mathbb{A}^\circ$, $\hat{X}$ and $\hat{B}_i^\pm$ are as in Lemma 10.6. Suppose that γ is a simple geodesic in $\mathbb{A}^\circ \setminus \hat{X}$ that is properly embedded in $\mathbb{A}^\circ$ and that intersects some $\hat{B}_i^+$ and some $\hat{B}_j^-$ infinitely often. Then $\hat{f}_\#(\gamma)$ and γ have non-trivial transverse intersection.*

Proof This follows easily from the fact that $\hat{B}_i^+$ and $\hat{B}_j^-$ translate in opposite directions. A detailed proof is given in the last paragraph of the proof of Theorem 2.3 of [20]. $\square$

We next divide $\mathbb{A}^\circ \setminus \hat{X}$ into a finitely many infinitely punctured subsurfaces on which we understand the dynamics and one finitely punctured subsurface on which we do not.

There is one infinitely punctured subsurface $V_i^\pm$ for each $\hat{B}_i^\pm$. Denote $\hat{B}_i^+ \cap \hat{X}$ by $\hat{X}_i^+$. Then V_i^+ is uniquely determined by the properties that ∂V_i^+ is a single geodesic line properly embedded in $\mathbb{A}^\circ$, that V_i^+ contains $\hat{f}_\#(\hat{B}_i^+)$ and that $V_i^+ \cap \hat{X} = \hat{f}(\hat{X}_i^+)$.

It is useful to think in terms of the following topological model. Identify $\mathbb{A}^\circ$ with $\mathbb{R}^2 \setminus \{0\}$, $\hat{B}_i^+$ with $[1, \infty) \times \{0\}$ and $\hat{X}_i^+$ with $\mathbb{N} \times \{0\}$. We may assume that all the other $\hat{B}_j^\pm$'s are far away from $\hat{B}_i^+$. Then V_i^+ can be identified with the set of points with distance at most $\frac{1}{2}$ from $[2, \infty) \times \{0\}$. Up to isotopy relative to $\hat{X}$, $\hat{f}|_{V_i^+}$ agrees with translation to the right by one unit followed by a vertical compression toward the x-axis. Thus up to isotopy relative to $\hat{X}$, $\hat{f}(V_i^+) \subset V_i^+$ and $\cap_{k=0}^\infty \hat{f}^k(V_i^+) = \emptyset$.

The V_i^-'s are defined analogously using $\hat{f}^{-1}$ and $\hat{B}_i^-$ instead of $\hat{f}$ and $\hat{B}_i^+$.

For each i and k, the line $\hat{f}_\#^k(\partial V_i^\pm)$ separates $\mathbb{A}^\circ$. Denote the complementary component that contains $\hat{f}^k(\hat{X}_i^\pm)$ by $\hat{f}_\#^k(V_i^\pm)$. Then $\hat{f}_\#(V_i^+) \subset V_i^+$ and $\cap_{k=0}^\infty \hat{f}_\#^k(V_i^+) = \emptyset$. Similarly, $\hat{f}_\#^{-1}(V_i^-) \subset V_i^-$ and $\cap_{k=0}^\infty \hat{f}_\#^{-k}(V_i^+) = \emptyset$.

The subsurface $W = \mathbb{A}^\circ \setminus (\hat{X} \cup (\bigcup_{i=1}^r \hat{V}_i^\pm))$ is finitely punctured. We write $\partial W = \partial_+ W \cup \partial_- W$ where $\partial_\pm W = \cup_{i=1}^r \partial V_i^\pm$. Then $\hat{f}_\#(\partial_+ W) \cap W = \emptyset$ and $\partial_- W \cap \hat{f}_\#(W) = \emptyset$. We say that W is the *Brouwer subsurface determined by the $B_i^\pm$'s.*

Denote $\mathbb{A}^\circ \setminus (\hat{X} \cup (\bigcup_{i=1}^r \hat{f}_\#(\hat{V}_i^\pm)))$ by $\hat{f}_\#(W)$. By part (2) of Lemma 3.5 of [20] there is a homeomorphism $\hat{f}'$, isotopic to $\hat{f}$, such that $\hat{f}'(L) = \hat{f}_\#(L)$ for each component L of ∂W. In particular $\hat{f}'(W) = \hat{f}_\#(W)$.

We recall some more definitions from [20].

Let $RH(W, \partial_+ W)$ be the set of non-trivial relative homotopy classes determined by embedded arcs $(\tau, \partial\tau) \subset (W, \partial_+ W)$. Given any collection $\mathcal{T}$ of elements of $RH(W, \partial_+ W)$ we will associate another collection $\hat{f}_\#(\mathcal{T}) \cap W$ of elements of $RH(W, \partial_+ W)$. We abuse notation slightly and sometime write $\mathcal{T} = \{t_i\}$ where each $t_i \in RH(W, \partial_+ W)$; we do not assume that the t_i's are distinct and it is essential that we allow multiplicity to occur.

Choose a homeomorphism $\hat{f}'$ as above. For any arc $\tau \subset W$ with endpoints on $\partial_+ W$, $\hat{f}'(\tau)$ is an arc in $\hat{f}_\#(W)$ with endpoints on $\hat{f}(\partial_+ W)$; in particular, $\hat{f}'(\tau) \cap \partial_- W = \emptyset$ and $\partial \hat{f}'(\tau) \cap W = \emptyset$. Let $\hat{f}_\#(\tau) \subset \hat{f}_\#(W)$ be the geodesic arc that is isotopic rel endpoints to $\hat{f}'(\tau)$. The components $\tau_1, \dots, \tau_r$ of $\hat{f}_\#(\tau) \cap W$ are arcs in W with endpoints in $\partial_+ W$. If $[\tau]$ denotes the element of $RH(W, \partial_+ W)$ determined by τ then we define $\hat{f}_\#([\tau] \cap W)$ to be $\{[\tau_1], \dots, [\tau_r]\}$. It is shown in [20] (see pages 249 - 250) that $\hat{f}_\#([\tau] \cap W)$ is well defined.

More generally if $\mathcal{T}$ is a finite collection of elements of $RH(W, \partial_+ W)$ then we define $\hat{f}_\#(\mathcal{T}) \cap W = \cup_{t \in \mathcal{T}} (\hat{f}_\#(t) \cap W)$. Note that $\hat{f}_\#(.) \cap W$ can be iterated. Inductively define $\hat{f}^n_\#(\tau) \cap W = \hat{f}^{n-1}_\#(\hat{f}_\#(\tau) \cap W) \cap W$.

We say that a finite collection $\mathcal{T} = \{t_i\}$ of distinct elements of $RH(W, \partial_+ W)$ is *fitted* if the t_i's are represented by simple disjoint arcs and if each element $s_j \in \hat{f}_\#(t_i \cap W)$ satisfies $\pm s_j \in \mathcal{T}$ where $\pm s_j$ means s_j or s_j with its orientation reversed. We say that $t \in RH(W, \partial_+ W)$ *disappears under iteration* if $\hat{f}^n_\#(t) \cap W = \emptyset$ for some $n > 0$. We say that a fitted family $\mathcal{T} = \{t_i\}$ disappears under iteration if each t_i does.

Lemma 10.8 *Assume that $\mathrm{Per}(\hat{f}|_{\mathbb{A}^\circ}) = \emptyset$ and that $\hat{X} = \cup_{i=1}^r \mathcal{O}(\hat{x}_i)$ where the $\alpha(\hat{x}_i)$'s are distinct and the $\omega(\hat{x}_i)$'s are distinct. If $\mathcal{T} \subset RH(W, \partial_+ W)$ is a fitted family that does not disappear under iteration then there exists $t \in \mathcal{T}$ and $l > 0$ such that $\hat{f}^l_\#(t) \cap W = \{t, s_1, \dots, s_m\}$ where each s_i disappears under iteration.*

Proof The first half (up through the bottom of page 252) of the proof of Theorem 5.5 of [20] applies to our current context without change to prove this lemma. □

A non-peripheral line $\hat\lambda \subset \mathbb{A}^\circ \setminus \hat X$ is a *reducing line for* $\hat f$ *relative to* $\hat X$ if it is properly embedded in $\mathbb{A}^\circ$ and is $\hat f$-invariant up to isotopy relative to $\hat X$. The isotopy class of $\hat f$ relative to $\hat X$ is *reducible* if there a reducing line for $\hat f$ relative to $\hat X$.

Lemma 10.9 *Suppose that W is a Brouwer subsurface for $\hat f\colon \mathbb{A}^\circ \to \mathbb{A}^\circ$ relative to $\hat X$ as above. If there exists $t \in RH(W, \partial_+ W)$ and $l > 0$ such that $\hat f^l_\#(t) \cap W = \{t, s_1, \dots, s_m\}$ where each s_i disappears under iteration then isotopy class of $\hat f^l$ relative to $\hat X$ is reducible and the reducing line for $\hat f^l$ relative to $\hat X$ can be chosen to be geodesic in a complete hyperbolic structure on $\mathbb{A}^\circ \setminus X$.*

Proof The last two paragraphs in the proof of Lemma 6.4 of [20] applies to our current context without change to prove this lemma. □

We now put together these results from [20] to prove our key technical lemma. Disjoint subsets of $\partial_h \mathbb{A}^\circ$ are said to *alternate around* $\partial_h \mathbb{A}^\circ$ if their intersections with each component of $\partial_h \mathbb{A}^\circ$ alternate around that component.

Proposition 10.10 *Assume that $\mathrm{Per}(\hat f|_{\mathbb{A}^\circ}) = \emptyset$ and that $\hat X = \cup_{i=1}^r \mathcal{O}(\hat x_i)$ where either $r = 1$ or the $2r$ points $\{\alpha(\hat x_i), \omega(\hat x_i) : 1 \le i \le r\}$ are distinct. Then there exists $l > 0$ such that the isotopy class of $\hat f^l$ relative to $\hat X$ is reducible. Moreover, if $r = 1$ or if the elements of $\{\alpha(\hat x_i)\}$ and $\{\omega(\hat x_i)\}$ alternate around $\partial_h \mathbb{A}^\circ$ then the reducing line $\hat\lambda$ for $\hat f^l$ relative to $\hat X$ can be chosen so that its ends converge to single points in $\partial_h \mathbb{A}^\circ$.*

Proof of Proposition 10.10 Let $B_i^\pm$ and $\beta_i^\pm$ be as in Lemma 10.6. Define W to be the Brouwer subsurface determined by the $B_i^\pm$'s.

Let $\mathcal{T}$ consist of one copy of each element of $RH(W, \partial_+ W)$ that occurs as a component of $\hat f^n_\#(\hat\beta_1^-) \cap W$ for some $n > 0$. Since the $\hat f^n_\#(\hat\beta_1^-)$'s are disjoint and simple and since W is finitely punctured, $\mathcal{T}$ is finite. It is fitted by construction.

If $\mathcal{T}$ disappears under iteration then $\hat B_1 = \cup_{n=-\infty}^{\infty} \hat f^n_\#(\hat\beta_1^-) \subset \mathbb{A}^\circ$ is a properly embedded $f_\#$-invariant line and a reducing line $\hat\lambda$ is obtained by pushing $\hat B_1$ off of itself into a non-contractible complementary component. The ends of $\hat\lambda$ converge to $\alpha(\hat x_1)$ and $\omega(\hat x_1)$ respectively. In this case $l = 1$.

Suppose now that $\mathcal{T}$ does not disappear under iteration. Lemmas 10.8 and 10.9 imply the existence of $l > 0$ and a reducing line λ for $\hat{f}^l$ relative to $\hat{X}$ that is realized as a geodesic in $\mathbb{A}^\circ \setminus \hat{X}$. Lemma 10.7 implies that each end of λ intersects at most one B_i^+ or B_j^+. There is therefore no obstruction to isotoping the end of λ relative to the union of the $B_i^\pm$ so that it converges to a single point in $\partial_h \mathbb{A}^\circ$. □

Proof of Lemma 10.2 If $\tilde{\gamma}(\tilde{x})$ projects to a geodesic $\gamma \in \Gamma(f)$ that is not simple, then there is a covering translation T such that $T(\tilde{\gamma}(\tilde{x}))$ has transverse intersections with $\tilde{\gamma}(\tilde{x})$. In particular, the pair $\{\alpha(T^n(\tilde{x})), \omega(T^n(\tilde{x}))\}$ links the pair $\{\alpha(T^{n+1}(\tilde{x})), \omega(T^{n+1}(\tilde{x}))\}$ for all n. Moreover, the endpoints $\alpha(\tilde{x})$ and $\omega(\tilde{x})$ of $\tilde{\gamma}(\tilde{x})$ are distinct from the endpoints of the axis of T so we may apply Proposition 10.10 to T and $\hat{X} = \mathcal{O}(\hat{x})$ where $\hat{x}$ is the projected image of $\tilde{x}$. The resulting line $\hat{\lambda}$ lifts to a line $\tilde{\lambda} \subset H$ whose ends converge to single points on S_∞. Up to isotopy relative to the full pre-image $\tilde{X} = \{\mathcal{O}(T^n(\tilde{x}_1)\} \subset H$ of $\hat{X}$ we may assume that $\tilde{\lambda}$ is $\tilde{f}^l$-invariant for some $l > 0$. But then $\tilde{\lambda}$ defines a non-trivial partition of $\tilde{X}$ in contradiction to the above linking property. This proves (1).

To prove (2) we assume that $\gamma_1 = \gamma(x_1)$ and $\gamma_2 = \gamma(x_2)$ are birecurrent and have transverse intersections and we argue to a contradiction. Let $\tilde{\gamma}_i = \tilde{\gamma}(\tilde{x}_i)$ be a lift of γ_i. If the endpoints of $\tilde{\gamma}_i$ differ by a covering translation, denote the image of the axis of this covering translation by $\beta_i \subset M$. Since γ_1 and γ_2 are birecurrent and have transverse intersections, there are finite subarcs $\gamma_0(x_i)$ of $\gamma(x_i)$ (where we view $\gamma(x_i)$ as a possibly periodic, infinite geodesic) such that $\gamma_0(x_1)\gamma_0(x_2)$ is an essential closed curve whose associated geodesic $\beta \subset M$ is not a multiple of either β_1 or β_2. Lift $\gamma_0(x_i)$ to $\tilde{\gamma}_0(\tilde{x}_i) \subset \tilde{\gamma}(\tilde{x}_i)$. After replacing $\tilde{\gamma}(\tilde{x}_2)$ with a translate, we may assume that the terminal endpoint of $\tilde{\gamma}_0(\tilde{x}_1)$ equals the initial endpoint of $\tilde{\gamma}_0(\tilde{x}_2)$. Let T be the covering translation that carries the initial endpoint of $\tilde{\gamma}_0(\tilde{x}_1)$ to the terminal endpoint of $\tilde{\gamma}_0(\tilde{x}_2)$.

Denote $\alpha(T^n(\tilde{x}_i))$ by α_i^n and $\omega(T^n(\tilde{x}_i))$ by ω_i^n. Then $\{\alpha_2^n, \omega_2^n\}$ links both $\{\alpha_1^n, \omega_1^n\}$ and $\{\alpha_1^{n+1}, \omega_1^{n+1}\}$ for all n. We claim that the α_i^n's and ω_i^n's are all disjoint. We will focus on ω_i^0, the other cases being analogous. Since T has no periodic points other than its two fixed points $\omega_1^0 \neq \omega_1^n$ for $n \neq 0$. Since the axis of T projects to β, our choice of β guarantees that $\omega_1^0 \neq \alpha_1^n$ for any n. The $T^j(\tilde{\gamma}(\tilde{x}_2))$'s are disjoint because $\gamma(x_2)$ is simple. By construction, $T^j(\tilde{\gamma}(\tilde{x}_2))$ intersects the component of $H \setminus \tilde{\gamma}(\tilde{x}_2)$ that is disjoint from ω_1^0 and so is entirely contained in this component for $j \neq 0$. In particular, ω_1^0 is not equal to α_2^j or ω_2^j for $j \neq 0$. We already know that ω_1^0 is not equal to α_2^0 or ω_2^0 so we have verified the claim.

Define $\mathbb{A}^\circ$ using T and let $\hat{x}_i$ be the image of $\tilde{x}_i$. By our verified claim the α and ω limit sets of $\hat{x}_1$ and $\hat{x}_2$ are four distinct points in $\partial_h \mathbb{A}^\circ$ and we may apply Proposition 10.10. The proof now concludes as in the previous case. □

Proof of Lemma 10.5 Let $\tilde{x} \in H$ be a lift of $x \in M$. Given a near cycle T of H, if either $\alpha(\tilde{x})$ or $\omega(\tilde{x})$ is $T^{\pm}$ there is nothing to prove. Assume then that $\{\alpha(\tilde{x}), \omega(\tilde{x})\} \cap \{T^{\pm}\} = \emptyset$ and let $\hat{f}\colon \mathbb{A}^\circ \to \mathbb{A}^\circ$, $\hat{X} = \mathcal{O}(\hat{x})$, l and $\hat{\lambda}$ be as in Proposition 10.10. As in the proof of Lemma 10.2, there is a lift $\tilde{\lambda}$ that separates $\mathcal{O}(\tilde{x})$ from $\mathcal{O}(T(\tilde{x}))$. After an isotopy relative to $\mathcal{O}(\tilde{x}) \cup \mathcal{O}(T(\tilde{x}))$ we may assume that $\tilde{f}^l$ preserves $\tilde{\lambda}$.

We claim that if $\{\alpha(\tilde{x}), \omega(\tilde{x})\}$ is contained in a single component of $S_\infty \setminus \{T^{\pm}\}$ then $\tilde{f}$ is Reeb class relative to $\mathcal{O}(\tilde{x}) \cup \mathcal{O}(T\tilde{x})$. By Theorem 8.3, it suffices to show that $\tilde{f}$ is not a translation class relative to $\mathcal{O}(\tilde{x}) \cup \mathcal{O}(T\tilde{x})$ and therefore it suffices to show that $\tilde{f}^l$ is not a translation class relative to $\mathcal{O}(\tilde{f}^l, \tilde{x}) \cup \mathcal{O}(\tilde{f}^l, T\tilde{x})$. This follows from the observation that if $\{\alpha(\tilde{x}), \omega(\tilde{x})\})$ is contained in a single component of $S_\infty \setminus \{T^{\pm}\}$ then $\tilde{\gamma}(\tilde{x})$ and $T\tilde{\gamma}(\tilde{x})$ are anti-parallel.

On the other hand, T is a near cycle so there is a free disk containing both $T(\tilde{x})$ and some $\tilde{f}^t(\tilde{x})$. This contradicts Lemma 8.7(2). □

11 Birecurrent points

As in the previous section, $\tilde{f}\colon H \to H$ is the identity lift of a fixed point free homeomorphism $f\colon M \to M$ of a hyperbolic surface. We now also assume that $\mathrm{Per}(f) = \emptyset$. Let $\mathcal{B} \subset M$ be the set of birecurrent points of f and let $\tilde{\mathcal{B}} \subset H$ be the full pre-image of $\mathcal{B}$.

The following definition is necessary to account for isolated punctures in M.

Definition 11.1 We say that $\mathcal{O}(x)$ *rotates about the isolated puncture* c if for some (and hence all) lifts $\tilde{x} \in H$ there is a parabolic covering translation T whose fixed point P projects to c such that every near cycle for every point in $\mathcal{O}(\tilde{x})$ is a positive iterate of T.

Lemma 11.2 *If $\tilde{x} \in \tilde{\mathcal{B}}$ and $\tilde{\alpha}(\tilde{x}) = \tilde{\omega}(\tilde{x}) = P$, then P projects to an isolated puncture c in M and $\mathcal{O}(x)$ rotates about c.*

Proof Suppose that $\tilde{\alpha}(\tilde{x}) = \tilde{\omega}(\tilde{x}) = P$. Lemma 10.5 implies that P is T^+ or T^- (or both) for an indivisible covering translation T and that all near cycles for elements of $\mathcal{O}(\tilde{x})$ are iterates of T.

We recall from [11] that a *disk chain* for $\tilde{f}$ is an ordered set $(U_0, U_1, \dots, U_n)$ of embedded free disks in the surface satisfying:

(1) If $i \neq j$ then either $U_i = U_j$ or $U_i \cap U_j = \emptyset$
(2) For $1 \leq i \leq n$, there exists $m_i > 0$ with $\tilde{f}^{m_i}(U_i) \cap U_{i+1} \neq \emptyset$.

If $U_0 = U_n$ we say that $(U_0, \dots, U_n)$ is a *periodic disk chain.* Proposition (1.3) of [11] asserts that if an orientation preserving homeomorphism of $\mathbb{R}^2$ possesses a periodic disk chain then it has a fixed point. Thus $\tilde{f}$ has no periodic disk chains.

For any disk chain $D = (U_0, \dots, U_n)$, we denote the disk chain $(TU_0, \dots, TU_n)$ by TD. If $D_1 = (U_0, \dots, U_n)$ and $D_2 = (U_n, \dots, U_m)$ where all the U_i's are disjoint or equal, then we denote the disk chain $(U_0, \dots, U_m)$ by $D_1 \cdot D_2$.

Suppose that both positive and negative iterates of T occur as near cycles for elements of $\mathcal{O}(\tilde{x})$. Then there exist free disks $\tilde{U}_1, \tilde{U}_2$, elements $\tilde{x}_i \in \tilde{U}_i \cap \mathcal{O}(\tilde{x})$ and integers $p > 0$ and $q < 0$ such that $D_1 = (\tilde{U}_1, T^p(\tilde{U}_1))$ and $D_2 = (\tilde{U}_2, T^q(\tilde{U}_2))$ are disk chains. Since all near cycles for elements of $\mathcal{O}(\tilde{x})$ are iterates of T, every translate of $\tilde{U}_i$ that intersect $\mathcal{O}(\tilde{x})$ equals $T^a(\tilde{U}_i)$ for some integer a. If it is possible to choose $\tilde{U}_1$ and $\tilde{U}_2$ so that they have the same projection in M, do so. If this is not possible, then birecurrence allows us to replace $\tilde{U}_i$ with an arbitrarily small disk neighborhood of $\tilde{x}_i$. We may therefore assume without loss that the projections $U_i \subset M$ of $\tilde{U}_i$ are either disjoint or equal. Birecurrence implies that there exist integers r and s such that the forward orbit of $\tilde{x}_1$ intersects $T^r\tilde{U}_2$ and the forward orbit of $\tilde{x}_2$ intersects $T^s\tilde{U}_1$. In particular, $D_3 = (\tilde{U}_1, T^r\tilde{U}_2)$ and $D_4 = (\tilde{U}_2, T^s\tilde{U}_1)$ are disk chains. For $i > -\frac{r+s}{q}$ the disk chain $D_5 = (D_3, \cdot T^r D_2 \cdot T^{r+q} D_2 \cdots T^{r+(i-1)q} D_2 \cdot T^{r+iq} D_4)$ begins at $\tilde{U}_1$ and ends at $T^a\tilde{U}_1$ with $a < 0$. But then D_1 and D_5 can be used to produce a disk chain that begins and ends at $\tilde{U}_1$ in contradiction to the fact that there are no periodic disk chains for $\tilde{f}$. This contradiction proves that all near cycles for $\mathcal{O}(\tilde{x})$ are positive iterates of T or all are negative iterates of T. After replacing T by its inverse if necessary, we may assume that all near cycles are positive iterates of T.

If T is parabolic, then we're done so suppose that T is hyperbolic. Thus $T^+ = P$ and $T^- \neq P$. Let us now switch our focus to $\tilde{f}^{-1}$. All of its near cycles for $\tilde{x}$ are negative iterates of T. But this contradicts the fact that $\tilde{x}$ converges to $P \neq T^-$ under iteration by $\tilde{f}^{-1}$. □

We will need the following facts about near cycles.

Lemma 11.3 *Suppose that T is a near cycle for $\tilde{x}$ of period m.*

(1) *There is a neighborhood $\tilde{U}$ of $\tilde{x}$ such that T is a near cycle for $\tilde{y}$ of period m for each $\tilde{y} \in \tilde{U}$.*

(2) *For any covering translation S, STS^{-1} is a near cycle for $S\tilde{x}$ of period m.*

(3) *There is a neighborhood $\tilde{U}$ of $\tilde{x}$ so that if $\tilde{f}^s(x) \in S(\tilde{U})$ for some covering translation S, then STS^{-1} is a near cycle for $\tilde{f}^s(\tilde{x})$ of period m.*

Proof Items 1 and (2) follow from the definition of near cycle. If $\tilde{U}$ is as in (1) and $\tilde{f}^s(x) \in S(\tilde{U})$ then T is a near cycle of period m for $\tilde{f}^s(S^{-1}x)$. Item (3) now follows from (2). □

Lemma 11.4 *If $\tilde{x} \in \tilde{\mathcal{B}}$ and $\alpha(\tilde{x}) \neq \omega(\tilde{x})$ then one of the following holds.*

(1) *Every near cycle for every point in $\mathcal{O}(\tilde{x})$ has axis $\tilde{\gamma}(\tilde{x})$.*

(2) *For every neighborhood V of $\tilde{\alpha}(\tilde{x})$ or $\tilde{\omega}(\tilde{x})$ there is a near cycle for an element of $\mathcal{O}(\tilde{x})$ whose fixed points on S_∞ are contained in V.*

Proof We will prove either that (1) holds or that (2) holds for neighborhoods of $\omega(\tilde{x})$. Applying this with f^{-1} replacing f completes the proof.

If $\omega(\tilde{x})$ is the fixed point of a parabolic near cycle for an element of $\mathcal{O}(\tilde{x})$ then (2) holds for neighborhoods of $\omega(\tilde{x})$ and we are done. Suppose then that $\omega(\tilde{x})$ is not the fixed point of a parabolic near cycle for an element of $\mathcal{O}(\tilde{x})$. If there is a parabolic near cycle for an element of $\mathcal{O}(\tilde{x})$, then Lemma 10.5 implies that its fixed point must be $\alpha(\tilde{x})$.

For $t > 0$, define $N_t = \{k \in \mathbb{Z} :$ there is a near cycle for $\tilde{f}^k(\tilde{x})$ with period at most $t\}$. Fix $t > 0$ such that $N_t \neq \emptyset$. Lemma 11.3 and birecurrence implies that N_t is bi-infinite. For each $k \in N_T$ choose a near cycle T_k for $\tilde{f}^k(\tilde{x})$with period at most t. Denote the axis of T_k by A_k.

There is a uniform bound to the distance that T_k moves $\tilde{f}^k(\tilde{x})$. This has several consequences. First, $\lim_{k\to\infty} T_k(\omega(\tilde{x})) = \omega(\tilde{x})$. This implies that T_k is hyperbolic for all large k and hence by Lemma 11.3, hyperbolic for all k. Additionally, both the translation length of T_k and the distance between $\tilde{f}^k(\tilde{x}_k)$ and A_k are uniformly bounded. The set of axes of covering translations with uniformly bounded translation length is discrete and closed in H. Thus either there is sequence $k_i \in N_t$ tending to infinity such that $A_{k_i} \to \omega(\tilde{x})$ or there is a covering translation T whose axis A has $\omega(\tilde{x})$ as an endpoint and satisfies

$A_k = A$ for all sufficiently large $k \in N_t$. If the former case holds we are done, so assume that $A_k = A$ for all sufficiently large k, say $k > K$.

Choose any $k_1 < k_2 \in N_t$. For any $\epsilon > 0$, there exists $k > K$ and a covering translation S such that $\tilde{f}^{k+j}(\tilde{x})$ is ϵ close to $\tilde{f}^{k_1+j}(S\tilde{x})$ for $0 \le j \le k_2 - k_1 + t$. In other words, the initial $k_2 - k_1 + t$ points in the orbits of $\tilde{f}^k(\tilde{x})$ and of $\tilde{f}^{k_1}(S\tilde{x})$ are within ϵ of each other. For sufficiently small ϵ, Lemma 11.3 implies that $k, k + (k_2 - k_1) \in N_t$, that $SA_k = A_{k_1}$ and that $SA_{k+(k_2-k_1)} = SA_{k_2}$. Since $A_k = A_{k+(k_2-k_1)}$, $A_{k_1} = A_{k_2}$.

We have now shown that $A_k = A$ is independent of k. The same argument that established $\lim_{k\to\infty} T_k(\omega(\tilde{x})) = \omega(\tilde{x})$ also proves that $\lim_{k\to-\infty} T_k(\alpha(\tilde{x})) = \alpha(\tilde{x})$. Thus $\alpha(\tilde{x})$ is an endpoint of A. Since the choice of T_k was arbitrary we have verified (1) for fixed t. Since t was arbitrary, we have verified (1). □

Definition 11.5 A simple closed geodesic is a *partial cross-section* to $\Gamma(f)$ if it is only crossed in one direction by elements of $\Gamma(f)$ and if there is at least one such crossing.

Lemma 11.6 **(1)** *The map that assigns the unordered pair $\{\tilde{\alpha}(\tilde{x}), \tilde{\omega}(\tilde{x})\}$ to $\tilde{x} \in \tilde{\mathcal{B}}$ is continuous.*

(2) *If $x \in \mathcal{B}$ and $\gamma(x)$ is defined then $\gamma(x)$ is birecurrent.*

(3) *If $x \in \mathcal{B}$ and $\gamma(x)$ is defined but is not a simple closed curve, then $\gamma(x)$ intersects a partial cross-section α. In this case, the set $\{y \in \mathcal{B} : \gamma(y)$ crosses $\alpha\}$ is open in $\mathcal{B}$.*

Proof We first show that if $\gamma(x)$ is defined, then continuity at $\tilde{x}$ implies that $\gamma(x)$ satisfies (2) and (3).

Choose $x_i = f^{n_i}(x)$ with $n_i > 0$ such that $x_i \to x$; choose also lifts $\tilde{x}_i \to \tilde{x}$. The geodesics $\tilde{\gamma}(\tilde{x}_i)$ are translates of $\tilde{\gamma}(\tilde{x})$. If $\tilde{\gamma}(\tilde{x})$ is the axis of a covering translation, then $\gamma(x)$ is a simple closed curve and we are done. We may therefore assume that the $\tilde{\gamma}(\tilde{x}_i))$'s are distinct as unoriented lines. We may also assume there is a free disk that contains x and each x_i.

The continuity assumption implies that $\tilde{\gamma}(\tilde{x}_i) \to \tilde{\gamma}(\tilde{x})$ as unoriented lines. From Lemma 10.2(1) it follows that the $\tilde{\gamma}(\tilde{x}_i)$'s are disjoint. Lemmas 8.9 and 8.10 therefore imply that for sufficiently large i, say $i > N$, the $\tilde{\gamma}(\tilde{x}_i)$'s are consistently oriented (i.e. the lines are parallel and not anti-parallel). Let T_i be the covering translation that carries $\tilde{f}^{n_i}(\tilde{x})$ to $\tilde{x}_i$ and hence also $\tilde{\gamma}(\tilde{x})$ to $\tilde{\gamma}(\tilde{x}_i)$. Choose $\tilde{y} \in \gamma(\tilde{x})$. Since $\tilde{f}^{n_i}(\tilde{x}) \to \omega(\tilde{x})$ and $\tilde{\gamma}(\tilde{x}_i) \to \tilde{\gamma}(\tilde{x})$ there are points

$\tilde{y}_i \in \tilde{\gamma}(\tilde{x})$ such that $\tilde{y}_i \to \omega(\tilde{x})$ and $T_i(\tilde{y}_i) \to \tilde{y}$. This proves that the accumulation set of the forward end of $\gamma(x)$ contains $\tilde{y}$ and hence all of $\gamma(x)$ since $\tilde{y}$ was arbitrary. By symmetry, $\gamma(x)$ is also contained in the accumulation set of its backward end so $\gamma(x)$ is birecurrent and we have verified (2).

Lemma 10.2(2) now implies that $\gamma(x)$ does not intersect any other $\gamma(y)$. Lemmas 8.9 and 8.10 therefore imply that if $\tilde{\gamma}(\tilde{y})$ is contained in the region bounded by $\tilde{\gamma}(\tilde{x}_i)$ and $\tilde{\gamma}(\tilde{x}_j)$ for $i,j > N$, then the orientation on $\tilde{\gamma}(\tilde{y})$ agrees with the orientation $\tilde{\gamma}(\tilde{x}_i)$ and $\tilde{\gamma}(\tilde{x}_j)$. Let $\tilde{\alpha}_0$ be a geodesic arc connecting $\tilde{\gamma}(\tilde{x}_i)$ and $\tilde{\gamma}(\tilde{x}_j)$ and let α_0 be its projected image. Then elements of $\Gamma(f)$ cross α_0 only in one direction. Extend α_0 to a simple closed curve α' by following $\gamma(x)$ from one intersection with α_0 to the next. The geodesic α corresponding to α' is a partial cross-section. This establishes the first part of (3). The second part follows from the fact that a pair of geodesics in H have non-empty transverse intersection if and only if their endpoints link in S_∞ and the fact that linking a particular pair of points in S_∞ is an open condition on pairs of points in S_∞.

It remains to establish (1). Suppose that $\tilde{x}_i, \tilde{x} \in \tilde{\mathcal{B}}$ and that $\tilde{x}_i \to \tilde{x}$. We assume at first that $\alpha(\tilde{x}) \neq \omega(\tilde{x})$. According to Lemma 11.4 there are two cases to consider. In the first, every neighborhood U^+ of $\omega(\tilde{x})$ contains T_1^+ and T_1^- where T_1 is a near cycle for an element of $\mathcal{O}(\tilde{x})$ and every neighborhood U^- of $\alpha(\tilde{x})$ contains T_2^+ and T_2^- where T_2 is a near cycle for an element of $\mathcal{O}(\tilde{x})$. For sufficiently large i both T_1 and T_2 are near cycles for elements of $\mathcal{O}(\tilde{x}_i)$. Corollary 10.5 implies that $\{\alpha(\tilde{x}_i), \omega(\tilde{x}_i)\}$ intersects both U^+ and U^- for all sufficiently large i. This completes the proof of (1) in the first case and so also the proof of (2) and (3) in the first case.

In the second case, Lemma 11.4 implies that $\tilde{\gamma}(\tilde{x})$ is the axis of a near cycle for $\tilde{x}$ so we have proved (2) and (3). As in the first case, for all sufficiently large i, say $i > N$, $\tilde{\gamma}(\tilde{x})$ is the axis of a near cycle for an element of $\mathcal{O}(\tilde{x}_i)$. Lemma 10.2 and (2) imply that $\tilde{\gamma}(\tilde{x}_i)$ and $\tilde{\gamma}(\tilde{x})$ are either disjoint or equal for $i > N$. Lemma 10.5 therefore implies that they are equal.

It remains to establish (1) when $\tilde{\alpha}(\tilde{x}) = \tilde{\omega}(\tilde{x})$. By Lemma 11.2, $\tilde{\alpha}(\tilde{x}) = \tilde{\omega}(\tilde{x}) = T^\pm$ for a parabolic covering translation and every near cycle for $\mathcal{O}(\tilde{x})$ is an iterate of T. For sufficiently large i, iterates of T are near cycles for elements of $\mathcal{O}(\tilde{x}_i)$. By Lemma 10.5 either $\alpha(\tilde{x}_i)$ or $\omega(\tilde{x}_i)$ is $T^\pm$. This rules out the possibility that $\gamma(x_i)$ is defined and birecurrent. The only other possibility is that $\tilde{\alpha}(\tilde{x}_i) = \tilde{\omega}(\tilde{x}_i) = T^\pm$. □

12 Proof of Theorem 1.1

The following result is implicit in the paper of Atkinson [1]. Since it is not explicit there we give the (short) proof here. This proof is essentially the same as an argument in [1].

Proposition 12.1 *Suppose $T\colon X \to X$ is an ergodic automorphism of a probability space (X,ν) and let $\phi\colon X \to \mathbb{R}$ be an integrable function with $\int \phi \; d\nu = 0$. Let $S(n,x) = \sum_{i=0}^{n-1} \phi(T^i(x))$. Then for any $\varepsilon > 0$ the set of x such that $|S(n,x)| < \varepsilon$ for infinitely many n is a full measure subset of X.*

Proof Let A denote the set of x such that $|S(n,x)| < \varepsilon$ for only finitely many n. We will show the assumption $\mu(A) > 0$ leads to a contradiction. Suppose $\mu(A) > 0$ and let A_m denote the subset of A such that $|S(i,x)| < \varepsilon$ for m or fewer values of i. Then $A = \cup A_m$ and there is an $N > 0$ such that $\mu(A_N) > p$ for some $p > 0$.

The ergodic theorem applied to the characteristic function of A_N implies that for almost all x and all sufficiently large n (depending on x) we have

$$\frac{card(A_N \cap \{T^i(x) \mid 0 \leq i < n\})}{n} > p.$$

We now fix an $x \in A_N$ with this property. Let $B_n = \{i \mid 0 \leq i \leq n \text{ and } T^i(x) \in A_N\}$ and $r = card(B_n)$; then $r > np$. Any interval in $\mathbb{R}$ of length ε which contains $S(i,x)$ for some $i \in B_n$ contains at most N values of $\{S(j,x) : j > i\}$. Hence any interval of length ε contains at most N elements of $\{S(i,x) \mid i \in B_n\}$. Consequently an interval containing the r numbers $\{S(i,x) \mid i \in B_n\}$ must have length at least $r\varepsilon/N$. Since $r > np$ this length is $> np\varepsilon/N$. Therefore

$$\sup_{0 \leq i \leq n} |S(i,x)| > \frac{np\varepsilon}{2N},$$

and hence by the ergodic theorem, for almost all $x \in A_N$

$$\left| \int \phi \; d\mu \right| = \lim_{n\to\infty} \frac{|S(n,x)|}{n} = \limsup_{n\to\infty} \frac{|S(n,x)|}{n} > \frac{p\varepsilon}{2N} > 0.$$

This contradicts the hypothesis so our result is proven. □

Corollary 12.2 *Suppose $T\colon X \to X$ is an automorphism of a Borel probability space (X,μ) and $\phi\colon X \to \mathbb{R}$ is an integrable function. Let $S(n,x) = \sum_{i=0}^{n-1} \phi(T^i(x))$ and suppose $\mu(P) > 0$ where $P = \{x \mid \lim_{n\to\infty} S(n,x) = \infty\}$. Let*

$$\hat{\phi}(x) = \lim_{n\to\infty} \frac{S(n,x)}{n}.$$

Then $\int_P \hat{\phi}\, d\mu > 0$. *In particular* $\hat{\phi}(x) > 0$ *for a set of positive* μ*-measure.*

Proof By the ergodic decomposition theorem there is a measure m on the space $\mathcal{M}$ of all T invariant ergodic Borel measures on X with the property that for any μ integrable function $\psi\colon X \to \mathbb{R}$ we have $\int \psi\, d\mu = \int_{\mathcal{M}} I(\psi,\nu)\, dm$ where $\nu \in \mathcal{M}$ and $I(\psi,\nu) = \int \psi\, d\nu$.

The set P is T invariant. Replacing $\phi(x)$ with $\phi(x)\mathfrak{X}_P(x)$, where $\mathfrak{X}_P$ is the characteristic function of P, we may assume that ϕ vanishes outside P. Then clearly $\hat{\phi}(x) \geq 0$ for all x for which it exists. Let $\mathcal{M}_P$ denote $\{\nu \in \mathcal{M} \mid \nu(P) > 0\}$. If $\nu \in \mathcal{M}_P$ the fact that $\hat{\phi}(x) \geq 0$ and the ergodic theorem imply that $I(\phi,\nu) = \int \phi\, d\nu = \int \hat{\phi}\, d\nu \geq 0$. Also Proposition 12.1 implies that $\int \phi\, d\nu = 0$ is impossible so $I(\phi,\nu) > 0$. Then $\mu(P) = \int I(\mathfrak{X}_P,\nu)\, dm = \int \nu(P)\, dm = \int_{\mathcal{M}_P} \nu(P)\, dm$. This implies $m(\mathcal{M}_P) > 0$ since $\mu(P) > 0$.

Hence
$$\int \hat{\phi}\, d\mu = \int \phi\, d\mu = \int I(\phi,\nu)\, dm \geq \int_{\mathcal{M}_P} I(\phi,\nu)\, dm > 0$$
since $I(\phi,\nu) > 0$ for $\nu \in \mathcal{M}_P$ and $m(\mathcal{M}_P) > 0$. □

We are now prepared to complete the proof of the following result from §1.

Theorem 1.1 *Suppose* $F\colon S \to S$ *is a non-trivial, Hamiltonian surface diffeomorphism and that if* $S = S^2$ *then* $\mathrm{Fix}(F)$ *contains at least three points. Then* F *has periodic points of arbitrarily high period.*

Proof We first observe that F has infinite order. If $R(F) \neq \emptyset$ then this follows from the fact that the isotopy class of F relative to $\mathrm{Fix}(F)$ has infinite order. If $R(F) = \emptyset$ this follows from Lemma 6.3.

Suppose that the periods of points in $\mathrm{Per}(F)$ are bounded. We will show that this leads to a contradiction. Replacing F with an iterate we may assume $\mathrm{Fix}(F) = \mathrm{Per}(F) \neq \emptyset$. By [4] and Proposition 7.1, $M = S \setminus \mathrm{Fix}(F)$ is F invariant and $f = F|_M\colon M \to M$ is isotopic to the identity.

No component of M can be a disk by Theorem 3.1. We first handle the case that some component U is an annulus. In this case by Lemma 5.1 U can be compactified to a closed annulus $\bar{U}$ in such a way that f extends to the identity on at least one boundary component (this is where the assumption that $\mathrm{Fix}(f)$ has at least three points if $S = S^2$ is used). If the lift of $f|_{\bar{U}}$ which fixes points on the boundary were to have mean rotation number zero, $f|_{\bar{U}}$ would have an interior fixed point by Theorem 2.1 of [15], which is impossible. Hence this lift has non-zero mean rotation number and consequently a point of non-zero rotation number. Also $f|_{\bar{U}}$ is chain recurrent since it preserves a finite measure positive on open sets. It now follows from Theorem 3.2 that f has points of arbitrarily high period.

We are left with the case that all components of M have a hyperbolic structure. If M is not connected we replace it by one component. Let $\tilde{f}\colon H\cup S_\infty \to H\cup S_\infty$ be the identity lift of f.

We continue with the notation of the previous sections. Recall that by the Poincaré recurrence theorem the set of birecurrent points of M has full measure. Since there are only countably many simple closed geodesics and isolated punctures in M, Lemma 11.2 and Lemma 11.6 imply that there is a set $P \subset S$ of positive measure such that one of the following conditions is satisfied by all $x \in P$:

(1) $\gamma(x) = \gamma$ for some simple closed geodesic γ

(2) $\mathcal{O}(x)$ rotates about c for some isolated puncture c

(3) $\gamma(x)$ crosses α for some partial cross section α.

We consider these possibilities in order.

First assume $\gamma(x) = \gamma$ is a simple closed geodesic. Let $\tilde{\gamma}$ be a lift of γ and let $\tilde{P}(\tilde{\gamma}) = \{\tilde{x} \in H : \tilde{x}$ projects into P and $\tilde{\gamma}(\tilde{x}) = \tilde{\gamma}\}$. Recall that the canonical cyclic cover $\hat{f}\colon \mathbb{A} \to \mathbb{A}$ associated to γ is induced from $\tilde{f}$ by taking the quotient space of $H \cup S_\infty$ under the action of the indivisible covering translation T with axis equal to $\tilde{\gamma}$. By Lemma 5.4 it suffices to show that $\tilde{f}$ considered as a lift of $\hat{f}$ has points of two different rotation numbers. Denote the image of $\tilde{P}(\tilde{\gamma})$ in $\mathbb{A}$ by $\hat{P}$. If T_0 is another covering translation, not a power of T, then $\tilde{\gamma}(T_0(\tilde{x})) = T_0(\tilde{\gamma}) \neq \tilde{\gamma}$ so the covering map $\pi\colon \operatorname{Int}(\mathbb{A}) \to M$ restricts to a bijection between $\hat{P}$ and P. In particular there is a finite measure μ on $\hat{P}$ defined by $\mu(A) = \omega(\pi(A))$. The identity lift of f is a lift of $\hat{f}$. We can now apply Corollary 12.2 to $\hat{f}\colon \mathbb{A} \to \mathbb{A}$ with measure μ and function $\phi = \delta$, a homological displacement function, so that $\hat{\phi}(x) = \rho(x, \tilde{f})$. It follows that there is a point of $\mathbb{A}$ with positive rotation number. Since any point of $\partial\mathbb{A}$ has zero rotation number for $\hat{f}$, this completes the proof in the first case.

Next we consider the case that $\mathcal{O}(x)$ rotates about c for some isolated puncture c. Recall from Definition 11.1 that this means if $x \in P$ then for a lift $\tilde{x} \in H$ there is a parabolic covering translation T whose fixed point p corresponds to c such that every near cycle for every point in $\mathcal{O}(\tilde{x})$ is a positive iterate of T.

Let $\tilde{P}(p) = \{\tilde{x} \in H : \tilde{x}$ projects into P and $\tilde{\omega}(\tilde{x}) = \tilde{\alpha}(\tilde{x}) = p\}$. Let $\hat{f}\colon \mathbb{A} \to \mathbb{A}$ be the canonical cyclic cover associated to γ and recall that $\hat{f}|_{\mathrm{int}\mathbb{A}}$ is induced from $\tilde{f}|_H$ by taking the quotient space under the action of T. As in the previous case the image $\hat{P} \subset \mathbb{A}$ of $\tilde{P}(p)$ has a finite invariant measure defined by $\mu(A) = \omega(\pi(A))$. By Lemma 5.2 and Lemma 5.4 it suffices to show that $\tilde{f}$ considered as a lift of $\hat{f}$ has points of positive rotation number.

Let $\hat{U} \subset \mathrm{int}\mathbb{A}$ be a lift of an open free disk $U \subset M$ for f such that $\mu(\hat{U}_P) > 0$ where $\hat{U}_P = \hat{U} \cap \hat{P}$. A full measure set of $U_P = U \cap P$ is recurrent. But the fact that $\mathcal{O}(x)$ rotates about c for $x \in P$ means that if $x \in U_P \cap f^k(U_P)$, for $k > 0$, and $\hat{x} \in \hat{U}_P \cap \pi^{-1}(x)$ then $\hat{x} \in \hat{U}_P \cap \hat{f}^k(\hat{U}_P)$. It follows that if $\hat{V} = \cup_{n \ge 0} \hat{f}^n(\hat{U}_P)$ and $V = \cup_{n \ge 0} f^n(U_P)$ then $\pi|_{\hat{V}}\colon \hat{V} \to V$ is a measure preserving bijection on a set of full measure. In particular $\mu(\hat{V})$ is finite.

Let $\delta\colon \mathbb{A} \to \mathbb{R}$ be a homological displacement function for the lift $\hat{f}$ and let $\delta_U\colon \hat{U}_P \to \mathbb{R}$ be defined by

$$\delta_U(x) = \sum_{m=0}^{n-1} \delta(\hat{f}^m(x)).$$

where $x \in \hat{U}_P$, $\hat{f}^n(x) \in \hat{U}_P$, and $\hat{f}^k(x) \notin \hat{U}_P$ for $1 \le k < n$. Note in this case that if $\tilde{U}_P$ is a lift of $\hat{U}_P$ and $\tilde{x} \in \tilde{U}_P$ is a lift of x, then $\tilde{f}^n(\tilde{x}) \in T^k(\tilde{U}_P)$ for some $k > 0$. Hence if we identify $H_1(\mathbb{A}, R)$ with $\mathbb{R}$ we can consider δ_U as a real valued function which takes on only strictly positive integer values. In effect δ_U is a homological displacement function for the first return map of $\hat{f}$ on U_P.

We claim that $\int_{\hat{V}} \delta \ d\mu = \int_{\hat{U}_p} \delta_U \ d\mu$. This can be seen by setting

$$W_n = \{x \in \hat{U}_P | \hat{f}^n(x) \in \hat{U}_P \text{ and } \hat{f}^k(x) \notin \hat{U}_P \text{ for } 0 < k < n\}$$

and letting $V_n = \cup_{m=0}^{n-1} \hat{f}^m(W_n)$. Then

$$\int_{W_n} \delta_U \ d\mu = \int_{W_n} \sum_{m=0}^{n-1} \delta(\hat{f}^m(x)) \ d\mu = \sum_{m=0}^{n-1} \int_{\hat{f}^m(W_n)} \delta \ d\mu = \int_{V_n} \delta \ d\mu.$$

Since the W_n's are pairwise disjoint as are the V_n's and since $\hat{U}_P = \cup W_n$ and $\hat{V} = \cup V_n$ the claim $\int_{\hat{V}} \delta \ d\mu = \int_{\hat{U}_p} \delta_U \ d\mu$ follows.

Hence

$$\int_{\hat{V}} \rho(x, \tilde{f}) \ d\mu = \int_{\hat{V}} \delta \ d\mu = \int_{\hat{U}_p} \delta_U \ d\mu > 0,$$

and it follows that $\rho(x, \tilde{f}) > 0$ on a set of positive measure. Since $\hat{f}$ is covered by the identity lift it has fixed points in the boundary of $\mathbb{A}$ and hence it has points of two distinct rotation numbers. By Lemma 5.4 $\hat{f}$ has points of arbitrarily high period.

We are left with the case that $\gamma(x)$ crosses the partial cross section α, which we will show contradicts the hypothesis that the flux homomorphism is zero. We assume that all crossings of α by elements of $\Gamma(f)$ are in the positive direction. Let F_t be an isotopy of S with $F_0 = id$ and $F_1 = F$. Let $\delta\colon S \to H_1(S,R)$ be a homological displacement function. Then

$$\rho_\mu(F_t) \wedge [\alpha] = \int \rho(x, F_t) \wedge [\alpha]\ d\mu = \int \delta(x) \wedge [\alpha]\ d\mu.$$

Also since there is no x for which $\gamma(x)$ crosses α in the negative sense we have $\rho(x, F_t) \wedge [\alpha] \geq 0$ for all x in a full measure subset of S and hence

$$\rho_\mu(F_t) \wedge [\alpha] = \int \rho(x, F_t) \wedge [\alpha]\ d\mu \geq \int_P \rho(x, F_t) \wedge [\alpha]\ d\mu.$$

We want now to apply Corollary 12.2 with $\phi(x) = \delta(x) \wedge [\alpha]$, which will mean that $\hat{\phi}(x) = \rho(x, F_t) \wedge [\alpha]$. If $S(n,x) = \sum_{i=0}^{n-1} \delta(F^i(x)) \wedge [\alpha]$ the fact that $\gamma(x) = \gamma$ for all $x \in P$ implies that $\lim\limits_{n\to\infty} S(n,x)) = \infty$ for all $x \in P$. Thus by Corollary 12.2 we conclude

$$\rho_\mu(F_t) \wedge [\alpha] \geq \int_P \rho(x, F_t) \wedge [\alpha]\ d\mu = \int_P \hat{\phi}\ d\mu > 0.$$

Therefore by Proposition 2.11 $\mathcal{F}(F_t)([\alpha]) \neq 0$ contradicting the hypothesis that F is Hamiltonian. □

Proof of Theorem 1.3 Replacing F by a power we may assume that F has more than $2g+2$ fixed points where g is the genus of S. Let Q be a finite set consisting of more than $2g+2$ fixed points. By Theorem 1.1, we can choose a periodic point x whose period is $p > 1$. Letting $P = Q \cup orb(x)$ we observe that the Thurston canonical form of F relative to P cannot be periodic. This is because any non-trivial periodic diffeomorphism of S can have at most $2g+2$ fixed points (see [8]).

Hence some iterate of the Thurston canonical form of F relative to P must be reducible with either a pseudo-Anosov component or a non-trivial Dehn twist along a reducing curve. In either of these cases F^p is not isotopic to the identity relative to P. □

Proof of Theorem 1.4 This is an immediate consequence of Theorem 1.3 and Proposition 7.1.

Acknowledgements

John Franks is supported in part by NSF grant DMS0099640.
Michael Handel is supported in part by NSF grant DMS0103435.

References

[1] **G Atkinson**, *Recurrence of cocycles and random walks*, J. Lond. Math. Soc. 13 (1976) 486-488

[2] **A Banyaga**, *The Structure of Classical Diffeomorphism Groups*, Mathematics and its Applications Kluwer Academic Publishers, (1997)

[3] **S Bleiler**, **A Casson**, *Automorphisms of surfaces after Nielsen and Thurston*, London Mathematical Society Student Texts, 9, Cambridge University Press (1988)

[4] **M Brown**, **J Kister**, *Invariance of complementary domains of a fixed point set*, Proc. Amer. Math. Soc. 91 (1984) 503–504

[5] **C Conley**, **E Zehnder**, *The Birkhoff-Lewis Fixed point Theorem and a Conjecture of V.I. Arnold*, Invent. Math. 73 (1983) 33–49

[6] **C J Earle**, **J Eells**, *The diffeomorphism group of a compact Riemann surface*, Bull. Amer. Math. Soc. 73 (1967) 557–559

[7] **A Fathi**, **F Laudenbach**, **V Poenaru**, *Travaux de Thurston sur les surfaces*, Asterisque 66–67 (1979)

[8] **S.D Fisher**, **J Franks**, *The Fixed Points of an Analytic Self Mapping*, Proc. Amer. Math. Soc. 99 (1987) 76-78

[9] **J Franks**, *Homology and Dynamical Systems*, Amer. Math. Soc. CBMS Regional Conference Series 49 (1982), 120 pages.

[10] **J Franks**, *Recurrence and fixed points of surface homeomorphisms*, Ergod. Th. Dynam. Sys. 8* (1988) 99–107

[11] **J Franks**, *Generalizations of the Poincaré-Birkhoff Theorem*, Annals of Math. 128 (1988) 139–151

[12] **J Franks**, *Geodesics on S^2 and periodic points of annulus homeomorphisms*, Inventiones Math. 108 (1992) 403–418

[13] **J Franks**, *Area Preserving Homeomorphisms of Open Surfaces of Genus Zero*, New York Jour. of Math. 2 (1996) 1–19

[14] **J Franks**, *Rotation vectors and fixed points of area preserving surface diffeomorphisms*, Trans. Amer. Math. Soc. 348 (1996) 2637–2662

[15] **J Franks**, *Rotation numbers for Area Preserving Homeomorphisms of the Open Annulus*, Proceedings of the International Conference Dynamical Systems and Related Topics, (K Shiraiwa, editor) World Scientific (1991) 123–128

[16] **J Franks**, **M Handel**, *Area preserving group actions on surfaces*, Geometry and Topology 7 (2003) 757–771

[17] **Mary-Elizabeth Hamstrom**, *The space of homeomorphisms on a torus*, Illinois J. Math. 9 (1965) 59–65

[18] **Mary-Elizabeth Hamstrom**, *Homotopy groups of the space of homeomorphisms on a 2-manifold*, Illinois J. Math. 10 (1966) 563–573

[19] **M Handel**, *Zero entropy surface diffeomorphisms*, preprint

[20] **M Handel**, *A fixed point theorem for planar homeomorphisms*, Topology 38 (1999) 235–264

[21] **M Handel**, *Commuting homeomorphisms of* S^2, Topology 31 (1992) 293–303

[22] **M Handel**, **W P Thurston**, *New proofs of some results of Nielsen*, Advances in Math. 56 (1985) 173–191

[23] **A Katok**, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Inst. Hautes Etudes Sci. Publ. Math. 51 (1980) 137 –173

[24] **J Mather**, *Invariant subsets of area-preserving homeomorphisms of surfaces*, Advances in Math. Suppl. Studies, 7B (1994) 331–351

[25] **J Mather**, *Topological proofs of some purely topological consequences of Caratheodory's Theory of prime ends*, in Selected Studies, (Th M Rassias, G M Rassias, editors) North-Holland (1982) 225–255

[26] **J Milnor**, *Lectures on the h-Corbordism Theorem* (notes by L Siebenmann and J Sondow,) Princeton Mathematical Notes, Princeton University Press (1965)

[27] **L Polterovich**, *Growth of maps, distortion in groups and symplectic geometry*, `arXiv:math.DS/0111050`

[28] **J Tits**, *Systèmes générateurs de groupes de congruence*, C. R. Acad. Sci. Paris Sér. A-B 283 (1976) Ai, A693–A695

[29] **W P Thurston**, *On the geometry and dynamics of diffeomorphisms of surfaces*, Bull. Amer. Math. Soc. 19 (1988) 417–431