

# A Geometric Criterion for Positive Topological Entropy

Keith Burns<sup>1</sup>, Howard Weiss<sup>2</sup>

<sup>1</sup>Department of Mathematics, Northwestern University, Evanston, IL 60208, USA.

Email: burns@math.psu.edu

<sup>2</sup>Department of Mathematics, Pennsylvania State University, University Park, PA 16802, USA.

Email: Weiss@math.psu.edu

Received: 29 April 1994

**Abstract:** We prove that a diffeomorphism possessing a homoclinic point with a topological crossing (possibly with infinite order contact) has positive topological entropy, along with an analogous statement for heteroclinic points. We apply these results to study area-preserving perturbations of area-preserving surface diffeomorphisms possessing homoclinic and double heteroclinic connections. In the heteroclinic case, the perturbed map can fail to have positive topological entropy only if the perturbation preserves the double heteroclinic connection or if it creates a homoclinic connection. In the homoclinic case, the perturbed map can fail to have positive topological entropy only if the perturbation preserves the connection. These results significantly simplify the application of the Poincaré–Arnold–Melnikov–Sotomayor method. The results apply even when the contraction and expansion at the fixed point is subexponential.

## 0. Introduction

Let  $f$  be a  $C^1$  diffeomorphism. A homoclinic point for  $f$  is a point that is both forward and backwards asymptotic to a hyperbolic periodic point of  $f$  (and is not the periodic point itself). In other words,  $q$  is a homoclinic point if there is a hyperbolic periodic point  $p$  such that  $q \in W^s(p) \cap W^u(p)$  and  $q \neq p$ . A homoclinic point is called transverse if the stable and unstable manifolds intersect transversely there. Poincaré [Po] discovered that the existence of a transverse homoclinic point causes very complicated dynamics. Smale [S] proved that a diffeomorphism  $f$  that possesses a transverse homoclinic point has a horseshoe. A horseshoe is a set  $A$  that is invariant under  $f^{n_0}$  for some  $n_0 \geq 1$  and has the property that  $f^{n_0}$  restricted to  $A$  is topologically conjugate to a nontrivial subshift of finite type, i.e., there exists a *homeomorphism*  $\pi : A \rightarrow \Sigma$  such that  $\pi \circ f^{n_0} = \sigma \circ \pi$ , where  $\Sigma$  is the shift space and  $\sigma : \Sigma \rightarrow \Sigma$  denotes the shift map. By a nontrivial subshift of finite type, we

---

The first author was partially supported by a Sloan Foundation Fellowship and an N.S.F. grant. The second author was partially supported by a National Science Foundation Postdoctoral Research Fellowship. Both authors would like to thank MSRI for their support during the period that much of this paper was written.

mean a subshift that is defined by an irreducible matrix that is not a permutation; the full two shift is the prototypical example. It is easy to show that a map that possesses a horseshoe has positive topological entropy; see Sect. 1.

It is natural to consider homoclinic points where the stable and unstable manifolds cross, but may not do so transversely. What one means by “cross” is clear in simple contexts, and can be made precise with ideas from elementary topology; a definition using homology is given in Sect. 4.

In the case of a surface diffeomorphism with a homoclinic point where the stable and unstable manifolds cross with finite order contact, Conley [Co] showed that there exist transverse homoclinic points in every neighborhood of the nontransverse homoclinic point and hence the map has a horseshoe and positive topological entropy (see [CR2] for a proof). Conley’s result should extend to higher dimensions, but we do not know a reference. The case of a crossing with infinite order contact is much more delicate. Various pathologies develop which do not appear in the finite order case. It seems extremely difficult to establish either the existence of a nearby transverse homoclinic point or the presence of a horseshoe.

The difficulty in finding a horseshoe lies in showing that the map  $\pi : A \rightarrow \Sigma$  is  $1 : 1$ . Let us say that a map  $f$  has a **horseshoe factor** if some power of  $f$  has a nontrivial subshift of finite type as a topological factor, i.e. there is a subset  $A$  that is invariant under  $f^{n_0}$  for some  $n_0 \geq 1$  and a continuous map  $\pi : A \rightarrow \Sigma$  that is *onto but not necessarily*  $1 : 1$  such that  $\pi \circ f^{n_0} = \sigma \circ \pi$ , where again  $\Sigma$  is the shift space and  $\sigma$  is the shift map. Any diffeomorphism  $f$  possessing a homoclinic point at which the stable and unstable manifolds have a topological crossing has a horseshoe factor. We give an elementary proof of this fact in the case of surface diffeomorphisms in Sect. 2 and generalize our argument to all dimensions in Sect. 4 (see Theorems 2.1 and 4.3). Gedeon, McCord and Mischaikow have recently proved a generalization of this result [GMM], using results of [MiMr].

The above result has a natural generalization to the case of a heteroclinic cycle, a chain of hyperbolic fixed points linked by heteroclinic points (as illustrated in Fig. 0.1). In arbitrary dimensions, we require that each pair of stable and unstable manifolds forming the chain should cross topologically (Theorem 4.8.). In dimension two, this requirement can be relaxed a little (Theorem 2.2).

Our results also apply to crossings of center manifolds for indifferent fixed points, and furnish examples of maps having positive topological entropy with no *apparent* hyperbolicity.

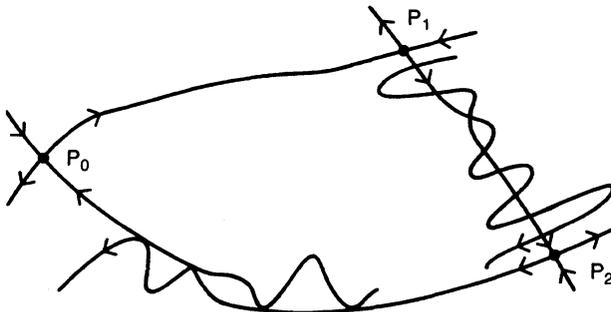


Fig. 0.1. A Heteroclinic Cycle.

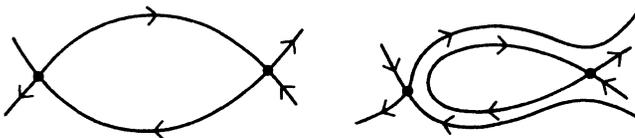
It follows from the above results that the diffeomorphisms in question have positive topological entropy. In dimension two, but not in higher dimensions, one can then apply a theorem of Katok [K1, K2] to conclude that if  $f$  is  $C^{1+\alpha}$  and has positive topological entropy, then  $f$  possesses a horseshoe that carries most of the entropy of  $f$ .

The proofs of the above theorems contain two stages. The first stage is to describe a method by which one proves that a set is being mapped across itself in the appropriate way to produce a horseshoe factor. The crucial second stage is to identify the relevant sets in each case when we wish to apply the general method. For the first stage of the proof, we use only elementary topology in the case of dimension two, and basic homology theory in arbitrary dimensions. Topological methods which can be used to find horseshoe factors are not new, although we believe that ours is more elementary than previous techniques. Our method is similar to an elegant and very general cohomological method of Easton [E].

Easton applied his method to a number of examples, but the only discussion of homoclinic points in [E] is in the context of the standard Smale horseshoe. Churchill and Rod [C, CR] have given cohomological methods, which apply to surface diffeomorphisms. The theorem on p. 548 of [Ch] is related to Theorem 2.1 of the present paper.

We apply the results of Sect. 2 to study area-preserving maps of  $\mathbb{R}^2$ . Let  $C_{AP}^1$  be the group of area-preserving  $C^1$  diffeomorphisms of  $\mathbb{R}^2$  with the  $C^1$  topology. Suppose  $f \in C_{AP}^1$  and  $f$  has two hyperbolic fixed points with a double heteroclinic connection (see Sect. 1). Such a map is illustrated in Fig. 0.2. Consider a small perturbation of  $f$  in  $C_{AP}^1$ . We show that the perturbed map can fail to have positive topological entropy only if the perturbation preserves the double heteroclinic connection or if it creates a homoclinic connection (see Sect. 1). A perturbation of the latter type is shown in Fig. 0.2. We also show that a small perturbation in  $C_{AP}^1$  of a map with a homoclinic connection must produce a map with positive topological entropy, unless the perturbation preserves the homoclinic connection. These theorems are formulated precisely in Sect. 3. They generalize ad hoc arguments used in [KW].

We believe that these results will have numerous applications. There are many papers in which the authors show that the Hamiltonian flow for a small Hamiltonian perturbation of a completely integrable Hamiltonian system with 2 degrees of freedom has positive topological entropy. This amounts to showing that a small area-preserving perturbation of a Poincaré map having a double heteroclinic connection has positive topological entropy. All authors apply a Poincaré–Arnold–Melnikov–Sotomayor (PAMS) type method [A, M, Po (Sect. 403), So] to each branch of the connection and need to verify that the PAMS function has a *simple* zero. This implies that the stable and unstable manifolds for the perturbed systems intersect



**Fig. 0.2.** A Map with a double heteroclinic connection and a perturbation with zero topological entropy.

with transversal crossings. This is usually quite difficult to verify, and with few exceptions (e.g. [Z]), has only been established for non-autonomous equations with periodic or quasi-periodic forcing. Obviously, this places a great limitation on the utility of the method. The results in Sect. 3 obviate the need to show the existence of a simple zero if one wishes to show that the topological entropy is positive; one needs only to show that the PAMS function has a zero (that implies crossing) and is not identically zero (that implies that the stable and unstable manifolds do not coincide) to conclude that the perturbed map has positive topological entropy. This remark considerably simplifies the process of showing a perturbed map has positive topological entropy.

In [KW], the authors construct a convex surface whose geodesic flow has positive topological entropy. They start with a triaxial ellipsoid whose geodesic flow possesses a Poincaré map having a double heteroclinic connection. The authors construct an explicit small local conformal perturbation and show that one connection breaks and the perturbed stable and unstable manifolds intersect (not necessarily transversely or with a topological crossing). It follows immediately from the results in Sect. 3 that the geodesic flow for the perturbed metric has positive topological entropy.

Another application of the results in Sect. 3 was discovered by Gabriel Paternain [Pa]. It follows from work of Kozlov [Ko], Tatarinov [T], and Ziglin [Z], that one can find a suitable Poincaré section for the geodesic flow on a Poisson sphere such that the Poincaré map has a double heteroclinic connection and that a small perturbation of the geodesic flow by a potential of Hess–Appel’rot type preserves one branch of the connection and splits the other branch such that the stable and unstable manifolds intersect transversely. Theorem 3.2 immediately implies that the perturbed flow has positive topological entropy. Paternain then applies the Maupertuis Principle to obtain an explicit real analytic metric on  $S^2$  with positive curvature whose geodesic flow has positive topological entropy.

## 1. Preliminaries

Let  $M$  be a smooth manifold and  $f : M \rightarrow M$  a diffeomorphism. Suppose  $p$  is a hyperbolic periodic point of  $f$  at which the stable and unstable subspaces have dimensions  $k$  and  $l$  respectively. The sets  $W^s(p) = \{x \in M : f^n(x) \rightarrow p\}$  and  $W^u(p) = \{x \in M : f^n(x) \rightarrow p\}$  are immersed smooth manifolds diffeomorphic to  $\mathbb{R}^k$  and  $\mathbb{R}^l$  respectively. They are called the **stable** and **unstable manifolds** of  $p$  respectively. Since we often want to exclude the point  $p$  from consideration, it is convenient to define the **deleted stable** and **deleted unstable manifolds**  $\widehat{W}^s(p) = W^s(p) \setminus \{p\}$  and  $\widehat{W}^u(p) = W^u(p) \setminus \{p\}$ . Note that  $\widehat{W}^s(p)$  and  $\widehat{W}^u(p)$  are connected, unless they are one dimensional, in which case they have two components.

A **homoclinic point** is a point  $q$  such that  $q \in \widehat{W}^s(p) \cap \widehat{W}^u(p)$  for some hyperbolic periodic point  $p$ . A **heteroclinic point** is a point  $q$  such that  $q \in \widehat{W}^s(p) \cap \widehat{W}^u(p')$  for distinct hyperbolic periodic points  $p$  and  $p'$ . Homoclinic and heteroclinic points are called **transverse** if they are points of transverse intersection of the relevant stable and unstable manifolds. A **heteroclinic cycle** is a sequence  $p_0, \dots, p_{r-1}$  of hyperbolic periodic points such that  $\widehat{W}^s(p_i) \cap \widehat{W}^u(p_{i+1}) \neq \emptyset$  for  $i = 0, \dots, r-1$ . In dealing with heteroclinic cycles, we make the convention that indices denoting which point in the cycle we are considering should be interpreted modulo  $r$ .

Consider now hyperbolic periodic points in a two dimensional manifold that have one dimensional stable and unstable manifolds. A periodic point  $p$  of this type has a **homoclinic connection** if one of the components of  $\widehat{W}^s(p)$  coincides with one of the components of  $W^u(p)$ . Two distinct periodic points  $p$  and  $p'$  of this type have a **heteroclinic connection** if a component of the deleted stable manifold of one of the points coincides with a component of the deleted unstable manifold of the other point, and have a **double heteroclinic connection** if one component of  $\widehat{W}^s(p)$  coincides with one component of  $\widehat{W}^u(p')$  and one component of  $\widehat{W}^s(p)$  coincides with one component of  $\widehat{W}^s(p')$ .

The (two-sided) subshift of finite type  $\Sigma_A$  defined by an  $n \times n$  matrix  $A$  of 0's and 1's is the subset of  $\Sigma_n = \{1, \dots, n\}^{\mathbb{Z}}$  consisting of all sequences which satisfy the condition that a symbol  $l \in \{1, \dots, n\}$  can follow a symbol  $k \in \{1, \dots, n\}$  if and only if  $A_{kl} = 1$ . The matrix  $A$  is **irreducible** if and only if for each  $i$  and  $j$  there is an  $n \geq 1$  such that the  $ij^{\text{th}}$  entry of  $A^n$  is positive. The Perron–Frobenius theorem [Ga, Sect. 13.2] tells us that  $A$  has an eigenvalue  $\lambda_A$  that is real and satisfies  $\lambda_A \geq |\lambda|$  for all eigenvalues  $\lambda$  of  $A$ . The following proposition is well known (see [W, Theorem 7.13(ii)]).

**Proposition 1.1.** *Let  $A$  be an irreducible  $n \times n$  matrix of 0's and 1's. Then the topological entropy of the shift on  $\Sigma_A$  is  $\log \lambda_A$ .*

The **full  $n$  shift** is the special case when  $A$  is an  $n \times n$  matrix with all entries 1 and  $\Sigma_A = \Sigma_n$ . It has entropy  $\log n$ .

**Lemma 1.2.** *Let  $A$  be an irreducible  $n \times n$  matrix of 0's and 1's. Then  $\lambda_A > 1$  unless  $A$  is a permutation matrix.*

*Proof.* If  $A$  is a permutation matrix, then  $(1, \dots, 1)$  is an eigenvector with eigenvalue 1. Now suppose that  $A$  is not a permutation matrix. Let  $s$  and  $S$  be the minimum and maximum row sums respectively of  $A$ . Then  $s \leq \lambda \leq S$ , cf. [Ga, Sect. 13.2]. Since  $A$  is irreducible, every row contains at least one 1, and hence  $\lambda_A \geq s \geq 1$ . We also know that every column of  $A$  contains at least one 1 and, since  $A$  is not a permutation matrix, at least one column of  $A$  contains a second 1. The Perron–Frobenius theorem tells us that  $A$  has an eigenvector  $v$  all of whose entries are positive. Consider the sum of the entries of  $Av$ . Each component of  $v$  must appear at least once, and at least one component must appear more than once, so this sum is greater than the sum of the entries of  $v$ . Hence 1 is not an eigenvalue of  $A$  and we must have  $\lambda_A > 1$ .  $\square$

It follows that a nontrivial subshift of finite type (as defined in the introduction) has positive topological entropy.

**Lemma 1.3.** *Suppose that the diffeomorphism  $f$  has a horseshoe factor. Then  $f$  has positive topological entropy.*

*Proof.* We are assuming that there is a subset  $A$  that is invariant under  $f^{n_0}$  for some  $n_0 \geq 1$  and has the property that  $f^{n_0}|_A$  is semiconjugate to a nontrivial subshift of finite type  $\sigma$ . Now  $h_{TOP}(\sigma) > 0$  and it follows from well known properties of topological entropy [W, Theorems 7.2, 7.10] that

$$h_{TOP}(f) \geq \frac{1}{n_0} h_{TOP}(f^{n_0}|_A) \geq \frac{1}{n_0} h_{TOP}(\sigma) > 0. \quad \square$$

We now develop the geometric machinery that will be used in the proofs of Theorems 2.1 and 4.3. Let  $f : M \rightarrow M$  be a diffeomorphism with a hyperbolic fixed point  $p$ . Let  $k = \dim W^s(p)$ ,  $l = \dim W^u(p)$  and  $m = k + l = \dim M$ . By choosing a suitable basis for  $T_p M$ , we may think of  $df(p)$  as a linear map  $L : \mathbb{R}^m \rightarrow \mathbb{R}^m$  which preserves the splitting  $\mathbb{R}^m = \mathbb{R}^k \oplus \mathbb{R}^l$ , contracts the  $\mathbb{R}^k$  factor and expands the  $\mathbb{R}^l$  factor. By the Hartman–Grobman theorem, there are a neighbourhood  $N$  of  $p$  and a homeomorphism  $h$  of  $N$  into  $\mathbb{R}^m$  with  $h(p) = (0, 0)$ , such that if  $x \in N$  and  $f(x) \in N$ , then  $h(f(x)) = L(h(x))$ . We may assume that  $N$  and  $h$  have been chosen so that  $h(N) = D^k(1) \times D^l(1)$ , where  $D^n(r)$  is the closed disc of radius  $r$  about the origin in  $\mathbb{R}^n$ . We may also assume that  $D^k(1) \times \{0\}$  and  $\{0\} \times D^l(1)$  lie in  $W^s(p) \cap N$  and  $W^u(p) \cap N$  respectively. In order to simplify our notation, we shall henceforth identify  $N$  with  $D^k(1) \times D^l(1)$  and suppress the homeomorphism  $h$ . Unless otherwise specified, distances in  $N$  will be measured with respect to the product of the Euclidean metrics on  $D^k(1)$  and  $D^l(1)$ . We shall call discs in  $N$  of the form  $\{x\} \times D^l(1)$  and  $D^k(1) \times \{y\}$  **vertical** and **horizontal** respectively.

**Lemma 1.4.** *Let  $V$  be a compact subset of  $W^u(p) \cap \text{Int} N$ . Suppose we are given  $\rho \in (0, 1)$  and  $\varepsilon > 0$ . Then for any large enough  $n$  the following hold;*

$$(1) f^{-n}(V) \subset \{0\} \times D^l(\rho),$$

$$(2) \text{If } (0, y) \in f^{-j}(V), \text{ then } f^n(D^k(1) \times \{y\}) \subset N \text{ and has diameter less than } \varepsilon.$$

*Proof.* We may assume  $\varepsilon < \text{dist}(V, \partial N)$ .

Choose a Riemannian metric  $g$  on the whole manifold  $M$  on which the diffeomorphism  $f$  acts. The restriction of  $g$  to  $N$  may not be equivalent to the Euclidean metric that we are using on  $N$ , because  $h : N \rightarrow D^k(1) \times D^l(1)$  may not be differentiable. The two metrics do, however, define the same topology on  $N$ . In particular, there is  $\alpha > 0$  such that if  $z, z' \in N$  and  $\text{dist}_g(z, z') < \alpha$  then  $\text{dist}(z, z') < \varepsilon$ . Also given  $\delta > 0$ , there is  $\beta(\delta) > 0$  such that for any  $z, z' \in N$  with  $\text{dist}(z, z') \leq \beta(\delta)$  we have  $\text{dist}_g(z, z') \leq \delta$ .

There is certainly an  $n_1$  such that for any  $n \geq n_1$  we have  $f^{-n}(V) \subset \{0\} \times D^l(1)$ . Set  $B = \sup\{\|Df^n(x)\| : x \in N\}$ . For  $n \geq 0$ , let

$$R_n = \{x \in N : f^k(x) \in D^k(1) \times D^l(1) \text{ for } 0 \leq k \leq n\}.$$

It is clear that  $R_n$  is a union of entire horizontal discs, the image under  $f^n$  of a horizontal disc in  $R_n$  lies in a horizontal disc, and the sets  $R_n$  and  $f^n R_n$  both decrease with  $n$ .

Choose  $n_2$  so that for any  $n \geq n_2$ ,  $R_n \subset D^k(1) \times D^l(\rho)$  and the image under  $f^n$  of each disc  $D^k(1) \times \{y\}$  in  $R_n$  lies in a disc  $D^k(\beta(\alpha/B)) \times \{y'\}$  for some  $y' \in D^l(1)$ . Then any  $n \geq n_1 + n_2$  has the desired properties.  $\square$

## 2. Topological Crossing of Stable and Unstable Manifolds in a Surface

It is intuitively evident what it means for two smooth curves in a surface to cross topologically: they should contain arcs with the behaviour shown in the left half of Fig. 2.1. This intuitive notion agrees with the two dimensional version of the definition given in Sect. 4. In the course of this section and the next, we shall need

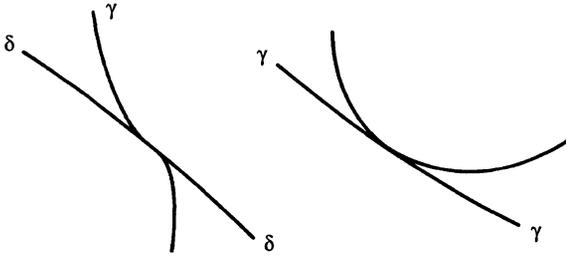


Fig. 2.1. Topological crossing and topological tangency.

to consider three different ways in which two curves in a surface can intersect: they may coincide; have a topological crossing; or have a topological tangency, in which case neither of the previous possibilities occurs (see the right half of Fig. 2.1).

**Theorem 2.1.** *Let  $f : M \rightarrow M$  be a diffeomorphism of a surface with a hyperbolic periodic point  $p$ . Assume that one component of  $\widehat{W}^s(p)$  and one component of  $\widehat{W}^u(p)$  have a topological crossing. Then some power of  $f$  has the full two shift as a topological factor.*

*Remarks.* We emphasize that the topological crossing need to be transverse and could be of infinite order.

Next suppose that  $p_0, \dots, p_{r-1}$  is a heteroclinic cycle in a surface. Then there are components  $V^s(p_i)$  of  $\widehat{W}^s(p_i)$  and  $V^u(p_i)$  of  $\widehat{W}^u(p_i)$ , such that  $V^u(p_i) \cap V^s(p_{i+1}) \neq \emptyset$  for  $i = 0, \dots, r - 1$ . (We remind the reader that in this context indices should be interpreted modulo  $r$ .)

**Theorem 2.2.** *Let  $p_0, \dots, p_{r-1}$  be a heteroclinic cycle in a surface  $M$  and let  $V^s(p_i)$  and  $V^u(p_i)$  be defined as above. Suppose that  $V^u(p_0)$  and  $V^s(p_1)$  cross topologically. Assume also that  $V^u(p_i)$  and  $V^s(p_{i+1})$  lie on the same side of any arc from  $p_i$  to  $p_{i+1}$  that can be formed by first traversing a piece of  $V^u(p_i)$  and then traversing a piece of  $V^s(p_{i+1})$ . Then some power of  $f$  has the full two shift as a topological factor and thus  $h_{TOP}(f) > 0$ .*

We now describe the method that will be used to prove the above theorems. Let  $N \subseteq M$  be homeomorphic to  $[-1, 1] \times [-1, 1]$ . In the following we shall identify  $N$  with  $[-1, 1] \times [-1, 1]$  and suppress the homeomorphism. Let  $R = [-1, 1] \times [-\rho, \rho]$ , where  $\rho \in (0, 1)$ .

We shall call a set  $V \subset \text{Int } N$  **essential** if  $V$  contains a curve joining the two components of  $N \setminus R$ .

A set  $S \subset R$  will be called a **horizontal strip** if

- (1)  $S$  is closed and path connected;
- (2)  $S$  contains a curve joining the left edge  $\{-1\} \times [-\rho, \rho]$  and the right edge  $\{1\} \times [-\rho, \rho]$  of  $R$ ;
- (3)  $\partial S$  is a Jordan curve which is the union of a finite number of arcs all of whose endpoints lie on the left edge or the right edge of  $R$ .

It is easily seen that  $\partial S$  contains exactly two curves joining the left edge  $\{-1\} \times [-\rho, \rho]$  and the right edge  $\{1\} \times [-\rho, \rho]$  of  $R$ , and  $S$  lies in the region of  $R$  bounded by these curves; see Fig. 2.2. We shall call the curve on which the second coordinate is larger  $c_{upper}$  and the other curve  $c_{lower}$ .

**Definition 2.3.** Let  $n$  be a positive integer and  $S$  a horizontal strip. We shall say that  $f^n$  stretches  $S$  across  $R$  if  $f^n S \subset \text{Int } N$ ,  $f^n(\partial S \cap \text{Int } R) \subset N \setminus R$ , and  $f^n$  maps  $c_{upper}$  and  $c_{lower}$  into opposite components of  $N \setminus R$ .

**Theorem 2.4.** Suppose  $N$  contains two disjoint closed horizontal strips  $S_0$  and  $S_1$  that are stretched across  $R$  by  $f^{n_0}$  for some  $n_0 \geq 1$ . Then  $f^{n_0}$  has the full two shift as a topological factor.

*Proof.* Let  $A = \bigcap_{k=-\infty}^{\infty} f^{kn_0}(S_0 \cup S_1)$ . Define  $\pi : A \rightarrow \Sigma_2$  so that for any  $z \in A$  the  $k^{\text{th}}$  term of the sequence  $\pi(z)$  is 0 if  $f^{kn_0}z \in S_0$  and 1 if  $f^{kn_0}z \in S_1$ . The continuity of  $f$  and  $f^{-1}$  implies that for any  $K > 0$  there is a neighbourhood  $N(z)$  of  $z$  in  $R$  such that if  $z' \in N(z) \cap A$ , then  $(\pi(z))_k = (\pi(z'))_k$  for  $-K \leq k \leq K$ . Thus  $\pi$  is continuous. It is easily verified that  $\pi \circ f^{n_0}|_A = \pi \circ \sigma$ , where  $\sigma$  is the shift on  $\Sigma_2$ . It remains to show that  $\pi$  is surjective.

**Key Lemma 2.5.** Let  $S$  be a horizontal strip and  $V$  an essential subset of  $N$ . Suppose that  $f^n$  stretches  $S$  across  $R$ . Then  $f^n(S \cap V)$  is essential.

*Proof.* It is obvious that  $V$  contains an arc  $\alpha : [0, 1] \rightarrow R$  such that  $\alpha(0)$  lies on  $c_{upper}$  and  $\alpha(1)$  lies on  $c_{lower}$  and  $\alpha(t) \in \text{Int } R$  for  $0 < t < 1$ . Let us call a component of  $\partial S \cap \text{Int } R$  positive if  $f^n$  maps it into the same component of  $N \setminus R$  as  $c_{upper}$  and negative otherwise. Let  $t_0 \in [0, 1)$  be the last time that  $\alpha(t)$  is in a positive component of  $\partial S \cap \text{Int } R$  and  $t_1 \in (t_0, 1]$  the next time that  $\alpha(t)$  is in  $\partial S$ . Observe that if  $\alpha$  exits  $S$  through a given component of  $\partial S \cap \text{Int } R$ , then  $\alpha$  must reenter  $S$  through the same component of  $\partial S \cap \text{Int } R$ . It is clear from this that the arc of  $\alpha$  between  $\alpha(t_0)$  and  $\alpha(t_1)$  must lie in  $S$ . Its image under  $f^n$  is a curve in  $\text{Int } N$  that joins the two components of  $N \setminus R$ . Hence  $f^n(S \cap V)$  is essential.  $\square$

**Lemma 2.6.** For each  $r \geq 1$ , the set  $f^{n_0}S_{\alpha_{-1}} \cap f^{2n_0}S_{\alpha_{-2}} \cap \dots \cap f^{rn_0}S_{\alpha_{-r}}$  is essential for every  $\alpha \in \Sigma_2$ .

*Proof.* Use induction of  $r$ . The case  $r = 1$  is trivial. Suppose the lemma holds for  $r$ . Then  $V = f^{n_0}S_{\alpha_{-2}} \cap \dots \cap f^{rn_0}S_{\alpha_{-r-1}}$  is essential for any  $\alpha \in \Sigma_2$ . It follows from

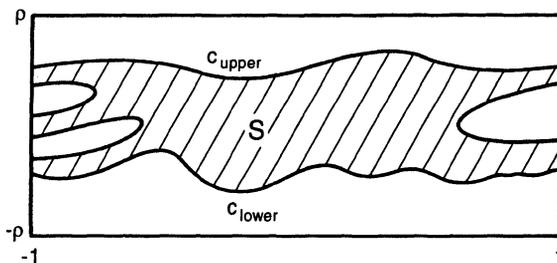


Fig. 2.2. A Horizontal Strip.

the Key Lemma that

$$f^{n_0}(S_{x_{-1}} \cap V) = f^{n_0}S_{x_{-1}} \cap f^{2n_0}S_{x_{-2}} \cap \dots \cap f^{(r+1)n_0}S_{x_{-r-1}}$$

is essential.  $\square$

**Corollary 2.7.** *For each  $\alpha \in \Sigma_2$  and each  $r \geq 0$ , the set  $A_r(\alpha) = \bigcap_{k=-r}^r f^{kn_0} S_{x_{-k}} \neq \emptyset$ .*

*Proof.* Let  $\beta = \sigma^{r+1}\alpha$ , where  $\sigma$  is the shift on  $\Sigma_2$ . Then

$$A_r(\alpha) = f^{-(r+1)n_0}(f^{n_0}S_{\beta_{-1}} \cap \dots \cap f^{2rn_0}S_{\beta_{-2r-1}}),$$

which is nonempty by the previous lemma.  $\square$

Since the sets  $A_r(\alpha)$  are compact and  $A_0(\alpha) \supset A_1(\alpha) \supset \dots$ , there is a point  $z \in \bigcap_{r=0}^\infty A_r(\alpha)$  for each  $\alpha$ . It is obvious that  $\pi(z) = \alpha$ . Thus  $\pi$  is surjective and the proof of Theorem 2.4 is complete.

We now apply Theorem 2.4 to prove Theorems 2.1 and 2.2.  $\square$

*Proof of Theorem 2.1.* After possibly replacing  $f$  by a power of  $f$ , we may assume that  $p$  is a fixed point of  $f$ . We use the notation introduced at the end of Sect. 1.

The first observation is that  $W^u(p) \cap N$  contains two arcs  $\alpha_1$  and  $\alpha_2$  that cross  $[-1, 1] \times \{0\}$  topologically. We choose  $\alpha_1$  to be  $[-1, 1] \times \{0\}$ . Now let  $\alpha$  be any arc that lies in  $\widehat{W}^s(p)$  and crosses  $\widehat{W}^u(p)$  topologically. After replacing  $\alpha$  by its image under a large enough power of  $f$ , we may assume that the intersection of  $\alpha$  with  $\widehat{W}^s(p)$  is contained in  $[-1, 1] \times \{0\}$ . One of the components of  $\alpha \cap N$  can now be chosen to be  $\alpha_2$ .

Choose  $\rho > 0$  such that the endpoints of  $\alpha_0$  and  $\alpha_1$  lie outside  $R \stackrel{\text{def}}{=} [-1, 1] \times [-\rho, \rho]$ .

**Lemma 2.8.** *For any large enough  $n$  there are disjoint closed horizontal strips  $S_0(n) \supset f^{-n}\alpha_0$  and  $S_1(n) \supset f^{-n}\alpha_1$  that are stretched across  $R$  by  $f^n$ .*

*Proof.* Choose  $\varepsilon > 0$  so that the closed  $\varepsilon$ -neighbourhoods of the end points of  $\alpha_0$  and  $\alpha_1$  lie in  $\text{Int}N \setminus R$  and the closed  $\varepsilon$ -neighbourhoods of  $\alpha_0$  and  $\alpha_1$  are disjoint and lie in  $\text{Int}N$ . Let  $n$  be large enough so that the conclusions of Lemma 1.4 hold with  $V = \alpha_0$  or  $\alpha_1$ . For  $i = 0, 1$ , let

$$S_i(n) = [-1, 1] \times \{y \in [-\rho, \rho] : (0, y) \in f^{-n}(\alpha_i)\}.$$

Each of the sets  $S_i$  is a rectangle of the form  $[-1, 1] \times [y_{i0}, y_{i1}]$ . These are certainly horizontal strips; in  $S_i$  the curve  $c_{upper}$  will be  $[-1, 1] \times \{y_{i1}\}$  and  $c_{lower}$  will be  $[-1, 1] \times \{y_{i0}\}$ . It is clear from our construction that  $S_0(n)$  and  $S_1(n)$  are closed and disjoint, and that  $\alpha_i$  lies in  $f^n S_i(n)$  for  $i = 0, 1$ . Each segment  $[-1, 1] \times \{y\}$  in one of the  $S_i$  is mapped by  $f^n$  to a curve of length less than  $\varepsilon$  which contains a point on one of the  $\alpha_i$ , and consequently lies in  $\text{Int}N$ . The edges  $[-1, 1] \times \{y_{i0}\}$  and  $[-1, 1] \times \{y_{i1}\}$  of  $S_i$  map into the  $\varepsilon$ -neighbourhoods of the endpoints of  $\alpha_i$ , and thus map into different components of  $N \setminus R$ .  $\square$

Theorem 2.1 follows immediately from the above lemma and Theorem 2.4.  $\square$

*Proof of Theorem 2.2.* After replacing  $f$  by one of its powers, and the original cycle by a possibly longer one containing the orbits of the periodic points involved in the original cycle, we may assume that all of the periodic points  $p_0, \dots, p_{r-1}$  are fixed. We shall assume that the surface is oriented and the diffeomorphism  $f$  is orientation preserving. If the surface is not orientable, we can replace  $f$  by a lift to the double cover of the surface; and if  $f$  is not orientation preserving we can replace  $f$  by  $f^2$ .

Choose a smooth embedding of  $[-1, 1] \times [-1, 1]$  whose image  $N$  is as shown in Fig. 2.3. The embedding is chosen so that  $\{-1\} \times [-1, 1]$  maps into  $V^u(p_0)$  and there is  $\rho \in (0, 1)$  such that

(1)  $[-1, 1] \times \{-\rho\}$  and  $[-1, 1] \times \{\rho\}$  map into  $V^s(p_1)$ .

(2) The image  $R$  of  $[-1, 1] \times [-\rho, \rho]$  lies on the same side of  $V^u(p_0)$  as  $V^s(p_0)$  and on the same side of  $V^s(p_1)$  as  $V^u(p_1)$ .

Our strategy is to show that any sufficiently large iterate  $f^n$  maps  $R$  across itself twice as shown in Fig. 2.3. We can then apply the argument in the proof of Theorem 1. Let  $\alpha$  and  $\beta$  be the **vertical sides** of  $R$ , namely  $\{-1\} \times [-\rho, \rho]$  and  $\{1\} \times [-\rho, \rho]$  respectively. We first show that  $\alpha$  and  $\beta$  map across  $R$  in the proper way.

In the following discussion all curves are embedded curves. Pick a number  $\sigma \in (\rho, 1)$  such that  $f^n \alpha \cup f^n \beta$  is transverse to  $[-1, 1] \times [-\sigma, \sigma]$  for all  $n$ .

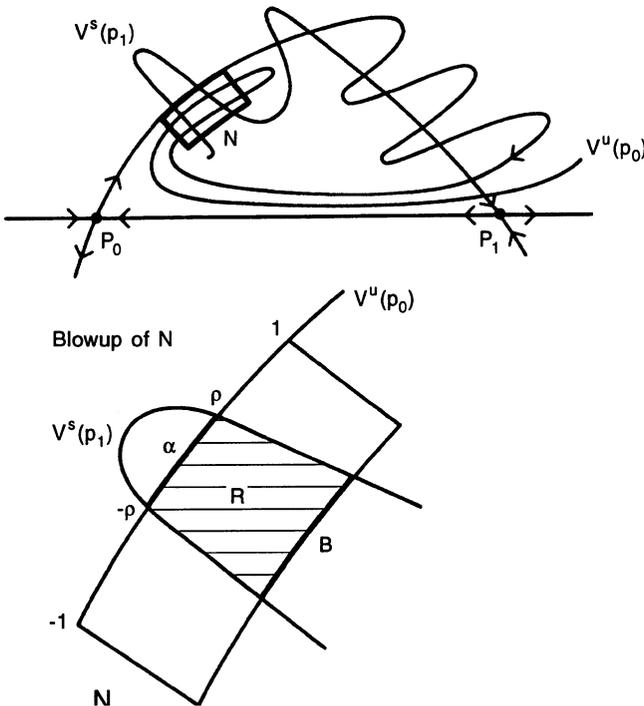


Fig. 2.3. Proof of Theorem 2.2.

**Definition 2.9.** Let  $\gamma : [0, 1] \rightarrow M$  be a curve. We shall call  $\gamma$  **good** if for all sufficiently large  $n$ , there is  $t_n \in [0, 1]$  such that

- (1)  $f^n \circ \gamma(0) \notin N$ ;
- (2)  $f^n \circ \gamma([0, t_n])$  does not intersect any part of the boundary of  $N$  except  $(-1, 1) \times \{-1\}$ ;
- (3)  $f^n \circ \gamma(t_n) \in (-1, 1) \times (\sigma, 1)$ .

This is obviously an open property with respect to the Hausdorff topology on compact curves.

**Definition 2.10.** Let  $\eta$  be a compact curve and  $\gamma$  a curve. We say that  $f^n \gamma$  accumulates on  $\eta$  if there exist compact curves  $\gamma_n \subset \gamma$  such that  $f^n \gamma_n$  converges to  $\eta$  in the Hausdorff topology on compact curves.

We shall use the following obvious lemma.

**Lemma 2.11.** Suppose  $\gamma$  is a curve that accumulates on a compact curve  $\eta$ . If  $\eta$  is good, then  $\gamma$  is good.

**Proposition 2.12.** Any curve which has one endpoint on  $V^s(p_1)$  and lies on the same side of  $V^s(p_1)$  as  $V^u(p_1)$  is good.

*Proof.* Divide the integers modulo  $r$  into blocks of consecutive integers so that  $i$  and  $j$  are in the same block if and only if  $V^u(p_k) = V^s(p_{k+1})$  for  $i \leq k < j$ . Observe that 1 is the first element of a block.

Let  $i_1 = 1, \dots, i_k$  be the first elements of the blocks. Observe that if  $\gamma$  is a curve with an endpoint in  $V^s(p^{i_j})$  that lies on the same side of  $V^s(p_{i_j})$  as  $V^s(p_{i_j})$ , then  $f^n \gamma$  accumulates onto every compact curve contained in  $\{p_{i_j}\} \cup V^u(p_{i_j}) \cup \dots \cup \{p_{i_{j+1}-1}\} \cup V^u(p_{i_{j+1}-1})$ .

We now prove by a descending induction that if  $1 \leq j \leq k$ , then any curve with an endpoint on  $V^s(p_{i_j})$  that lies on the same side of  $V^u(p_{i_j})$  as  $V^s(p_{i_j})$  is good. The case  $j = k$  is obvious. The inductive step is contained in the following lemma.  $\square$

**Lemma 2.13.** Suppose  $\gamma$  is a curve that has an endpoint on  $V^s(p_{i_j})$  and lies on the same side of  $V^s(p_{i_j})$  as  $V^u(p_{i_j})$ . Then at least one of the following occurs:

- (1)  $f^n \gamma$  accumulates on a curve  $\beta$  that has an endpoint on  $V^s(p_{i_{j+1}})$  and lies on the same side of  $V^s(p_{i_{j+1}})$  as  $V^u(p_{i_{j+1}})$ .
- (2)  $f^n \gamma$  contains a curve  $\alpha$  that has an endpoint on  $V^s(p_{i_{j+1}})$  and lies on the same side of  $V^s(p_{i_{j+1}})$  as  $V^u(p_{i_{j+1}})$ .

*Proof.* There are ways (illustrated in Fig. 2.4) in which  $V^u(p_{i_{j+1}-1})$  and  $V^s(p_{i_{j+1}})$  can intersect. In Case I they have with a topological crossing; both phenomena (1) and (2) occur. In case II they have a topological tangency and  $V^u(p_{i_{j+1}-1})$  lies on the same side of  $V^s(p_{i_{j+1}})$  as  $V^u(p_{i_{j+1}})$ ; only phenomenon (1) occurs. In case III they have a topological tangency and  $V^u(p_{i_{j+1}-1})$  lies on the opposite side of  $V^s(p_{i_{j+1}})$  from  $V^u(p_{i_{j+1}})$ ; only phenomenon (2) occurs. The hypothesis in the theorem, that  $V^u(p_i)$  and  $V^s(p_{i+1})$  lie on the same side of any arc from  $p_i$  to  $p_{i+1}$  that can be formed by first traversing a piece of  $V^u(p_i)$  and then traversing a piece

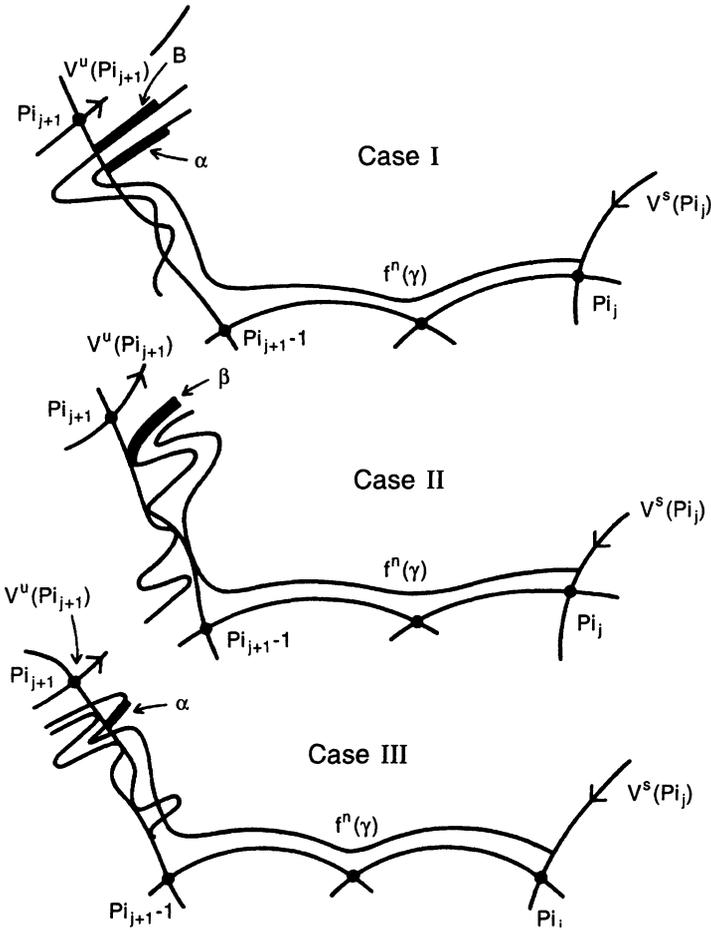


Fig. 2.4. The three cases in Lemma 2.13.

of  $V^s(p_{i+1})$ , guarantees that  $V^s(p_i)$  lies as shown in Fig. 2.4 and not on the other side of  $V^u(p_i)$ .  $\square$

It follows from Proposition 2.12 that for any large enough  $n$ , both  $f^n\alpha$  and  $f^n\beta$  contain arcs that lie in the interior of  $N$  and join  $[-1, 1] \times \{\sigma\}$  to  $[-1, 1] \times \{-\sigma\}$ . Henceforth we consider a fixed value of  $n$  for which such arcs exist. Arcs of  $f^n\alpha \cup f^n\beta$  that join  $[-1, 1] \times \{\sigma\}$  to  $[-1, 1] \times \{-\sigma\}$ , lie in the interior of  $N$ , and do not contain any shorter curves with these properties will be called **connecting curves**. Two connecting curves cannot intersect because they lie in  $f^n\alpha \cup f^n\beta$  which is the homeomorphic image of two intervals. There are only finitely many connecting curves, because any horizontal line  $[-1, 1] \times \{y\}$  that is transverse to  $f^n\alpha \cup f^n\beta$  contains only finitely many points that can lie on connecting curves. The first coordinate in  $N$  induces an ordering on the connecting curves in an obvious way.

**Lemma 2.14.** *Two connecting curves are adjacent in this ordering if and only if they are joined by a curve in  $[-1, 1] \times (-\sigma, \sigma)$  that does not intersect  $f^n\alpha \cup f^n\beta$ .*

*Proof.* It is obvious that two connecting curves are adjacent if they are joined by a curve in  $[-1, 1] \times (-\sigma, \sigma)$  that does not intersect  $f^n\alpha \cup f^n\beta$ .

Conversely, suppose that  $c_1$  and  $c_2$  are adjacent curves. Choose a segment  $\gamma$  of the form  $[x_1, x_2] \times \{y_1\}$  that joins  $c_1$  to  $c_2$  and is transverse to  $f^n\alpha \cup f^n\beta$ . Starting at  $(x_1, y_1)$ , follow  $\gamma$  until it intersects  $f^n\alpha \cup f^n\beta$ . On one side of  $\gamma$  or the other we can now move along  $f^n\alpha$  or  $f^n\beta$  and intersect  $\gamma$  before hitting the lines  $[1, 1] \times \{\pm 1\}$ . We then continue along  $\gamma$  to the next intersection with  $f^n\alpha \cup f^n\beta$ . After finitely many steps we reach  $(x_2, y)$ .  $\square$

**Lemma 2.15.** *Suppose  $\gamma$  lies in  $N$ , joins  $f^n\alpha$  and  $f^n\beta$  and does not contain any curve that joins  $f^n\alpha$  and  $f^n\beta$ . Then  $f^{-n}\gamma$  lies in  $R$ .*

*Proof.* Clearly  $f^{-n}\gamma$  cannot cross  $\alpha$  or  $\beta$ . Also  $f^{-n}\gamma$  cannot cross the piece  $\delta$  of  $V^s(p_1)$  marked in Fig. 2.5, for  $f^n\delta$  lies near  $p_1$  and is disjoint from  $N$ .  $\square$

We divide the connecting curves into two classes—those of type  $\alpha$  that lie in  $f^n\alpha$  and those of type  $\beta$  that lie in  $f^n\beta$ .

**Lemma 2.16.** *There are connecting curves  $\hat{\alpha}_1, \hat{\alpha}_2$  of type  $\alpha$  and  $\hat{\beta}_1, \hat{\beta}_2$  of type  $\beta$  such that each pair  $\hat{\alpha}_i, \hat{\beta}_i, i = 1, 2$ , is adjacent with respect to the ordering described above.*

*Proof.* Let  $x$  and  $y$  be the points where the two arcs of  $f^n\alpha$  that emanate from  $W^s(p_1)$  first intersect  $[-1, 1] \times \{\sigma\}$ . These arcs, the piece of  $W^s(p_1)$  that join the initial points, and the segment of  $[-1, 1] \times \{\sigma\}$  form a Jordan curve. The two arcs  $\beta_1, \beta_2$  of  $f^n\beta$  that emanate from  $W^s(p_1)$  and end at the first intersection with  $[-1, 1] \times \{\sigma\}$  enter the interior of the Jordan curve described above and do not cross it. Let  $\alpha_i^*$  (resp.  $\beta_i^*$ ) be the connecting curves of type  $\alpha$  (resp.  $\beta$ ) contained in  $\alpha_i$  (resp.  $\beta_i$ ). It is clear that  $\beta_1^*$  and  $\beta_2^*$  lie in between  $\alpha_1^*$  and  $\alpha_2^*$ ; with the right choice of numbering we may assume that they occur in the order  $\alpha_1^*, \beta_1^*, \beta_2^*$ , and  $\alpha_2^*$ . There must be connecting curves  $\hat{\alpha}_i$  and  $\hat{\beta}_i$  between each pair  $\alpha_i^*$  and  $\beta_i^*$  which have the desired properties.  $\square$

**Lemma 2.17.** *Suppose that  $\hat{\alpha}$  and  $\hat{\beta}$  are adjacent connecting curves. Then the region in  $[-1, 1] \times [-\sigma, \sigma]$  between  $\hat{\alpha}$  and  $\hat{\beta}$  contains the image under  $f^n$  of a horizontal strip that is stretched across  $R$  by  $f^n$ .*

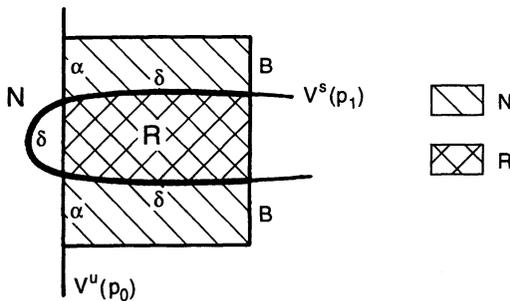


Fig. 2.5. Proof of Lemma 2.15.

*Proof.* By Lemma 2.14, there is a curve  $\gamma \subset [-1, 1] \times (-\sigma, \sigma)$  that joins  $\hat{\alpha}$  to  $\hat{\beta}$  and does not intersect  $f^n\alpha \cup f^n\beta$ . Let  $C$  be the component of  $[-1, 1] \times (-\sigma, \sigma) \setminus (f^n\alpha \cup f^n\beta)$  that contains  $\gamma$ . Let  $S = f^{-n}C$ . Note that  $(\alpha \cup \beta) \cap S = \emptyset$ . The other sides of  $R$ , namely  $\{-1\} \times [-\rho, \rho]$  and  $\{1\} \times [-\rho, \rho]$  lie on  $W^s(p_1)$  and do not map into  $N$  under  $f^n$ . Hence  $f^n(\partial R) \cap C = \emptyset$  and  $S \cap \partial R = \emptyset$ . But  $S$  is connected and contains  $f^{-n}\gamma$ , which lies in  $R$  by Lemma 2.15. Hence  $S \subset R$ .

Since  $f^n(\alpha \cup \beta)$  is transverse to  $[-1, 1] \times \{-\sigma, \sigma\}$ , we see that  $\partial C$  is a Jordan curve consisting of a finite number of arcs belonging to  $f^n\alpha, f^n\beta, [-1, 1] \times \{-\sigma\}$  and  $[-1, 1] \times \{\sigma\}$ . It follows that  $\partial S$  is a Jordan curve consisting of a finite number of arcs belonging to  $\alpha, \beta, f^n([-1, 1] \times \{-\sigma\})$  and  $f^n([-1, 1] \times \{\sigma\})$ . The end-points of these arcs must lie on  $\alpha \cup \beta$ . Furthermore  $S$  is obviously closed and path connected and contains the curve  $f^{-n}\gamma$  that joins  $\alpha$  to  $\beta$ . Thus  $S$  is a horizontal strip.

We now show that  $f^n$  stretches  $\bar{S}$  across  $R$ . It is clear that  $f^n$  maps  $S$  into  $\text{Int } N$  and maps  $\partial S \cap \text{Int } R$  into  $[-1, 1] \times \{-\sigma, \sigma\} \subset N \setminus R$ , so we have only to show that  $c_{upper}$  and  $c_{lower}$  map into opposite components of  $N \setminus R$ . The curves  $c_{upper}$  and  $c_{lower}$  do not belong to  $\alpha \cup \beta$ , so  $f^n$  must map them into  $[-1, 1] \times \{-\sigma, \sigma\}$ . Since  $c_{upper}$  and  $c_{lower}$  lie on opposite sides of  $f^{-n}\gamma$  in  $R$ , their images under  $f^n$  lie on opposite sides of  $\gamma$  in the region of  $[-1, 1] \times [-\sigma, \sigma]$  that lies between  $\hat{\alpha}$  and  $\hat{\beta}$ . Hence  $c_{upper}$  and  $c_{lower}$  map into opposite components of  $N \setminus R$ .  $\square$

We see from the two previous lemmas that the hypotheses of Theorem 2.4 are satisfied. This completes the proof of Theorem 2.2.  $\square$

*Remarks.*

(i) In the proofs of the above theorems we have not used hyperbolicity of the fixed points: it would suffice for them to have stable manifolds that contract under  $f$ , and unstable manifolds that contract under  $f^{-1}$  at a subexponential rate. It is not difficult to devise explicit examples that have fixed points with these properties. There are also natural examples in celestial mechanics.

(ii) The hypothesis in Theorem 2.1 that there be a topological crossing is essential. Gavrilov and Šilnikov [GŠ] have constructed a surface diffeomorphism with zero topological entropy which has a homoclinic point at which the stable and unstable manifolds have a topological tangency.

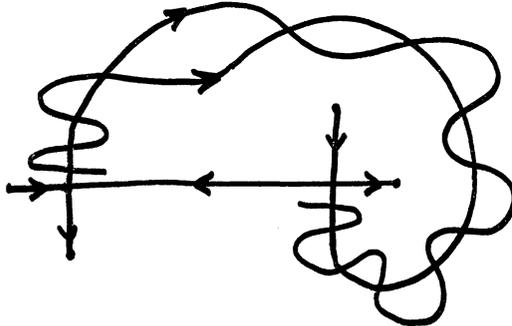


Fig. 2.6.

(iii) The hypothesis in Theorem 2.1 that  $V^u(p_i)$  and  $V^s(p_{i+1})$  lie on the same side of any arc from  $p_i$  to  $p_{i+1}$  that can be formed by first traversing a piece of  $V^u(p_i)$  and then traversing a piece of  $V^s(p_{i+1})$  is necessary. Figure 2.6 shows an example in which both this hypothesis and the conclusion of Theorem 2.2 are false.

### 3. Perturbations of Area-Preserving Maps of $\mathbb{R}^2$

Let  $C_{AP}^1$  denote the space of  $C^1$  area-preserving diffeomorphisms of  $\mathbb{R}^2$ . Suppose  $f \in C_{AP}^1$  has a hyperbolic fixed point  $p$  with a homoclinic connection. The structural stability of hyperbolic fixed points ensures that any diffeomorphism  $g$  that is close enough to  $f$  in the  $C^1$  topology will have a fixed point  $p_g$  that is close to  $p$ .

Suppose that  $V^s(p)$  and  $V^u(p)$  are the components of  $\widehat{W}^s(p)$  and  $\widehat{W}^u(p)$  which coincide. Let  $V_g^s(p_g)$  and  $V_g^u(p_g)$  be the components of  $\widehat{W}_g^s(p_g)$  and  $\widehat{W}_g^u(p_g)$  that are close to  $V^s(p)$  and  $V^u(p)$  respectively.

**Theorem 3.1.** *There exists a neighborhood  $\mathcal{U}$  of  $f$  in  $C_{AP}^1$  such that if  $g \in \mathcal{U}$  and  $V_g^s(p_g) \neq V_g^u(p_g)$ , then  $h_{TOP}(g) > 0$ .*

*Proof.* In view of Theorem 2.1, we have only to show that  $p_g$  must have a homoclinic point if  $g$  is close enough to  $f$ . It is a general fact, which goes back to Poincaré, that homoclinic points persist under small perturbations in  $C_{AP}^1$  (see [McMe] for an exposition of the proof).

We now give a proof that any  $g$  close enough to  $f$  has a homoclinic point, which we will modify in the proof of the next theorem.

Let  $\Gamma^s$  and  $\Gamma^u$  denote  $V_g^s(p_g)$  and  $V_g^u(p_g)$  respectively. Pick nearby points  $x_s \in \Gamma^s$  and  $x_u \in \Gamma^u$ . Let  $\Delta_s$  and  $\Delta_u$  be the arcs of  $\Gamma^s$  and  $\Gamma^u$  respectively joining  $x_s, x_u$  to  $g(x_s), g(x_u)$ . Then  $\Delta_s$  and  $\Delta_u$  are  $C^1$  close and we can find a short arc  $\alpha$  from  $x_s$  to  $x_u$  such that

$$\alpha \cap (g^{-1}(\Delta_s) \cup \Delta_s) = \{x_s\} \quad \text{and} \quad \alpha \cap (g^{-1}(\Delta_u) \cup \Delta_u) = \{x_u\}.$$

Suppose now that  $p_g$  does not have a homoclinic point. Then  $\alpha \cup \Delta_s \cup g(\alpha) \cup \Delta_u$  forms a Jordan curve that bounds a compact region  $R_0$  with positive area.

Let  $R$  be the compact region bounded by the arc of  $\Gamma^s$  from  $p_g$  to  $x_s, \alpha$ , and the arc of  $\Gamma^u$  from  $g(x_u)$  to  $p_g$ , as shown in Fig. 3.1. Then  $g(R)$  is the compact

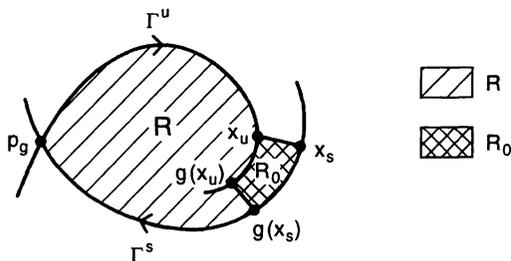


Fig. 3.1. Proof of Theorem 3.1.

region bounded by the arc of  $\Gamma^s$  from  $p_g$  to  $g(x_s), g \circ \alpha$ , and the arc of  $\Gamma^u$  from  $g(x_u)$  to  $p_g$ . One sees that either  $R \subset gR$  or  $gR \subset R$ , and that in either case their areas differ by area  $(R_0)$ . This is impossible if  $g$  is area-preserving.  $\square$

Now suppose  $f \in C_{AP}^1$  has a double heteroclinic connection between hyperbolic fixed points  $p$  and  $p'$ . We describe natural and easily verifiable geometric conditions that will ensure that a nearby area-preserving map will have positive topological entropy. As explained in the introduction the essence of the result is that a small area-preserving perturbation of  $f$  can fail to produce positive topological entropy only if it preserves the double heteroclinic connection or if it creates a homoclinic connection.

The structural stability of hyperbolic fixed points ensures that any diffeomorphism  $g$  that is close enough to  $f$  in the  $C^1$  topology will have fixed points  $p_g$  and  $p'_g$  that are close to  $p$  and  $p'$  respectively. Suppose that  $V^s(p), V^u(p), V^s(p')$ , and  $V^u(p')$  are the components of  $\widehat{W}^s(p), \widehat{W}^u(p), \widehat{W}^s(p')$ , and  $\widehat{W}^u(p')$  which are involved in the double heteroclinic connection described above, so that  $V^s(p) = V^u(p')$  and  $V^u(p) = V^s(p')$ . Let  $V_g^s(p_g)$  be the component of  $\widehat{W}_g^s(p_g)$  that is close to  $V^s(p)$ , and define  $V_g^u(p_g), V_g^s(p'_g)$ , and  $V_g^u(p'_g)$  analogously.

**Theorem 3.2.** *There exists a neighborhood  $\mathcal{U}$  of  $f$  in  $C_{AP}^1$  such that if  $g \in \mathcal{U}$  and*

- (1)  $V_g^s(p_g) \neq V_g^u(p'_g)$  or  $V_g^u(p_g) \neq V_g^s(p'_g)$  and
- (2)  $V_g^s(p_g) \neq V_g^u(p_g)$  and  $V_g^s(p'_g) \neq V_g^u(p'_g)$ , then  $h_{TOP}(g) > 0$ .

The following corollary, which contains the result used in [KW], is an easy consequence of Theorem 3.2.

**Corollary 3.3.** *There exists a neighborhood  $\mathcal{U}$  of  $f$  in  $C_{AP}^1$  such that if  $g \in \mathcal{U}, V_g^s(p_g) \neq V_g^u(p'_g)$  and  $V_g^s(p_g) \cap V_g^u(p'_g) \neq 0$ , then  $h_{TOP}(g) > 0$ .*

*Remarks.*

- (1) The above results include the case where one of the connections is not broken.
- (2) The area-preserving serving hypothesis is easily seen to be essential in these results.
- (3) We emphasize that we are considering a small perturbation of  $f$ .

In our proof of Theorem 3.2 we shall use the fact that  $g$  can be linearized near  $p_g$  by a  $C^1$  change of coordinates that depends continuously on  $g$ . This  $C^1$  linearization is not essential for the proof of this theorem but it significantly simplifies the exposition.

**Linearization Theorem 3.4.** *Let  $f$  be a diffeomorphism of  $\mathbb{R}^2$  with a hyperbolic fixed point  $p$ . For  $g$  close to  $f$  in the  $C^1$  topology, let  $p_g$  be the hyperbolic fixed point of  $g$  that is close to  $p$ . Let  $\lambda_g$  and  $\mu_g$  be the eigenvalues of  $D_{p_g}g$  on the stable and unstable subspaces at  $p_g$  respectively. Let  $A_g$  be the linear map of  $\mathbb{R}^2$  defined by the matrix*

$$\begin{pmatrix} \lambda_g & 0 \\ 0 & \mu_g \end{pmatrix}.$$

Then there exists a neighborhood  $U$  of  $p$  diffeomorphic to the open unit ball  $B$  in  $\mathbb{R}^2$ , a neighborhood  $\mathcal{N} \subset \text{Diff}^1(\mathbb{R}^2)$  of  $f$ , and a continuous mapping  $H : \mathcal{N} \rightarrow \text{Diff}^1(U, B)$  such that  $A_g \circ H(g) = H(g) \circ g$  on  $U \cap g^{-1}U$  for each  $g \in \mathcal{N}$ .

*Remarks.* This theorem is due to Hartman [Ha]. It is false in higher dimensions and is false if one asks for the linearization to be  $C^2$  or smoother. We outline a proof. One may assume that  $p_g$  is the origin for all  $g$ . Change  $g$  outside of a ball around 0 so that it agrees with a hyperbolic linear map outside of a still larger ball around 0. One can easily show that the stable and unstable manifolds are  $C^1$  ([HPS]). One then applies the one-dimensional Sternberg Linearization Theorem [St] to the local stable and unstable manifolds of 0. The theorem immediately follows by observing the continuous dependence of the conjugating map in Sternberg’s proof in the  $C^1$  topology. We thank A. Katok for helpful remarks about this proof.

*Proof of Theorem 3.2.* Choose  $U$  and the linearizing map  $H$  as above. We may assume that  $H(f)$  maps the region inside the double heteroclinic connection for  $f$  into the first quadrant of  $\mathbb{R}^2$ .

Define the coordinates  $x_g$  and  $y_g$  on  $U$  by  $H(g)(q) = (x_g(q), y_g(q))$ . Let  $\mathcal{F}$  and  $\mathcal{F}^\perp$  be the foliations of  $U \setminus \{p_g\}$  that are the level curves of the functions  $Q(q) = x_g(q)y_g(q)$  and  $q \rightarrow x_g(q)^2 - y_g(q)^2$  respectively. We say that  $q_1$  is **above**  $q_2$  if  $q_1$  and  $q_2$  lie on the same leaf of  $\mathcal{F}^\perp$  and  $Q(q_1) > Q(q_2)$ . Suppose that  $\alpha_1$  and  $\alpha_2$  are  $C^1$  arcs which are transverse to  $\mathcal{F}^\perp$  and have the same endpoints. We say that  $\alpha_1$  and  $\alpha_2$  are **equivalent** if the regions  $R_1 = \{q : q \text{ lies above a point of } \alpha_1 \text{ and below a point of } \alpha_2\}$  and  $R_2 = \{q : q \text{ lies above a point of } \alpha_2 \text{ and below a point of } \alpha_1\}$  have the same area. It is easily seen that  $p_g$  cannot lie between two equivalent curves. Furthermore two equivalent curves must either cross topologically or coincide.

Now suppose that  $W$  is a connected  $C^1$  curve in  $U$  that is part of a curve that is invariant under  $g$ . We say that  $W$  is **carried by** a leaf  $L$  of  $\mathcal{F}$  if there is a point  $q \in W \cap L$  such that the arc of  $W$  from  $q$  to  $g(q)$  lies in  $U$ , is transverse to  $\mathcal{F}^\perp$  and is equivalent to the arc of  $L$  from  $q$  to  $g(q)$ . See Fig. 3.2.

An easy continuity argument proves the following lemma.

**Lemma 3.5.** *Let  $W$  be a connected  $C^1$  curve in  $U$  that is part of a curve that is invariant under  $g$ . Suppose that for some  $q \in W$  the arc of  $W$  from  $g^{-1}(q)$  to*

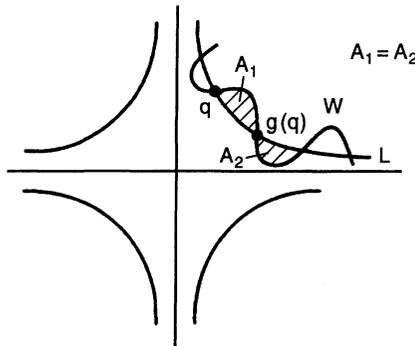


Fig. 3.2.  $W$  is carried by  $L$ .

$g(q)$  lies in  $U \setminus \{p_g\}$  and is transverse to  $\mathcal{F}^\perp$ . Then there is a unique leaf  $L$  of  $\mathcal{F}$  that carries  $W$ .

Choose a compact neighborhood  $U_0$  of  $p_g$  with the property that  $\bigcup_{n=-100}^{100} g^n(U_0) \subset U$ . Let  $q_s$  be the first point on  $V_g^s(p'_g)$  (as one moves along  $V_g^s(p'_g)$  starting from  $p'_g$ ) that lies in  $U_0$ . Let  $W^s$  be the path component of  $U \cap V_g^s(p'_g)$  that contains  $q_s$ . Define  $W^u$  analogously. Let  $L^s$  and  $L^u$  be the leaves that carry  $W^s$  and  $W^u$  respectively.

If  $L^s = L^u$ , then  $p_g$  will have a homoclinic point in  $U$ , and it follows from Theorem 3.1 that either  $g$  has positive topological entropy or there exists a homoclinic connection for  $p'_g$ . Henceforth we will assume that  $L^s \neq L^u$ .

Let  $\Gamma^s$  be the curve that starts at  $p'_g$ , follows  $V_g^s(p'_g)$  until the first intersection with  $L^s$  and then follows  $L^s$  (in the direction from  $q$  towards  $g(q)$ , where  $q$  is the point of changeover from  $V_g^s(p'_g)$  to  $L^s$ ). Define  $\Gamma^u$  analogously. Suppose that we can choose points  $x_s$  on  $\Gamma^s$  and  $x_u$  on  $\Gamma^u$  such that the arcs from  $x_s$  to  $g(x_s)$  and from  $x_u$  to  $g(x_u)$  are  $C^1$  close. Then we can use the argument in the proof of Theorem 3.1 to see that area preservation will be violated unless  $p_g$  has a homoclinic connection. Even though  $\Gamma^s$  and  $\Gamma^u$  are no longer invariant, their construction ensures that the area-preservation argument still applies.

It is clear that we can choose  $x_s$  and  $x_u$  with the desired properties if  $L^s$  or  $L^u$  lies in the interior of the first quadrant, as illustrated in Fig. 3.3.

If neither  $L^s$  nor  $L^u$  lies in the interior of the first quadrant,  $V_g^s(p_g)$  and  $V_g^u(p_g)$  lie “inside”  $\Gamma^s$  and  $\Gamma^u$ . If we interchange the roles of  $p_g$  and  $p'_g$  in the above argument, we will then find that the  $L^s$  and  $L^u$  constructed from  $V_g^s(p_g)$  and  $V_g^u(p_g)$  lie in the closure of the quadrant bounded by  $V_g^s(p_g)$  and  $V_g^u(p_g)$  in the linearizing

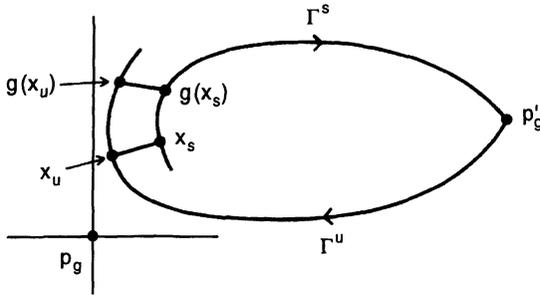


Fig. 3.3.

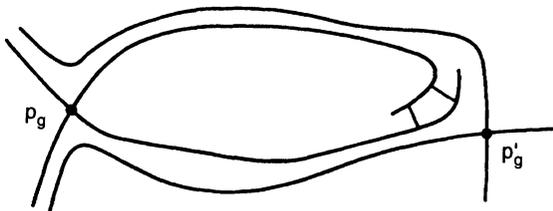


Fig. 3.4.

coordinates around  $p'_g$ ; see Fig. 3.4. Moreover, if one of the original  $L^s$  and  $L^u$  lay outside the closure of the quadrant bounded by  $V_g^s(p_g)$  and  $V_g^u(p_g)$  (in the coordinates near  $p_g$ ), one of the new  $L^s$  and  $L^u$  will lie in the interior of the first quadrant (in the coordinates near  $p'_g$ ), and we are done.

The only remaining case is when  $L^s$  and  $L^u$  form the boundary of the quadrant bounded by  $V_g^s(p_g)$  and  $V_g^u(p_g)$ . But this is only possible if  $p_g$  and  $p'_g$  form a heteroclinic cycle of length 2. Moreover each of the pairs  $V_g^s(p_g), V_g^u(p'_g)$  and  $V_g^u(p_g), V_g^s(p'_g)$  will either coincide or cross topologically. It follows from Theorem 2.2 that  $g$  has positive topological entropy or a double heteroclinic connection between  $p_g$  and  $p'_g$ .  $\square$

#### 4. Topological Crossing of Higher Dimensional Stable and Unstable Manifolds

Suppose that  $Y$  and  $Z$  are compact oriented manifolds with boundary, smoothly embedded in an oriented manifold  $X$ . Suppose further that  $\partial Y \cap Z = \emptyset = \partial Z \cap Y$  and  $\dim Y + \dim Z = \dim X$ . We wish to define the (oriented) **intersection number of  $Y$  with  $Z$ ,  $\#(Y, Z)$** .

Let us recall the definition in the case when  $Y$  and  $Z$  are without boundary (see e.g. [Hi]). The orientations of  $X, Y$  and  $Z$  allow one to assign an intersection number of 1 or  $-1$  to each transversal intersection of  $Y$  and  $Z$ . If  $Y$  and  $Z$  are transverse, there are finitely many intersection points and  $\#(Y, Z)$  is the sum of the intersection numbers at all intersection points. It is easily shown that if  $Y'$  is homotopic to  $Y$ ,  $Z'$  is homotopic to  $Z$ ,  $Y$  is transverse to  $Z$ , and  $Y'$  is transverse to  $Z'$ , then  $\#(Y', Z') = \#(Y, Z)$ . One defines the intersection number in the general case by first performing a homotopy to make the submanifolds transversal, and then using the definition in the transversal case.

This procedure can be carried over to our situation, with one caveat. The invariance of intersection number under homotopy breaks down if the boundary of one of the submanifolds is allowed to intersect the other submanifold. One can, however, make the submanifolds transverse with a homotopy that moves points less than any prescribed distance. Thus we simply add to the above definition the provision that the homotopy used to make  $Y$  and  $Z$  transverse should move points by less than  $\varepsilon/2$ , where  $\varepsilon$  is chosen so that

$$(*) \quad 0 < \varepsilon < \min(\text{dist}(\partial Y, Z), \text{dist}(\partial Z, Y)).$$

Now suppose that  $W$  and  $W'$  are smoothly immersed oriented submanifolds of  $X$  with complementary dimension. It is natural to say that  $W$  and  $W'$  cross topologically if there are compact embedded submanifolds with boundary,  $V \subset W$  and  $V' \subset W'$ , such that  $\dim V = \dim W$  and  $\dim V' = \dim W'$ ,  $\partial V \cap V' = \emptyset = V \cap \partial V'$  and  $\#(V, V') \neq 0$ . Note that it would be no loss of generality to require  $V$  and  $V'$  to be connected, for if  $\#(V, V') \neq 0$  we can choose components  $V_0 \subset V$  and  $V'_0 \subset V'$  with  $\#(V_0, V'_0) \neq 0$ .

It is possible to generalize these ideas so that they can be applied to nonorientable manifolds. Suppose that  $Y$  and  $Z$  are compact orientable embedded submanifolds with boundary such that  $\dim Y + \dim Z = \dim X$ ; we no longer require  $X$  to be oriented or even orientable. Let  $U$  be an orientable open set containing  $Y \cap Z$  and  $\mathcal{O}$  a choice of orientations for  $Y, Z$  and  $U$ . Then we can define  $\#_\varepsilon(Y, Z; U)$  by following the above procedure (with the orientation of  $U$  replacing

that of  $X$ ) and requiring homotopies of  $Y$  and  $Z$  to move points by less than  $\varepsilon/2$  where, in addition to  $(*)$ ,  $\varepsilon$  also satisfies

$$\varepsilon < \min(\text{dist}(Y \setminus U, Z), \text{dist}(Z \setminus U, Y)).$$

This condition ensures that no intersection points appear outside  $U$ . We point out that in general both the sign and the absolute value of  $\#_{\mathcal{O}}(Y, Z; U)$  can change with  $\mathcal{O}$ .

**Definition 4.1.** *Let  $W$  and  $W'$  be smoothly immersed submanifolds with complementary dimension in a manifold  $M$ . Then  $W$  and  $W'$  **cross topologically** if there are an orientable open subset  $U$  of  $M$  and compact orientable embedded submanifolds with boundary,  $V \subset W$  and  $V' \subset W'$  such that*

- (1)  $\dim V = \dim W$  and  $\dim V' = \dim W'$ ,
- (2)  $\partial V \cap V' = \emptyset = V \cap \partial V'$ ,
- (3)  $V \cap V' \subset U$ ,

and there is a choice  $\mathcal{O}$  of orientations on  $V, V'$  and  $U$  such that  $\#_{\mathcal{O}}(V, V'; U) \neq 0$ .

It is convenient to give another version of this definition. Observe that if  $Y, Z$  and  $U$  are connected, then only the sign of  $\#_{\mathcal{O}}(Y, Z; U)$  can change with  $\mathcal{O}$ . Moreover if  $U'$  is another connected orientable open set containing  $Y \cap Z$ , and  $\mathcal{O}'$  is the choice of orientations for  $Y, Z$  and  $U'$  that is compatible with  $\mathcal{O}$ , then

$$\#_{\mathcal{O}}(Y, Z; U) = \#_{\mathcal{O}'}(Y, Z; U').$$

Thus if  $Y$  and  $Z$  are connected compact orientable embedded submanifolds with boundary that have complementary dimension, we can define

$$\#_+(Y, Z) = |\#_{\mathcal{O}}(Y, Z; U)|,$$

where  $U$  is any connected orientable open set containing  $Y \cap Z$  and  $\mathcal{O}$  is a choice of orientations for  $Y, Z$  and  $U$ . We hasten to point out that there *are* such open sets, for example the open  $\varepsilon$ -neighbourhood of  $Y$  for small enough  $\varepsilon$ .

Observe that  $V, V'$  and  $U$  can be chosen as in Definition 4.1 if and only if there are components  $V_0$  of  $V$  and  $V'_0$  of  $V'$  such that  $\#_+(V_0, V'_0) \neq 0$ . Thus Definition 4.1 is equivalent to:

**Definition 4.2.** *Let  $W$  and  $W'$  be smoothly immersed submanifolds with complementary dimension in a manifold  $M$ . Then  $W$  and  $W'$  **cross topologically** if there are compact connected orientable embedded submanifolds with boundary,  $V \subset W$  and  $V' \subset W'$ , such that  $\#_+(V, V') \neq 0$ .*

We call  $V$  and  $V'$  a **good pair** for  $W$  and  $W'$  if they satisfy the above properties. We are now ready to state the main theorem.

**Theorem 4.3.** *Let  $f : M \rightarrow M$  be a diffeomorphism with a hyperbolic fixed point  $p$ . Assume that  $W^s(p)$  and  $\widehat{W}^u(p)$  have a topological crossing. Then some power of  $f$  has the full two shift as a topological factor.*

*Remarks.* We emphasize that the topological crossing need not be transverse and could be of infinite order. It follows from Lemma 1.3 that  $h_{TOP}(f) > 0$ . As in the two dimensional case, we only require contraction of  $W^s(p)$  by  $f$  and of  $W^u(p)$  by  $f^{-1}$ , not exponential contraction. Gedeon, McCord and Mischaikow have recently proved a generalization of this result [GMM].

Before proving Theorem 4.3, we develop the topological machinery that will be needed. The ideas are analogous to those in Sect. 2, but we have to use homology in place of elementary two dimensional topology. Now  $N$  will be a subset of  $M$  that is homeomorphic to (and identified with)  $D^k(1) \times D^l(1)$  where  $k + l = \dim M$  and  $R = D^k(1) \times D^l(\rho)$  for some  $\rho \in (0, 1)$ .

We call a set  $V \subset N$  **essential** if  $inc_* : H_l(V, V \setminus R) \rightarrow H_l(N, N \setminus R)$  is nonzero. If  $S \subset N$ , the **interior of  $S$  as a subset of  $N$**  is

$$Int_N S = \{x \in S : B_\varepsilon(x) \cap N \subset S \text{ for some } \varepsilon > 0\}.$$

We shall call a set  $S \subset N$  a **horizontal strip** if  $S$  is connected,  $S \subset R$ , and  $S$  can be decomposed into two disjoint sets  $S^0$  and  $S^\partial$  such that:

- (1)  $\overline{S^0} \subset Int_N S$ ;
- (2)  $i = inc_* : H_l(N, N \setminus R) \rightarrow H_l(N, N \setminus S^0)$  is nonzero;

*Remark.* There are only two types of horizontal strip that arise in the results of this paper. The first are those described in Sect. 2, for which we can choose  $S^\partial$  to be a small neighbourhood in  $S$  of  $c_{lower} \cup c_{upper}$ . The second have the form  $D^k(1) \times Y$ , where  $Y \subset D^l(\rho)$ ; in this case we choose  $S^\partial$  to be  $D^k(1) \times Y^\partial$ , where  $Y^\partial$  is a small neighbourhood of  $\partial Y$  (see Lemma 4.7 below).

From (1) above and the excision theorem, we obtain:

**Lemma 4.4.** *The map  $i_X = inc_* : H_l(X \cap S, X \cap S^\partial) \rightarrow H_l(X, X \setminus S^0)$  is an isomorphism for every set  $X \subset N$ , in particular when  $X = N$ .*

Since  $H_l(N, N \setminus R)$  is one dimensional, we see from property (2) of a horizontal strip that there is a unique one dimensional subspace  $L(S) \subset H_l(S, S^\partial)$  such that

$$i(H_l(N, N \setminus R)) \subset i_N(L(S)).$$

Let  $n$  be a positive integer and  $S = S^0 \cup S^\partial$  a horizontal strip. We shall say that  $f^n$  **stretches  $S$  across  $R$**  if  $f^n S \subset N$ ,  $f^n S^\partial \subset N \setminus R$  and

$$f_*^n : L(S) \rightarrow H_l(N, N \setminus R) \text{ is nonzero.}$$

**Theorem 4.5.** *Suppose  $N$  contains two disjoint closed horizontal strips  $S_0$  and  $S_1$  that are stretched across  $R$  by  $f^{n_0}$  for some  $n_0 \geq 1$ . Then  $f^{n_0}$  has the full two shift as a topological factor.*

*Proof.* Except for the proof of the Key Lemma, the argument is identical to that in the two dimensional case. In the present context, the Key Lemma says:

*Let  $S = S^\partial \cup S^0$  be a horizontal strip and  $V$  an essential subset of  $N$ . Suppose that  $f^n$  stretches  $S$  across  $R$ . Then  $f^n(S \cap V)$  is essential.*

The proof of this claim is very similar to that of Lemma 3.4 in [E]. Consider the commuting diagram

$$\begin{array}{ccc} H_l(V \cap S, V \cap S^\partial) & \xrightarrow{j} & H_l(S, S^\partial) \\ f_*^n \downarrow & & f_*^n \downarrow \\ H_l(f^n(V \cap S), f^n(V \cap S) \setminus R) & \xrightarrow{k} & H_l(N, N \setminus R) \end{array}$$

in which  $j$  and  $k$  are induced by inclusions. We want to show that  $k$  is a nonzero map. This will be true if  $f_*^n \circ j$  is a nonzero map. Since  $f^n$  stretches  $S$  across  $R$ , it suffices to show that  $\text{image}(j) \cap L(S) \neq \emptyset$ . To that end, consider the commuting diagram

$$\begin{array}{ccccc} H_1(V, V \setminus R) & \xrightarrow{i_1} & H_1(N, N \setminus R) & \xrightarrow{i} & H_1(N, N \setminus S^0) \\ i_2 \downarrow & & & & \uparrow i_N \\ H_1(V, V \setminus S^0) & \xrightarrow{i_V^{-1}} & H_1(V \cap S, V \cap S^c) & \xrightarrow{j} & H_1(S, S^c) \end{array}$$

in which all maps are induced by inclusions except for the map  $i_V^{-1}$ , which is defined by Lemma 4.4. Since the diagram commutes, it suffices to show that  $\text{image}(i \circ i_1) \cap i_N(L(S)) \neq \emptyset$ . Observe that  $i_1$  is nonzero because  $V$  is essential,  $i$  is nonzero because  $S$  is a horizontal strip, and  $H_1(N, N \setminus R)$  is one dimensional. Hence  $i \circ i_1$  is nonzero. It now follows from the definition of  $L(S)$  that  $\text{image}(i \circ i_1) \cap i_N(L(S)) \neq \emptyset$ .

*Proof of Theorem 4.3.* As in the corresponding theorem for surfaces, we may assume that  $p$  is a fixed point. We use the notation and geometric structures introduced at the end of Sect. 1.

**Lemma 4.6.**  *$W^u(p) \cap \text{Int} N$  contains two compact embedded  $l$ -dimensional manifolds with boundary,  $V_0$  and  $V_1$ , such that*

- (1)  $V_0 \cap V_1 = \emptyset$ ;
- (2)  $\partial V_0 \cup \partial V_1 \subset N \setminus (D^k(1) \times \{0\})$ ;
- (3)  $\text{inc}_* : H_1(V_i, \partial V_i) \rightarrow H_1(N, N \setminus (D^k(1) \times \{0\}))$  is nonzero for  $i = 0, 1$ .

*Proof.* Let  $V_0 = \{0\} \times D^l(1/2)$ . Now we find  $V_1$ . Let  $V^s$  and  $V^u$  be a good pair for  $\widehat{W}^s(p)$  and  $\widehat{W}^u(p)$  according to Definition 4.2. Then  $f^n V^s$  and  $f^n V^u$  are a good pair for any  $n$ . Choose  $n$  large enough so that  $f^n V^s \subset \text{Int} N$ . Choose an open set  $X$  such that  $f^n V^s \subset X$ ,  $\overline{X} \subset \text{Int} N$  and  $\partial X$  is a smooth submanifold transverse to  $V^u$ . Then  $f^n V^u \cap \overline{X}$  is a smooth manifold with boundary. There is a component  $V_1$  of  $f^n V^u \cap \overline{X}$  such that  $f^n V^s$  and  $V_1$  are a good pair for  $\widehat{W}^s(p)$  and  $\widehat{W}^u(p)$ . It is easily seen that  $V_0$  and  $V_1$  have the desired properties.  $\square$

Choose  $\rho > 0$  such that  $\partial V_0 \cup \partial V_1 \subset N \setminus (D^k(1) \times D^l(\rho))$ , and set  $R = D^k(1) \times D^l(\rho)$ . Our choice of  $\rho$  makes the sets  $V_0$  and  $V_1$  that were defined in the previous lemma essential.

**Lemma 4.7.** *For any large enough  $n$  there are disjoint closed horizontal strips  $S_0(n) \supset f^{-n} V_0$  and  $S_1(n) \supset f^{-n} V_1$  that are stretched across  $R$  by  $f^n$ .*

*Proof.* Choose  $\varepsilon > 0$  so that the closed  $\varepsilon$ -neighbourhoods of  $\partial V_0$  and  $\partial V_1$  lie in  $\text{Int} N \setminus R$ . Let  $n$  be large enough so that the conclusions of Lemma 1.4 hold with  $V = V_0$  or  $V_1$ . For  $i = 0, 1$ , let

$$Y_i(n) = \{y \in D^l(\rho) : (0, y) \in f^{-n}(V_i)\} \quad \text{and} \quad S_i(n) = D^k(1) \times Y_i(n).$$

Our choice of  $\varepsilon$  and (2) of Lemma 1.4 ensure that we can split  $Y_i(n)$  into complementary subsets  $Y_i^c(n)$  and  $Y_i^0(n)$  such that  $\overline{Y_i^0} \subset \text{Int} Y_i(n)$  and  $f^n$  maps  $D^k(1) \times Y_i^c(n)$  into  $N \setminus R$ . Set  $S_i^c(n) = D^k(1) \times Y_i^c(n)$  and  $S_i^0(n) = D^k(1) \times Y_i^0(n)$ . It is clear from our construction that  $S_0(n)$  and  $S_1(n)$  are horizontal strips and are



## References

- [A] Arnold, V.: Instability of dynamical systems with several degrees of freedom. *Soviet Math. Dokl.* **5**, 581–585 (1964)
- [Bo] Bowen, R.: *On Axiom A Diffeomorphisms*. CBMS Lecture Series #35, American Math. Society, 1977
- [Ch] Churchill, R.C.: Invariant sets which carry cohomology. *J. Diff. Equations* **13**, 39–65 (1973)
- [Co] Conley, C.: *Topological Dynamics*, Joseph Auslander, ed., New York: Benjamin, 1968, pp. 129–153
- [CR1] Churchill, R.C., Rod, D.L.: Pathology in dynamical systems I: General theory. *J. Diff. Equations* **21**, 39–65 (1976)
- [CR2] Churchill, R.C., Rod, D.L.: Pathology in dynamical systems. III—analytic Hamiltonians. *J. Diff. Equations* **37**, 23–38 (1980)
- [E] Easton, R.W.: Isolating blocks and symbolic dynamics. *J. Diff. Equations* **17**, 23–38 (1975)
- [Ga] Gantmacher, F.R.: *The Theory of Matrices*. New York: Chelsea, 1959
- [GMM] Gedeon, T., McCord, C., Mischaikow, K.: *Electronic communication*
- [GŠ] Gavrilov, N., Šilnikov, L.: On 3-dimensional dynamical systems close to systems with a structurally unstable homoclinic curve, II. *Math. USSR Sb.* **90**, 139–156 (1973)
- [Ha] Hartman, P.: On local homeomorphisms of Euclidean spaces. *Bol. Soc. Mat. Mexicana* **5**, 220–241 (1960)
- [Hi] Hirsch, M.W.: *Differential Topology*. Graduate Texts in Mathematics **33**, New York-Heidelberg-Berlin: Springer, 1976
- [HPS] Hirsch, M., Pugh, C., Shub M.: *Invariant Manifolds*. Springer Lecture Notes in Mathematics, vol. **583**, New York-Heidelberg-Berlin: Springer Verlag, 1977
- [K1] Katok, A.: Lyapunov exponents, entropy and periodic orbits for diffeomorphisms. *Publ. Math. Hautes Études Sci* **51**, 137–173 (1980)
- [K2] Katok, A.: Nonuniform hyperbolicity and structure of smooth dynamical systems. *Proc. International Congress of Mathematicians Warszawa 1983*, **2**, pp. 1245–1254
- [Ko] Kozlov, V.V.: Integrability and non-integrability in Hamiltonian mechanics. *Russ. Math. Surveys* **38**, 1–76 (1983)
- [KW] Knieper, G., Weiss, H.: A surface with positive curvature and positive topological entropy. *J. Differential Geometry* **39**, 229–249 (1994)
- [McMe] McGehee, R., Meyer, K.: Homoclinic points of area preserving diffeomorphisms. *American J. Math.* **96**, 409–421 (1974)
- [MiMr] Mischaikow, K., Mrozek, M.: *Isolating neighbourhoods and chaos*. Preprint (1993)
- [M] Melnikov, V.K.: On the stability of the center for time periodic perturbations. *Trans. Moscow Math. Society* **12**, 1–57 (1963)
- [N] Newhouse, S.: *Dynamical System*. C.I.M.E. Lectures, Boston-Basel-Stuttgart: Birkhäuser, 1978, pp. 1–114
- [Pa] Paternain, G.: Real analytic convex surfaces with positive topological entropy and rigid body dynamics. Preprint (1991)
- [Po] Poincaré, H.: *Les Méthodes Nouvelles de la Mécanique Céleste*. New York: Dover Publications (reprint), 1957
- [S] Smale, S.: Diffeomorphisms with many periodic points. *Differential and Combinatorial Topology*. Cairnes, S.S. (ed.) (1965), Princeton, NJ: Princeton University Press, pp. 63–80
- [So] Sotomayor, G.: Structural stability of first order and Banach varieties. *Univ. Nac. Ingen. Inst. Mat. Puras Apl. Notas Mat.* **4**, 11–52 (1966)
- [Sp] Spanier, E.H.: *Algebraic Topology*. New York: McGraw-Hill, 1966
- [St] Sternberg, S.: On the structure of local homeomorphisms of Euclidean  $n$ -space, II. *Amer. J. Math.* **80**, 623–631 (1958)
- [T] Tatarinov, Ya.V.: A global view of the dynamics of a rigid body: Passage to the quotient system. *Moscow Univ. Mech. Bull.* **33**, 17–21 (1978)
- [W] Walters, P.: *An Introduction to Ergodic Theory*. Graduate Texts in Mathematics **79**, Berlin, Heidelberg, New York: Springer 1982
- [Z] Ziglin, S.L.: Splitting of separatrices, branching of solutions and non-existence of an integral in the dynamics of the rigid body. *Trans. Moscow Math. Soc.* **41**, 283–298 (1982)