Commun. Math. Phys. 90, 433–459 (1983)

Communications in Mathematical Physics

Bifurcation to Infinitely Many Sinks

Clark Robinson*

Mathematics Department, Northwestern University, Evanston, IL 60201, USA

Abstract. This paper considers one parameter families of diffeomorphisms $\{F_t\}$ in two dimensions which have a curve of dissipative saddle periodic points P_t, i.e. $F_t^n(P_t)=P_t$ and $|\det DF_t^n(P_t)|<1$. The family is also assumed to create new homoclinic intersections of the stable and unstable manifolds of P_t as the parameter varies through t_0. Gavrilov and Silnikov proved that if the new homoclinic intersections are created nondegenerately at t_0, then there is an infinite cascade of periodic sinks, i.e. there are parameter values t_n accumulating at t_0 for which there is a sink of period n [GS2, Sect. 4]. We show that this result is true for real analytic diffeomorphisms even if the homoclinic intersection is created degenerately. We give computer evidence to show that this latter result is probably applicable to the Hénon map for A near 1.392 and B equal -0.3.

Newhouse proved a related result which showed the existence of infinitely many periodic sinks for a single diffeomorphism which is a perturbation of a diffeomorphism with a nondegenerate homoclinic tangency. We give the main geometric ideas of the proof of this theorem. We also give a variation of a key lemma to show that the result is true for a fixed one parameter family which creates a nondegenerate tangency. Thus under the nondegeneracy assumption, not only is there a cascade of sinks proved by Gavrilov and Silnikov, but also a single parameter value t^* with infinitely many sinks.

1. Introduction

The existence of a cascade of sinks is important because it analyzes a sequence of bifurcation which is different than period doubling. The existence of infinitely many sinks in Theorem C shows that there are generic situations which often arise where points tend to infinitely many distinct attractors. It indicates that for certain parameter values near $A=1.392$ the Hénon map does not have a transitive strange attractor but actually many different periodic sinks with narrow basins of attraction. (See Example 2.4 below.)

* Partially supported by National Science Foundation Grant MSC 80-02177

These results for diffeomorphisms have similarities with bifurcations of maps of the intervals and also differences. A family of maps of the interval $\{f_t\}$ nondegenerately creates homoclinic intersections if there is an unstable periodic point P_t and a nondegenerate critical point Q_t in the unstable manifold of P_t, such that Q_t is mapped back to P_t by some power of f_t (Q_t is an eventually periodic critical point). If the family $\{f_t\}$ nondegenerately creates homoclinic intersections, then there is an infinite cascade of sinks as in Theorem A below, but there is no one parameter value with infinitely many sinks as in Theorem C. See Remark 8.3 below and [V] for further discussion of this comparison.

Newhouse originally proved there exists a specific example where there is a residual subset $\mathscr{R}$ of an open set of diffeomorphisms $\mathscr{N}$ such that each G in $\mathscr{R}$ has infinitely many sinks, [N2]. Later he proved this occurred near any dissipative diffeomorphism F which had a nondegenerate tangency of stable and unstable manifolds, [N4]. This later result follows from first proving that there is an open set of diffeomorphism in the C^2 topology $\mathscr{N}$, which is C^2 near F such that for each G in $\mathscr{N}$ there is a hyperbolic basic set $\Lambda = \Lambda(G)$ which has $W^u(\Lambda)$ having a nondegenerate tangency with $W^s(\Lambda)$. He also proved, [N4, Theorem 3], that for a fixed one parameter family of diffeomorphisms $\{F_t\}$ which creates a nondegenerate tangency of stable and unstable manifolds of a n periodic point P_t at $t=t_0$, where $|\det DF^n(P_t)|<1$, then for any $\varepsilon>0$ there is an interval $[t_1, t_2] \subset [t_0-\varepsilon, t_0+\varepsilon]$ such that for t in $[t_1, t_2]$ F_t has a nondegenerate tangency of the stable and unstable manifolds of a hyperbolic basic set Λ_t containing P_t. He has stated in talks and implies in [N4, Remark 1, p. 105], but does not state explicitly in his papers, that there is a residual subset J in $[t_1, t_2]$ such that for t in J the fixed one parameter family F_t has infinitely many sinks, [N4, Remark 1, p. 105]. Theorem E below states and proves that this is indeed true. See also [N5, Theorem 8.1], and [GH]. Earlier Garilov and Silnikov had proved that if a C^3 family nondegenerately creates the homoclinic intersection then there is an infinite cascade of sinks as in Theorem A below, [GS2, Sect. 4]. They also showed that hyperbolic invariant sets were created which are not conjugate to each other, so the system is inaccessible by a simple bifurcation from at least one side, [GS] or [GH].

This paper proves a key result, Proposition 3.3, that whenever there is a one parameter family $\{F_t\}$ and a box B such that the images of the box $F_t(B)$ are pulled across B in the shape of a horseshoe, then there is an interval of parameter values J such that for t in J, F_t has at least one periodic sink. Using this result it can be shown that a fixed one parameter family of real analytic diffeomorphism which creates homoclinic intersections (as defined in Sect. 2) has an infinite cascade of periodic sinks, Theorem A. Therefore there is a sequence of periodic sinks p_n for different parameter values t_n, but not necessarily infinitely many sinks for one parameter value t^*. This theorem has fairly weak hypothesis. It is almost certainly applicable to the Hénon map to show there is an infinite cascade of sinks for $B=-0.3$ and A near 1.39. See Example 2.4. Using the stronger hypothesis used in [N4] that the family nondegenerately creates homoclinic intersections, Proposition 3.3 is the only new ingredient needed to show that the one parameter family can be fixed and prove there are many parameter values which have infinitely many sinks, Theorem C. Most of the work goes into proving that if $\{F_t\}$

nondegenerately creates homoclinic intersections, then there is a persistent tangency of stable and unstable manifolds, Theorem D. This result is proved in [N4]. Since its proof is very involved, we present the main aspects of the proof with reference for the analytic details. (This result is stated in [N5] but those lectures do not discuss the proof.)

The outline of the paper is as follows. Section 2 states the main results and gives examples where the theorem applies. Section 3 states and proves the proposition on the creation of one sink. Section 4 proves that if a family of real analytic diffeomorphisms create intersections then they create odd order intersections. Section 5 proves the result on the existence of an infinite cascade of sinks. Section 6 discusses the persistence of intersection of "thick" Cantor sets. Section 7 shows how thick Cantor sets of stable manifolds arise from the nondegenerate creation of homoclinic tangency. Section 8 shows how the parameter value can be chosen to get both a thick Cantor set of stable manifolds and a nondegenerate tangency – hence a persistent tangency. Finally Sect. 9 proves that the persistence of nondegenerate tangency leads to infinitely many sinks.

2. Statement of Main Theorem and Applications

For a diffeomorphism F let $DF(p)$ denote the derivative of F at p, i.e. the matrix of partial derivatives. A point p is called a *periodic sink* (respectively *source*) for a diffeomorphism F if p is a periodic point, $F^n(p)=p$, such that all eigenvalues of the derivative of F^n, $DF^n(p)$, have absolute value less than one (respectively all have absolute value greater than one). Thus a sink is a periodic attractor and there is a neighborhood U of the orbit $\mathcal{O}(p)=\{F^j(p): j \text{ in } Z\}$ such that $\bigcap_{j\geqq 0} F^j(U)=\mathcal{O}(p)$. A point p is called a *periodic saddle point* for a diffeomorphism F in two dimensions if it is a periodic point with $F^n(p)=p$ for some n and the eigenvalues of $DF^n(p)$ are λ_u and λ_s, both real, with $|\lambda_u|>1$ and $|\lambda_s|<1$. The *stable and unstable manifolds* of a saddle point for a C^r diffeomorphism F are then C^r curves tangent to the eigendirections defined by $W^s(p,F)=\{q$: distance $F^j(q)$ to $F^j(p)$ goes to zero as j goes to infinity$\}$ and $W^u(p,F)=\{q$: distance $F^j(q)$ to $F^j(p)$ goes to zero as j goes to minus infinity$\}$.

We need to distinguish the direction of crossing of two oriented curves and the order of tangency. Let γ^1 and γ^2 be two oriented differentiable curves. We say that γ^1 has *positive* (respectively negative) *intersection* with γ^2 at a point p if p is on both curves and there are local coordinates (x,y) near p with $x(p)=y(p)=0$ and in which there are oriented parametrizations $\{(x_j(s),y_j(s)): |s|<\varepsilon\}$ of γ^j with (i) $y_2(s)\equiv 0$, $x_2(0)=0$, and $x_2'(0)>0$ and (ii) $y_1(0)=0=x_1(0)$, $y_1(s)<0$ for $-\varepsilon<s<0$ (respectively $0<s<\varepsilon$), and $y_1(s)>0$ for $0<s<\varepsilon$ (respectively $-\varepsilon<s<0$). We say that $\{\gamma^1,\gamma^2\}$ have *intersection of order* $n+1$ at p (or *tangency of order* n) if there are parametrizations as above with (i) $y_2(s)\equiv 0$ and (ii) $y_1(0), y_1'(0)=0,\dots,y_1^{(n)}(0)=0$, but $y^{(n+1)}(0)\neq 0$. Note that γ^1 and γ^2 are transverse at p if and only if they have intersection of order one if only if they have tangency of order zero.

We say that F_t *creates* (respectively destroys) *homoclinic intersections at* t_0 for a periodic saddle point P_t if there are $\varepsilon>0$, $Q_t=F_t^k(P_t)$ for some k, and continuously varying subarcs $\gamma_t^s \subset W^s(P_t, F_t)$ and $\gamma_t^u \subset W^u(Q_t, F_t)$ for $t_0 - \varepsilon \leqq t \leqq t_0 + \varepsilon$ such that

(i) $\gamma_t^s \cap \gamma_t^u = \emptyset$ for $t_0 - \varepsilon \leqq t < t_0$ (respectively, $t_0 < t \leqq t_0 + \varepsilon$),

(ii) for $t_0 < t \leqq t_0 + \varepsilon$ (respectively, $t_0 - \varepsilon \leqq t < t_0$) γ_t^s has both positive and negative intersections with γ_t^u.

We say that $\{F_t\}$ *creates odd order homoclinic intersections* at t_0 if condition (i) is satisfied and

(ii)′ for $t_0 < t \leqq t_0 + \varepsilon$ γ_t^u has at least one positive intersection with γ_t^s of odd order and at least one negative intersection with γ_t^s of odd order.

In condition (ii)′ there can be more than two intersections some of which are of even order. In the proof of Theorem A below show that if $\{F_t\}$ is a family of analytic diffeomorphisms which depend continuously on t and satisfying (ii)′ then it satisfies condition (ii). Also in the proof of Theorem B condition (ii) is not needed for all $t_0 < t \leqq t_0 + \varepsilon$ but only for a sequence of parameter values $t_j > t_0$ which accumulate on t_0. The definition of (ii)′ could be weakened accordingly.

The above condition is sufficient to prove there is a cascade of sinks, but to prove that there are infinitely many sinks for a single parameter value another condition is needed. We say that $\{F_t\}$ *nondegenerately creates homoclinic intersections at* t_0 if $\{F_t\}$ creates odd order homoclinic intersections [conditions (i) and (ii)′] and

(iii) $\gamma_{t_0}^s$ and $\gamma_{t_0}^u$ have intersection of order two (tangency of order one),

(iv) If coordinates are taken so γ_t^s lies on $y=0$ and $y^*(t)$ is the extreme value of y along γ_t^u, then $dy^*/dt \neq 0$ at $t = t_0$.

Note in this case, the intersections for $t > t_0$ are necessarily transverse, i.e. of order one. If $\{F_t\}$ nondegenerately creates homoclinic intersections, Newhouse uses the terminology that it *creates a nondegenerate tangency at* t_0. His terminology emphasizes the tangency in condition (iii), while ours emphasizes the topologically transverse intersections in condition (ii).

Theorem A. *Let $\{F_t\}$ be a one parameter family of real analytic diffeomorphisms in two dimensions which depend continuously on t. Assume it creates (or destroys) a homoclinic intersection at t_0 for the periodic points P_t of periodic n with $|\det DF_{t_0}^n(P_{t_0})| < 1$. Then F_t has an infinite cascade of sinks. More specifically there is a sequence of parameters values t_j converging monotonically to t_0 such that F_{t_j} has a periodic sink of period n_j. The orbits of the sinks pass near the point of tangency of $W^s(P_{t_0}, F_{t_0})$ and $W^u(F^k(P_{t_0}), F_{t_0})$. The periods n_j of the sinks grow like $n_{j+1} - n_j = n$ or $2n$ depending on whether F_t preserves the orientations on $W^s(P_t, F_t)$ and $W^u(P_t, F_t)$ or not.*

The result for real analytic diffeomorphisms follows quite directly from the following result about C^j diffeomorphisms.

Theorem B. *Assume $\{F_t\}$ is a one parameter family of C^j diffeomorphisms in Theorem A but assume it creates (or destroys) odd order homoclinic intersections at t_0 of oder j. Then the conclusion of Theorem A is true. Here $j \geqq 1$.*

2.1. Remark. Curry and Johnson, [CJ], calculated by means of a computer the asymptotic rate of the creation of sinks in the cascade for the family of maps studied in [ACHM]. They noted that

$$\lim_{n\to\infty} (t_n - t_{n-1})/(t_{n+1} - t_n) = \lambda,$$

where λ_u is the unstable eigenvalue of the saddle fixed point. They also include a proof with details to be supplied elsewhere. As noted above Gavrilov and Silnikov proved Theorem A under the assumption that a C^3 family nondegenerately creates homoclinic intersections. They also proved bounds on the parameter values for the existence of the sink of period n which imply the asymptotic rate of creation of sinks noted by Curry and Johnson, [GS2, 4.4 and 4.5]. At the end of Sect. 5, we indicate how this asymptotic rate is related to the proof of Theorem B.

The following theorems are essentially the results of Newhouse. In particular Theorem D is [N2, Theorem 2].

Theorem C. *Suppose $\{F_t\}$ is a fixed one parameter family of C^3 diffeomorphisms of a two manifold which nondegenerately creates homoclinic intersections at t_0 for the periodic points P_t of period n with $|\det DF^n_{t_0}(P_{t_0})|<1$ (respectively $|\det DF^n_{t_0}(P_{t_0})|>1$). Then given $\varepsilon>0$ there is a subinterval $[t_1, t_2] \subset [t_0-\varepsilon, t_0+\varepsilon]$ and a residual subset $J \subset [t_1, t_2]$ such that for t in J, F_t has infinitely many sinks (respectively sources).*

In Theorem C we assume $\{F_t\}$ is a C^1 curve of C^3 diffeomorphisms, i.e. the third derivative of F with respect to q in M has one continuous derivative with respect to the parameter t. F_t is assumed C^3 in order to almost C^2 linearize near P_t. See Sect. 7 for more details.

2.2. Remark. Theorem C can not be proved by showing the intervals of parameter values with sinks given in Theorem A overlap. In fact, Remark 5.2 indicates why no two of the sinks of Theorem A occur for the same parameter value. In terms of bifurcation subsets of the function space of diffeomorphisms, Theorem C, or more precisely [N4, Theorem 1], means that there is an open set of C^2 diffeomorphisms, $\mathcal{N}$, such that $\Sigma_1 = \{G \in \mathcal{N} : G$ has a generic saddle node$\}$ is dense in $\mathcal{N}$. Such bifurcations are codimension one, for each G in Σ_1 there is a codimension one submanifold $\Sigma_1(G)$ in $\mathcal{N}$ such that each H in $\Sigma_1(G)$ has a saddle node bifurcation. A generic arc crosses these bifurcations transversally. Theorem E does not prove there are transverse crossings but does prove there are infinitely many sinks. If the periodic point P_t has eigenvalues that are independent enough to C^2 linearize near P_t, then it appears that the bifurcations are actually generic saddle node bifurcations. This result would use a lemma like [N2, Lemma 2] and is not included in this paper.

Theorem C follows from Theorems D and E below. To state these results we need further definitions. See [N5] for more precise statements and examples. The creation of homoclinic intersections implies the creation of Smale horseshoes. In fact there are integers n, boxes B_n, and parameter values $t=t_n$ such that $F^n_t(B_n)$ crosses B_n in the shape of a horseshoe, see Fig. 5b below. Letting $G=F^n_t$ and $B=B_n$, the set $\Lambda = \{G^k(B): k$ is in $Z\}$ is the *maximal invariant set* for G in B, i.e. Λ is the set of all points q such that both the forward and backward orbit of q by G stays in B. Each point q in Λ also has a contracting (stable) direction E^s_q and an expanding (unstable) direction E^u_q much like the eigendirections at a saddle fixed point. More precisely, a closed invariant set Λ for G is said to have a *hyperbolic structure* if the tangent space of the ambient manifolds has a splitting at points q of Λ, $T_qM = E^s_q + E^u_q$, where the splitting varies continuously with q, and if there are constants $C>0$ and $\lambda>1$ such that for $k \geqq 0$ and for v^s in E^s_q, $|DG^k(q)v^s| \leqq C\lambda^{-k}|v^s|$, and for v^u in E^u_q, $|DG^{-k}(q)v^u| \leqq C\lambda^{-k}|v^u|$. If Λ has a hyperbolic structure for G, then the nonlinear map G has a family of invariant nonlinear manifolds tangent to the linear directions E^s_q and E^u_q which are contracted and expanded respectively by G. More precisely, for each point q in Λ, the *stable manifold of q for G* is the set $W^s(q,G)=\{m$: distance $G^k(q)$ to $G^k(m)$ goes to zero as k goes to infinity$\}$. The *local stable manifold of q of size* $\varepsilon>0$ is the set $W^s_\varepsilon(q,G)=\{m$ in $W^s(q,G)$: distance $G^k(q)$ to $G^k(m)$ is less than ε, for all $k\geqq 0\}$. Thus $W^s(q,G)=\bigcup\{G^{-j}W^s_\varepsilon(G^j(q),G): j\geqq 0\}$. If Λ has a hyperbolic structure for a C^r diffeomorphism G, then (i) $W^s_\varepsilon(q,G)$ is a C^r differentiable disk with dimension equal to $\dim E^s_q$, (ii) the disks $W^s_\varepsilon(q,G)$ vary continuously in the C^r topology as q varies in Λ, and (iii) the disks are invariant, $G(W^s_\varepsilon(q,G))\subset W^s_\varepsilon(G(q),G)$. The (global) stable manifold is an immersed C^r differentiable manifold. Similarly the *unstable manifold of q for G* is the set $W^u(q,G)=\{m$: distance $G^k(q)$ to $G^k(m)$ goes to zero as k goes to minus infinity$\}$ and the *local unstable manifold of size* ε is $W^u_\varepsilon(q,G)=\{m$ in $W^u(q,G)$: distance $G^k(q)$ to $G^k(m)$ is less than ε for all $k\leqq 0\}$. Again $W^u(q,G)=\bigcup\{G^j(W^u_\varepsilon(G^{-j}(q),G)): j\geqq 0\}$. A closed set Λ is called a *hyperbolic basic set* for G if (i) it is invariant for G, $G(\Lambda)=\Lambda$, (ii) it has a hyperbolic structure for G, (iii) there is a point q in Λ with a dense orbit, closure$(\mathcal{O}(q))=\Lambda$, and (iv) Λ has a local product structure, i.e. if $\varepsilon>0$ is sufficiently small and p, q are in Λ then $W^u_\varepsilon(p,G)\cap W^s_\varepsilon(q,G)\subset\Lambda$.

If Λ is a hyperbolic basic set then there is a neighborhood U of Λ such that $\bigcap_{j=-\infty}^{\infty} F^j(U)=\Lambda$. For G which is C^1 near F, $\Lambda(G)=\bigcap_j G^j(U)$ is a hyperbolic basic set for G. A hyperbolic basic set Λ is called a *wild hyperbolic set for F* (or has persistent tangencies of stable and unstable manifolds) if for any G which is C^2 near F there are points q_1 and q_2 in $\Lambda(G)$ for which $W^s(q_2,G)$ has a nondegenerate tangency with $W^u(q_1,G)$. (See [N5] for further definitions and more precise statements.)

Theorem C follows from the following two theorems.

Theorem D. *Suppose $\{F_t\}$ nondegenerately creates homoclinic intersections at t_0 for the curve of periodic points P_t with $|\det DF^n_{t_0}(P_{t_0})|\neq 1$ and each F_t is C^3. Then given $\varepsilon>0$ there is a subinterval $[t_1,t_2]\subset[t_0-\varepsilon,t_0+\varepsilon]$ such that for t in $[t_1,t_2]$, F_t has a wild hyperbolic set containing the corresponding periodic point P_t of period n.*

Theorem E. *Assume $\{F_t\}$ has an interval of parameter values $[t_1,t_2]$ such that for t in $[t_1,t_2]$, F_t has a wild hyperbolic set containing a periodic point P_t with $|\det DF^n_t(P_t)|<1$ (respectively >1). Then there is a residual subset $J\subset[t_1,t_2]$ such that for t in J, F_t has infinitely many sinks (respectively sources).*

2.3. Remark. If F_t are area preserving then Theorem D is unknown. It can be shown that F_t goes through a cascade of bifurcations producing elliptic points as in Sect. 5 below. See [N3].

2.4. Example. Recently there has been much interest in the Hénon map

$$F_{AB}(x,y)=(A-By-x^2,x).$$

For $B=0$, the map is like the graph of a parabola map of the interval. For certain values of A the end of the parabola gets mapped to the line $x=x_0$ containing one of the fixed points. This line is in the stable manifold of the fixed point. Thus as A

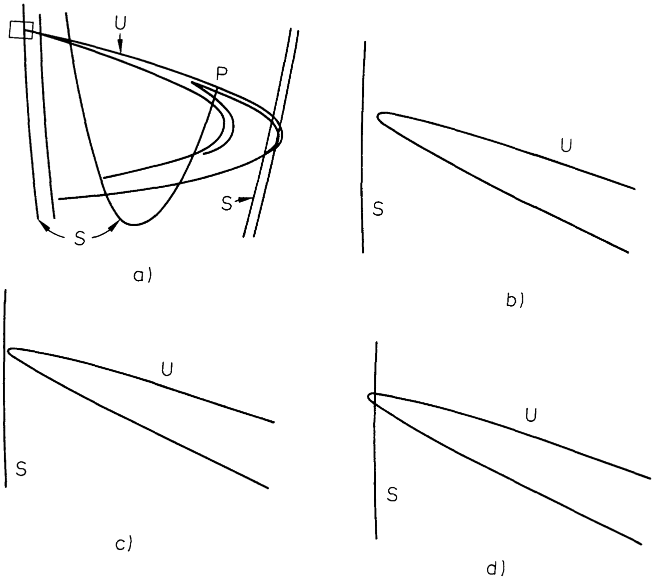


Fig. 1a–d. Hénon map. The stable and unstable manifolds of the fixed point P are labelled by S and U, respectively. The small box in **a** is enlarged in **b–d** for different values of A. $B=-0.3$. **a** $A=1.39$, **b** $A=1.39$, **c** $A=1.392$, **d** $A=1.395$

varies it creates a nondegenerate tangency. This holds for $|B|$ small enough. Therefore for small B there are values of A with infinitely many sinks. See [V, Theorem D] for details.

Earlier, [N4, Remark 1, p. 105] had indicated that for some parameter values the Hénon map has infinitely many sinks. It does not specify the values of A and B for which this is true, but oral communication indicated it is for this case with $|B|$ small.

More interestingly, for $B=-0.3$ and as A varies from 1.39 to 1.4, computer studies indicate F_A creates homoclinic intersections. See Fig. 1. This fact is probably verifiable either via more careful computer studies or analytically. Because everything is analytic, if $\{F_A\}$ does create homoclinic intersections then there is an infinite cascade of sinks. Therefore the computer studies strongly indicate that there is an infinite cascade of sinks.

Further the homoclinic tangency appears nondegenerate in the computer studies: the stable manifold has small curvature while the curvature of the unstable manifold is very large (it appears as a sharp point). See Fig. 1. Therefore it is indicated, but is unproved analytically or via computers, that there are values of A between 1.39 and 1.4 for $B=-0.3$ with infinitely many sinks. The nondegeneracy condition appears much more difficult to prove than the creation of homoclinic intersections.

The plots of iterates of a single point for A between 1.39 and 1.4 and $B=-0.3$ all appear like attractors. One aspect of the explanation of why the sinks are not visible is that the basis of attraction of the sinks are very narrow. Further theoretical and numerical explanation is needed.

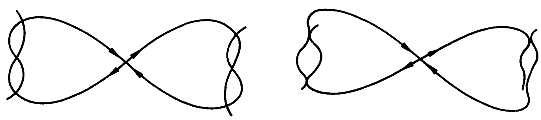


Fig. 2a and b. Forced Duffing equation with damping. **a** $\delta=0$, $\varepsilon>0$; **b** $\delta=\delta_0>0$, $\varepsilon>0$

These computer results contrast with the result of Misiurewicz, [Mis], for the Lozi map for which a hyperbolic attractor exists. The difference is that the smooth bend in the Hénon map causes a saddle node to be created as the image of the box is pulled across itself. The "piecewise linear" character of the Lozi map avoids this and immediately creates two saddle points.

The paper by Aronson et al. [ACHM], contains computer studies of an equation which models delayed regulation of population growth. Their studies like those for the Hénon map indicate the creation of odd order homoclinic intersections.

2.5. Example. Another type of example that has been studied is forced oscillators. Consider the forced Duffing equation with damping

$$\dot{x}=v\,,$$
$$\dot{v}=\beta x-\alpha x^3+\varepsilon(\gamma\cos\omega t-\delta v)\,.$$

For $\delta=0$, and $0<\varepsilon<\varepsilon_0$ there is a transverse homoclinic orbit for the time 2π map. This can be seen by using the Melnikov integral, [HM, Sect. 4]. As δ increases, the dissipation pulls the unstable manifold inward creating a nondegenerate tangency at $\delta=\delta_0$. (This corresponds to $\gamma=\gamma_0$ in [HM].) Thus for δ near δ_0 there are infinitely many sinks.

2.6. Example. Levi considered a forced van der Pol equation of the type studied by Levinson

$$\varepsilon\ddot{x}+\phi(x)\dot{x}+\varepsilon x=bp(t)\,,$$

where $\phi(x)$ and $p(t)$ are rounded off square periodic functions. He showed there were hyperbolic basic sets for certain values of b. Moreover he showed for other values of b, there are nondegenerate homoclinic tangency and so infinitely many sinks [L, p. 33 and Sect. 3.6].

3. Creation of One Sink

3.1. Example. Before considering the creation of sinks in two dimensions, consider a map of the real line $f_t:\mathbb{R}\to\mathbb{R}$ with $\partial^2 f_t/\partial x^2\geqq a>0$, $\partial f_t/\partial t<0$; and $f_0(x)>x$, e.g. $f_t(x)=x^2-t+2$. As t increases there is a first parameter value t_1 where $f_{t_1}(x_1)=x_1$. At this point $f'(x_1)=1$.

Then for $t_1<t<t_1+\varepsilon$, $f_t(x)=x$ has two solutions $x_t<x_t'$, and $0<f_t'(x_t)<1<f_t'(x_t')$. Thus x_t is a sink for f_t (and x_t' is a source).

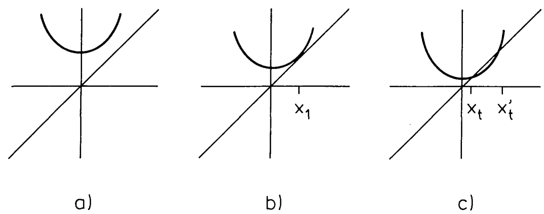


Fig. 3a–c. Graph of $y=f(x)$ with diagonal $y=x$. (**a**) $t<t_1$; (**b**) $t=t_1$; (**c**) $t>t_1$

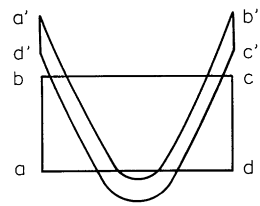


Fig. 4. The box $-2\leqq x\leqq 2$, $-2\leqq y\leqq 2$ with vertices a, b, c, d. The primed letters label the vertices which are the images of the vertices labelled by the corresponding unprimed letters, e.g., $a'=F_A(a)$

3.2. Example. Next consider a Hénon type map

$$F_A(x,y)=(y,\,-A-Bx+y^2)$$

with $0<B<1$ fixed, e.g. $B=0.3$. The fixed points (for appropriate A) are

$$x_1=y_1=\tfrac{1}{2}(1+B)-\tfrac{1}{2}[(1+B)^2+4A]^{1/2},$$
$$x_2=y_2=\tfrac{1}{2}(1+B)+\tfrac{1}{2}[(1+B)^2+4A]^{1/2}.$$

The eigenvalues at the fixed point (x,y) are given by $\lambda=y\pm[y^2-B]^{1/2}$. Thus the fixed point (x_1,y_1) has the following type:

$A<A_0\equiv-\tfrac{1}{4}(1+B)^2$,	there are no fixed points,
$A=A_0$,	saddle node, $\lambda=1, B$,
$A_0<A\leqq A_1$,	sink with real positive eigenvalues,
$A_1<A<A_2$,	sink with complex eigenvalues,
$A_2<A<A_3\equiv\tfrac{3}{4}(1+B)^2$,	sink with real negative eigenvalues,
$A=A_3$,	eigenvalues $\lambda=-1, -\mathrm{B}$,
$A_3<A$,	hyperbolic with reflection $\lambda_-<-1<\lambda_+<0$.

As far as the uses for this paper, the important bifurcation occurs at $A=A_0$, where a saddle node bifurcation creates a sink for A slightly larger than A_0. The difference between maps of the real line and the plane is apparent with the eigenvalues becoming complex for $A_1<A<A_2$. □

The next proposition gives a general criterion for the creation of one sink. The basic hypothesis is that there is a box B such that the family $\{F_t\}$ pulls the image of B across B creating a saddle point whose expanding eigenvalue is negative. (There is a reflection in the unstable direction.) See Fig. 5. This result is related to [N5, Lemma 8.2] but the argument there seems to apply only to a generic one parameter family. Also see [N2, Lemma 2]. This proposition is the only modification of the proof in [N4] needed to prove Theorem E.

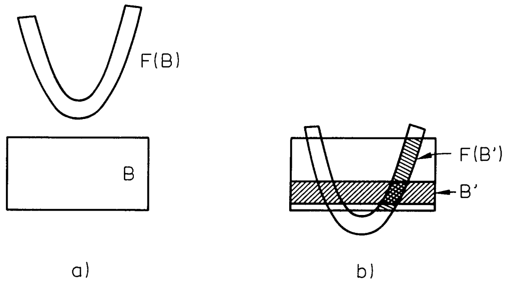


Fig. 5. **a** The image $F(B)$ for $F=F_{t_1}$ lies outside B; **b** The image $F(B)$ for $F=F_{t_2}$ crosses B at least twice. The saddle point with reflection is contained in $B' \cap F(B')$

3.3. Proposition. *Let $F_t(x,y)$ be a one parameter family of C^1 diffeomorphisms defined in a neighborhood of a box $B=\{(x,y): |x-x_0| \leqq \delta_1$ and $|y-y_0| \leqq \delta_2\}$ for $t_1 \leqq t \leqq t_2$. Assume $|\det DF_t(x,y)|<1$ for all $t_1 \leqq t \leqq t_2$ and (x,y) in B. Further assume*

(i) $F_{t_1} B \cap B = \emptyset$,

(ii) F_{t_2} *has a fixed point that is a saddle point with reflection,* i.e. *has real eigenvalues $\lambda \leqq -1$ and $|\lambda'|<1$,*

(iii) F_t *has no fixed points on the boundary of B for all $t_1 \leqq t \leqq t_2$.*

Then there is a t_0 with $t_1 < t_0 < t_2$ such that F_{t_0} has a fixed point sink. If $\det DF_t > 0$ *so F_t is orientation preserving, the value t_0 can be chosen so the eigenvalues have nonzero imaginary part.*

3.4. Remark. The way condition (iii) is satisfied in the applications of (3.3) is as follows. Assume $F_t(B)$ does not intersect the sides of the box, $\{(x_0 \pm \delta_1, y): |y-y_0| < \delta_2\}$ for $t_1 \leqq t \leqq t_2$, then there are no fixed points on the sides of B. If the top and bottom edges of B always have images outside the box, $F_t(x, y_0 \pm \delta_2)$ is not in B for $t_1 \leqq t \leqq t_2$, then there are no fixed points on the top and bottom of B. Thus with these two assumptions, there are no fixed points on the boundary of B.

Proof. The proof uses two key ideas: (1) eigenvalues vary continuously and (2) if F_t has a fixed point $p(t)$ whose eigenvalues are not equal to $+1$ then by the implicit function theorem there is an interval $(t-\varepsilon, t+\varepsilon)$ such that for s in this interval F_s has a fixed point $p(s)$. Let $p(t_2)$ be the fixed point given by assumption (ii) with eigenvalues $\lambda_1(t_2) \leqq -1$ and $|\lambda_2(t_2)|<1$. Let $t_3 < t \leqq t_2$ be the maximal interval of continuing the fixed point $p(t)$ in the interior of B. It follows that for $t=t_3$, F_t has a fixed point $p(t_3)$ because this is a closed condition. By assumption (iii), $p(t_3)$ can not be on the boundary of B, so is in its interior. By assumption (i), for $t=t_1$, F_t has no fixed points in B, so $t_1 < t_3$. By (2) above it follows that $p(t_3)$ has an eigenvalue equal to $+1$.

Either $\lambda_1(t_3)=1$ or $\lambda_2(t_3)=1$. If $\lambda_2(t_3)=1$, then at some intermediate value $t=t_4$, $|\det DF_t(p(t))| < |\lambda_2(t_4)| < 1$. Because $|\det DF_t(p(t))| = |\lambda_1(t_4)|\,|\lambda_2(t_4)|$, it follows that also $|\lambda_1(t_4)|<1$ and $p(t_4)$ is a sink. One the other hand if $\lambda_1(t_3)=1$, then at some parameter value t_5, $\operatorname{Re}\lambda_1(t_5)=0$. Since F_t is a diffeomorphism, $\lambda_1(t_5) \neq 0$ so it must be pure imaginary with $\lambda_2(t_5) = -\lambda_1(t_5)$. Then for $t=t_5$, $1 > |\det DF_t(p(t))| = |\lambda_2(t_5)|^2$ and $p(t_5)$ is a sink.

Note in the case where $\det DF_t(p) > 0$, the second case must occur so there must be a complex sink. □

4. Creation of Homoclinic Intersections for Real Analytic Maps

In this section we reduce the proof of Theorem A to Theorem B, i.e. we show that if $\{F_t\}$ is a real analytic family which creates homoclinic intersections, then it creates odd order homoclinic intersections.

Let P_t be the periodic points such that the stable and unstable manifolds of the orbit of P_t create homoclinic intersections. There is a k such that for $Q_t = F_t^k(P_t)$, the manifolds $W^u(Q_t, F_t)$ and $W^s(P_t, F_t)$ create homoclinic intersections. These manifolds are real analytic for fixed t because F_t is real analytic. These manifolds do not coincide for $t_0 < t \leqq t_0 + \varepsilon$, because they have two topologically transverse intersections, condition (ii). Because they are real analytic, all the intersections are of finite order. Any intersection of finite order which is topologically transverse is necessarily of odd order. Therefore by condition (ii) there are both positive and negative intersections of odd order.

5. Infinite Cascade of Sinks: Theorem B

For simplicity we assume the stable and unstable manifolds which create homoclinic intersections belong to a fixed point P_t. The case of a periodic point is not much different. See [N4]. We also assume both of the eigenvalues for P_t are positive. Again the case where one of these is negative can be handled by replacing F_t with F_t^2. Because $0 < \det DF_{t_0}(P_{t_0}) < 1$, the product of the eigenvalues is less than one, $0 < \mu\lambda < 1$, where $0 < \mu < 1$ and $1 < \lambda$. The tangencies also can occur from various sides and for different branches of $W^s(P_t, F_t)$ and $W^u(P_t, F_t)$. Some of these cases imply that $W^s(P_t, F_t)$ and $W^u(P_t, F_t)$ already have other intersections for $t = t_0$. See the work of Gavrilov and Silnikov, [GS] or [GH]. These differences make no difference in the proof given here. Compare the cases in Figs. 1 and 2 which are different.

For simplicity of discussion we assume that F_t can be linearized near P_t, i.e. there is a neighborhood U of P_t and C^j coordinates (x, y) on U so that $F_t(x, y) = (\mu x, \lambda y)$. Here j is the order of the intersection of stable and unstable manifolds. If F_t is C^∞ and satisfies nonresonance conditions of the eigenvalues at P_t, the Sternberg linearization gives such coordinates. In Appendix 5.4 at the end of the section, there is a discussion of how to obtain the necessary estimates without linearizing.

Let q_0 be the point at which $W^s(P_t, F_t)$ and $W^u(P_t, F_t)$ have a tangency at $t = t_0$. By looking along the orbit of q_0, we can assume $q_0 = (x_0, 0)$ is in U on the local stable manifold of P_{t_0}. Also by taking k large enough, $q_1 = F_{t_0}^{-k}(q_0) = (0, y_1)$ will also be in U.

We next form boxes B_n near q_0 to which Proposition 3.3 can be applied to get a sink. Take $\delta_1^u, \delta_1^s, \delta_0^u, \delta_0^s > 0$ and form

$$V_0 = \{(x, y) : |x - x_0| \leqq \delta_0^s,\ 0 \leqq (\operatorname{sign} y_1) y \leqq \delta_0^u\}$$

and

$$V_1 = \{(x, y) : |y - y_1| \leqq \delta_1^u,\ 0 \leqq (\operatorname{sign} x_0) x \leqq \delta_1^s\}.$$

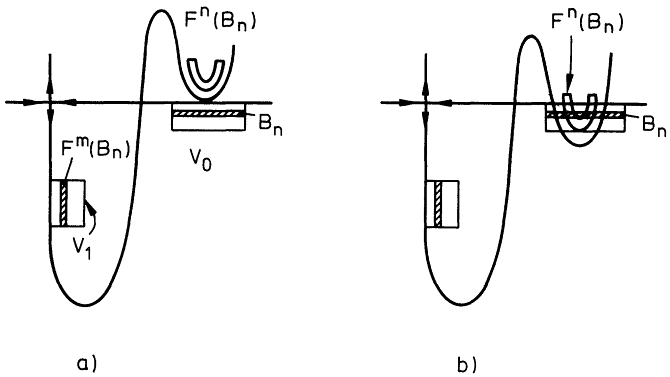


Fig. 6a and b. Location for boxes B_n and their images $F_t^m(B_n)$ and $F_t^n(B_n)$ for one case of homoclinic tangency. **a** $t=t_0$; **b** $t>t_0$

These choices can be made so that $V_0 \cap F_t(V_0)=\emptyset$, $V_1 \cap F_t(V_1)=\emptyset$, $\gamma_t^s \subset$ boundary V_0, and $\gamma_t^u \subset$ boundary $F_t^k(V_1)$, where $\gamma_t^s \subset W^s(P_t, F_t)$ and $\gamma_t^u \subset W^u(P_t, F_t)$ are as specified in the condition that $\{F_t\}$ creates odd order homoclinic intersections at t_0 near q_0. For N large enough, $n \geqq N$, and $m=n-k$,

$$B_n = \text{component}\,(V_0 \cap F_t^{-m}(V_1)) = \{(x, y) : |x-x_0| \leqq \delta_1^s,\ |y-\lambda^{-m} y_1| = \lambda^{-m}\delta_1^u\}$$

is a horizontal strip near q_0, where we take the component which is the first intersection along $F^{-m}(V_1)$. Then

$$F^m(B_n) = \{(x, y) : |y-y_1| \leqq \delta_1^u,\ |x-\mu^m x_0| \leqq \mu^m \delta_0^s\}$$

is a vertical strip near q_1, and $F_t^n = F_t^k \circ F_t^m(B_n)$ is a thin nonlinear box near q_0 which is parallel to $W^u(P_t, F_t)$.

The next step is to show Proposition 3.3 implies F_t^n has a sink in B_n for appropriate t. Take $t_0 < T < t_0 + \varepsilon$. Orienting $W^s(P_T, F_T)$ and $W^u(P_T, F_T)$, there is at least one intersection of these manifolds at some q_2 near q_0 which is an odd order intersection and the sign of the intersection is different at q_2 than at P_T. For large n, $W^u(P_T, F_T)$ intersects B_n monotonically near q_2 and $F_T^n(B_n)$ crosses B_n monotonically with a reversal in the y direction. We show below in Proposition 5.1 that F_t^n has a saddle point in B_n with eigenvalue $\lambda_1 < -1$. Remark 3.4 applies to show that there are no fixed points on the boundary of B_n as t varies. Also $F_T^n(B_n) \cap B_n = \emptyset$ for either $t=t_0$ or $t=t_0-\varepsilon$. Therefore Proposition 3.3 applies and there are t_n and $\varepsilon_n > 0$ such that for $t_n < t < t_n + \varepsilon_n$, F_t has a sink p_n in B_n of period n. Note that if $F_t^n(B_n) \cap B_n = \emptyset$ for $t=t_0$, then it follows from Proposition 3.3 that $t_0 < t_n < t_n + \varepsilon_n < T < t_0 + \varepsilon$. On the other hand if $F_t^n(B_n) \cap B_n \neq \emptyset$ for $t=t_0$, then $t_0 - \varepsilon < t_n < t_n + \varepsilon_n < t_0$.

As n increases B_n is closer to $W^s(P_t, F_t)$ and $F_t^n(B_n)$ is closer to $W^u(P_t, F_t)$. Therefore for larger n, T can be chosen nearer to t_0 in the argument above to show there is a sink of period n. Therefore as n increases t_n approaches t_0.

All that is left is to show that there is a saddle fixed point of F_T^n in B_n with eigenvalues $\lambda_1 < -1$ and $|\lambda_2| < 1$. This fact follows from the following proposition with $F = F_T$ and $P = P_T$.

5.1. Proposition. (a) *Assume F is C^j and has a fixed point P which has a homoclinic intersection of order j at q_2 with j odd. Then q_2 is the limit of hyperbolic saddle point z_n of period n, where z_n is in the box B_n as defined above.* (b) *If the sign of the intersection of $W^s(P,F)$ and $W^u(P,F)$ is different at q_2 than at P, then the saddle points z_n have negative unstable eigenvalue, $\lambda_1 < -1$.*

When the family nondegenerately creates intersections, then $j=1$ and this result follows from the usual transverse homoclinic point result. When j is larger than one, then the image $F^n(B_n)$ stretches across B_n but the slope is small and so the argument is delicate. In fact if the intersection is C^∞ flat, then the result is probably false. Therefore we give the details of this proof which takes most of the rest of the section (up to Remark 5.2).

Proof. We take the case as in part (b) of the proposition, but the other case is similar. Let $q_2=(x_2,0)$ and $q_3=F^{-k}(q_2)=(0,y_3)$ be in the linearized neighborhood of P. We take n large enough so that $F^n(B_n)$ crosses B_n monotonically. Therefore there is a horizontal subbox

$$B'_n=\{(x,y):|x-x_0|<\delta_0^s,\ A_1 \leqq y \leqq A_2\}$$

in B_n such that (i) $F^n(B'_n)$ stretches across B_n, $F_2^n(x,A_1) \geqq \lambda^{-m}y_1+\lambda^{-m}\delta_1^u \geqq A_2$ and $F_2^n(x,A_2) \leqq \lambda^{-m}y_1-\lambda^{-m}\delta_1^u \leqq A_1$, and (ii) $\partial F_2^n/\partial y(z)<0$ for z in B'_n. Because of the hyperbolic estimates on U, $|\partial F^n/\partial x(z)|<1$ for z in B_n.

In this situation it follows that F^n has a fixed point $z_n=(x_n,y_n)$ in B'_n. The existence of z_n can be shown using an index argument on B'_n. Alternatively it is possible to show there are (i) a horizontal curve of points in B'_n which have the y coordinate fixed, and (ii) a vertical curve of points in B'_n which have the x coordinate fixed. The point z_n where these two curves intersect is a fixed point. We need to show that z_n is a saddle point with reflection in the unstable direction. Since $|\lambda\mu|<1$, it follows that $|\det DF^n(z_n)|<1$. Thus if we show one of the eigenvalues $\lambda_1<-1$, the other eigenvalue λ_2 will have to have $|\lambda_2|<1$.

The graph of $W^u(P,F)$ near q_2 is given by $y=g(x)=a(x-x_2)^j+h(x)$. The slope is $g'(x)=ja(x-x_2)^{j-1}+h'(x)$. As a function of y, the slope at $z_n=(x_n,y_n)$ is about $ja(y_n/a)^{(j-1)/j}=ja^{1/j}y_n^{1-1/j}\equiv\xi_n$. Let $\xi_n^+=|\xi_n|$. We need an estimate below on ξ_n^+. Since $(|y_1+\delta_1^u|)\lambda^{-m} \geqq |y_n| \geqq |y_1-\delta_1^u|\lambda^{-m}$, there is a C independent of n such that $2C\lambda^{-m+m/j} \geqq \xi_n^+ \geqq C\lambda^{-m+m/j}$. To find the unstable eigenvector we take the sector of vectors whose slopes are about ξ_n,

$$S(z_n)=\{v=(v_1,v_2):\eta^{-1}<v_2/(v_1\xi_n)<\eta\},$$

where $\eta>1$ is independent of n. Thus for v in $S(z_n)$ the slope of v is between $\eta^{-1}\xi_n$ and $\eta\xi_n$. We show that for v in $S(z_n)$, $v''=DF^n(z_n)v$ is also in $S(z_n)$ and $\|v''\|>\|v\|$. Thus there is a vector v_0 in $S(z_n)$ such that $DF^n(z_n)v_0=\lambda_1 v_0$ and $|\lambda_1|>1$. Since $DF^n(z_n)$ reverses the second coordinate, $\lambda_1<-1$.

The proof of these facts about vectors v in $S(z_n)$ is like the proofs of Lemmas 7.5 and 7.6 below. There we prove there is a hyperbolic invariant set under slightly different assumptions. Let $v=(v_1,v_2)$ be in $S(z_n)$, $v'=(v'_1,v'_2)=DF^m(z_n)v=(\mu^m v_1,\lambda^m v_2)$, and $v''=(v''_1,v''_2)=DF^n(z_n)v$. The linear estimates in U and the bound

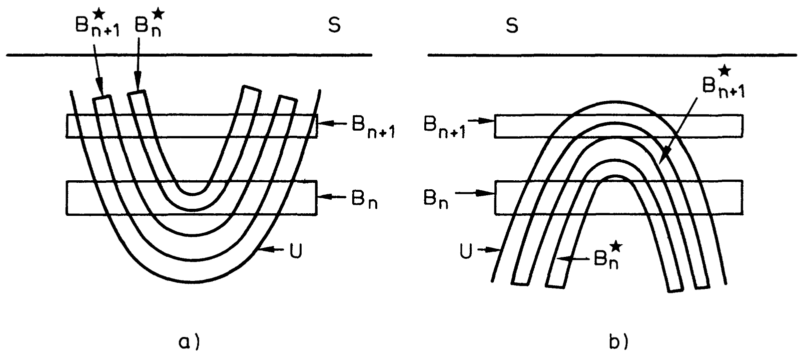


Fig. 7a and b. The stable and unstable manifolds of P_t are labelled S and U, respectively. The images $F_t^j(B_j)$ are labelled B_j^*. **a** F_t is as in Fig. 5 with $t_n<t<t_n+\varepsilon_n$; **b** F_t is as in Fig. 2 with $t_{n+1}<t<t_{n+1}+\varepsilon_{n+1}$

on the slope of v show that v' is nearly vertical for n large:

$$\begin{aligned}|v_2'/v_1'| &= \lambda^m\mu^{-m}|v_1/v_2| \geqq \lambda^m\mu^{-m}\eta^{-1}\xi_n^+\\ &\geqq \lambda^n\mu^{-m}\eta^{-1}C\lambda^{-m+m/j}\\ &\geqq \eta^{-1}C\mu^{-m}\lambda^{m/j}.\end{aligned}$$

Then v'' has slope about ξ_n, in fact the slope is between $\eta^{-1}\xi_n$ and $\eta\xi_n$ for large n. Thus v'' lies in $S(z_n)$. Moreover using the fact there is a constant $C_1>0$ such that for any z near q_3 and for any v', $\|v''\|=\|DF^k(z)v'\|\geqq C_1\|v'\|$; it follows that $\|v''\|>\|v\|$:

$$\begin{aligned}\|v''\| &\geqq C_1\|v'\| \geqq C_1|v_2'| \geqq C_1\lambda^m|v_2| \geqq C_1\lambda^m\eta^{-1}\xi_n(1+\eta^{-2}\xi_n^2)^{-1/2}\|v\|\\ &\geqq C_1\lambda^m\eta^{-1}\xi_n\|v\|/2 \geqq C_1\lambda^m\eta^{-1}C\lambda^{-m+m/j}\|v\|/2 \geqq \eta^{-1}C_1C\lambda^{m/j}\|v\|/2 \geqq \|v\|\end{aligned}$$

for large enough m. This completes the proof of Proposition 5.1 and Theorem D.

5.2. Remark. The intervals of parameter values $t_n<t<t_n+\varepsilon_n$ obtained above for the sinks p_n of period n do not overlap for large n, $n\geqq N$, at least in the case where $\{F_t\}$ nondegenerately creates homoclinic intersections. In the case of Figs. 7a and 6, by the time $t>t_n$, where the sink p_n exists and $F_t^n(B_n)\cap B_n\neq\emptyset$, the geometry forces $F_t^{n+1}(B_{n+1})\cap B_{n+1}$ to be a complete horseshoe and the periodic point p_{n+1} has become a saddle point with reflection. Therefore $t_0<\ldots<t_{n+1}<t_{n+1}+\varepsilon_{n+1}<t_n<t_n+\varepsilon_n$ for $n\geqq N$. In other cases such as Figs. 7b and 2 the geometry does not force the situation but a distance estimate does. The boxes B_n and B_{n+1} are roughly $\lambda^{-m}(1-\lambda^{-1})$ distance apart where $m=n-k$, while the images $F_t^n(B_n)$ and $F_t^{n+1}(B_{n+1})$ are roughly $\mu^m(1-\mu)$ distance apart. Since $\mu<\lambda^{-1}$, by the time $t_{n+1}<t<t_{n+1}+\varepsilon_{n+1}$, where the sink p_{n+1} exists and $F_t^{n+1}(B_{n+1})\cap B_{n+1}\neq\emptyset$, $F_t^n(B_n)\cap B_n$ is a complete horseshoe and p_n is a saddle point with reflection. Therefore $t_0>\ldots>t_{n+1}>t_{n+1}+\varepsilon_{n+1}>t_n>t_n+\varepsilon_n$ for $n\geqq N$. Therefore in any case the above argument can prove the existence of an infinite cascade of sinks, but it is not sufficient to prove there is one parameter value with an infinite number of sinks. The extra needed ingredient in the proof is the thickness of the Cantor set of stable manifolds of a basic set. The next section defines the thickness of a Cantor set and then Sect. 7 applies it to the stable manifolds.

5.3. Remark. The result of Curry and Johnson mentioned in Remark 2.3, [CJ], follows from the distance of the boxes B_n from $W^s(P_t, F_t)$ and $F_t^n(B_n)$ from $W^u(P_t, F_t)$. There distances are roughly $C_1\lambda^{-m}$ and $C_2\mu^m$ for $m=n-k$. If these boxes are pulled across at a linear rate then $t_n - t_0 \simeq C_3(C_1\lambda^{-m} \pm C_2\mu^m) \simeq C\lambda^{-m}$ since $\lambda^{-1} > \mu$. Thus

$$\lim_{n\to\infty} (t_n - t_{n-1})/(t_{n+1} - t_n) = \lambda.$$

See [CJ]. Also compare with the result in [GS2] about the parameter value having hyperbolic sets.

5.4. Appendix. Estimates without linearizing. Because we use the Taylor expansion of the graph of $W^u(P_T, F_T)$, we need C^j coordinates. Unless we make nonresonance assumptions on the eigenvalues, it is not possible to linearize the system. However it is still possible to find coordinates on which there are hyperbolic estimates which are good enough to prove z_n is a saddle point with reflection in the unstable direction.

There is a neighborhood u of $P=P_T$ and C^j coordinates (x, y) on U so that the local stable and unstable manifolds of P are given by $\mathrm{comp}(W^s(P, F) \cap U) = \{(x, 0) \cap U$ and $\mathrm{comp}(W^u(P, F) \cap U) = \{(0, y)\} \cap U$. Also

$$DF(x, y) = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},$$

where $\mu - \varepsilon \leqq b_{11} \leqq \mu + \varepsilon$, $\lambda - \varepsilon \leqq b_{22} \leqq \lambda + \varepsilon$, $|b_{12}| \leqq \varepsilon|x|$, and $|b_{21}| \leqq \varepsilon|y|$. These last two estimates follow because $b_{12}(0, y) = 0$, $(W^u = \{(0, y)\})$, so

$$|b_{12}(x, y)| \leqq |x| \sup|\partial^2 F_1/\partial x \partial y|$$

by the Mean Value Theorem. By a change of scale the second partial derivatives can all be made less than ε. The estimate on $b_{21}(x, y)$ is similar.

We need to show that $v' = (v'_1, v'_2) = DF^m(z_n)v$ is nearly vertical. The fact that $\|v''\| = \|DF^n(z_n)v\| > \|v\|$ follows directly as before. If $(x_r, y_r) = F^r(z_n)$ for $0 \leqq r \leqq m$, then

$$|x_r| = |F_1(x_{r-1}, y_{r-1}) - F_1(0, y_{r-1})| < |x_{r-1}| \sup|\partial F_1/\partial x| \leqq (\mu+\varepsilon)|x_{r-1}| \leqq (\mu+\varepsilon)^r |x_0|$$

by induction. Similarly $|y_r| < |y_m|(\lambda - \varepsilon)^{r-m}$. Let $DF(x_r, y_r) = (b^r_{ij})$. Then

$$\mu - \varepsilon \leqq b^r_{11} \leqq \mu + \varepsilon, \qquad |b^r_{12}| \leqq \varepsilon |x_0| (\mu+\varepsilon)^r,$$
$$|b^r_{21}| \leqq \varepsilon |y_m| (\lambda - \varepsilon)^{r-m}, \qquad \lambda - \varepsilon \leqq b^r_{22} \leqq \lambda + \varepsilon.$$

It suffices to consider v in $S(z_n)$ with $v_1 = 1$. Let $v^k = (v^r_1, v^r_2) = DF^r(z_n)v$. We know that $\eta^{-1}\xi_n^+ \leqq |v_2| \leqq \eta\varepsilon_n^+$ and $\xi_n^+ = A\lambda^{-m+m/j}$ with $C \leqq A \leqq 2C$. We prove by induction on r that

$$\eta^{-1} A(\lambda - 2\varepsilon)^{r-m+m/j} \leqq |v^r_2| \leqq A(\lambda + 2\varepsilon)^{r-m+m/j}, \tag{5.5}$$
$$|v^r_1| \leqq [(\mu + 2\varepsilon)(\lambda + 2\varepsilon)]^r.$$

Once we prove these inequalities then

$$|v'_2|/|v'_1| = |v^m_2|/|v^m_1| \geqq \eta^{-1} A(\lambda - 2\varepsilon)^{m/j} [(\mu + 2\varepsilon)(\lambda + 2\varepsilon)]^{-m}$$

is arbitrarily large as m goes to infinity (for small enough ε).

To show (5.5) for v_1^r

$$\begin{aligned}|v_1^{r+1}| &\leqq |b_{11}^r|\,|v_1^r| + |b_{12}^r|\,|v_2^r| \\ &\leqq (\mu+\varepsilon)[(\mu+2\varepsilon)(\lambda+2\varepsilon)]^r + \varepsilon|x_0|(\mu+\varepsilon)^r A(\lambda+2\varepsilon)^{r-m+m/j} \\ &\leqq [(\mu+2\varepsilon)(\lambda+2\varepsilon)]^{r+1}\{(\lambda+2\varepsilon)^{-1} + \varepsilon|x_0|\eta A(\mu+2\varepsilon)^{-1}(\lambda+2\varepsilon)^{-1-m+m/j}\} \\ &\leqq [(\mu+2\varepsilon)(\lambda+2\varepsilon)]^r\end{aligned}$$

for ε small enough. Similarly for v_2^r,

$$\begin{aligned}|v_2^{r+1}| &\geqq -|b_{21}|\,|v_1^r| + |b_{22}|\,|v_2^r| \\ &\geqq -\varepsilon|y_m|(\lambda-\varepsilon)^{r-m}[(\mu+2\varepsilon)(\lambda+2\varepsilon)]^r + (\lambda-\varepsilon)\eta^{-1}A(\lambda-2\varepsilon)^{r-m+m/j} \\ &\geqq \eta^{-1}A(\lambda-2\varepsilon)^{r+1-m+m/j}\{-\varepsilon|y_m|(\lambda-2\varepsilon)^{-1-m/j} + (\lambda-\varepsilon)(\lambda-2\varepsilon)^{-1}\} \\ &\geqq \eta^{-1}A(\lambda-2\varepsilon)^{r+1-m+m/j}\end{aligned}$$

for ε small enough. The upper estimate for $|v_2^{r+1}|$ is similar. This completes the necessary modifications without linearization.

6. Persistent Intersections of Cantor Sets

Given two Cantor sets A^1 and A^2 in the line $\mathbb{R}$, we want a criterion to imply that $A^1 \cap A^2 \neq \emptyset$. The relevant condition is the thickness as defined by Newhouse in [N1]. It is related to the Hausdorff dimension of the set, but different. See [N4, p. 107].

A Cantor set A in the line can be represented as $A = \bigcap_{i \geqq 0} A_i$, where A_0 is the smallest interval containing A and $A_i = A_0 - U_{1 \leqq j \leqq i} U_j$ and U_j are open intervals. Such a sequence of sets $\{A_i\}$ is called a *defining sequence* of A. It is obtained by specifying an ordering of the gaps removed from A_0 to form A.

The thickness is the ratio of the length of intervals left in A_i to the length of the adjacent gap U_i. Let I_{ij} for $j=1,2$ be the two components of A_i on either side of the gap U_i. Let ℓJ be the length of an interval J. The thickness of a defining sequence is defined by

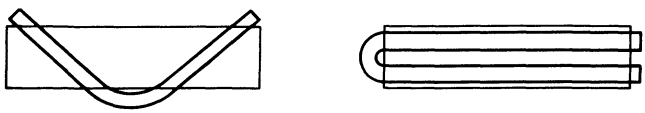
$$\tau(\{A_i\}) = \inf\{\ell(I_{ij})/\ell(U_i) : i \geqq 1,\ j=1,2\}.$$

This thickness depends on the choice of the defining sequence. *The thickness of the Cantor set* is defined as the thickness for the best choice of a defining sequence:

$$\tau(A) = \sup\{\tau(\{A_i\}) : \{A_i\} \text{ is a defining sequence for } A\}.$$

For a Cantor set A_α formed by removing the middle α of the remaining intervals at each step, the intervals at the k^{th} step have length $((1-\alpha)/2)^k$. The gaps formed at the k^{th} step are each α of the length of the intervals at the $(k-1)^{\text{th}}$ step, or $\alpha((1-\alpha)/2)^{k-1}$. Therefore the ratio is always

$$\frac{\left(\frac{1-\alpha}{2}\right)^k}{\alpha\left(\frac{1-\alpha}{2}\right)^{k-1}} = \frac{1-\alpha}{2\alpha}.$$



a) b)

Fig. 8. **a** $F_t^n(B_n)$ is shown crossing B_n for t as in Proposition 7.1; **b** $F_t^{-n}(B_n)$ is shown crossing B_n

This defining sequence is actually the best choice so $\tau(A_\alpha)=(1-\alpha)/2\alpha$. In particular $\tau(A_{1/3})=1$ for the middle third Cantor set. Also note $0<\tau(A)<\infty$ and $\tau(A_\alpha)$ goes to infinity as α goes to zero.

The reason for the above definition of thickness is the following lemma which is proven in [N4, Lemma 4, p. 107].

6.1. Lemma. *Let A^1 and A^2 be two Cantor sets in $\mathbb{R}$ with A^1 not contained in any gap of A^2 and vice versa. In particular if A_0^j is the smallest interval containing A^j, then assume $\operatorname{int} A_0^1 \cap \operatorname{int} A_0^2 \neq \emptyset$. Further assume $\tau(A^1)\tau(A^2)>1$. Then $A^1 \cap A^2 \neq \emptyset$. In fact if $\{A_i^j\}$ are defining sequences for A^j for $j=1,2$ such that $\tau(\{A_i^1\})\tau(\{A_i^2\})>1$, then for each $i \geqq 1$, $\operatorname{int}(A_i^1 \cap A_i^2) \neq \emptyset$.*

7. Thick Cantor Sets of Stable Manifolds from a Nondegenerate Tangency

If Λ is a hyperbolic set for F, the *stable thickness* of Λ, $\tau^s(\Lambda)$, is defined by $\tau^s(\Lambda) = \limsup_{\varepsilon\to 0}\{\tau(\gamma_\varepsilon \cap W^s(\Lambda))\}$, where γ is any C^1 arc transverse to the stable manifolds $W^s(\Lambda)$ at q and γ_ε is the arc of length ε in γ centered at q. The result [N4, Proposition 5] shows this definition of thickness is independent of γ and q for F which is C^2 and γ is C^1. Thus it is well defined. The *unstable thickness*, $\tau^u(\Lambda)$, is defined in a similar manner.

It is not hard to construct a diffeomorphism which has a horseshoe Λ with $W^s(\Lambda)\cap\gamma^s$ and $W^u(\Lambda)\cap\gamma^u$ any desired middle α and β Cantor sets. Thus $\tau^s(\Lambda)$ and $\tau^u(\Lambda)$ are arbitrarily large. Moreover it can be done so that the fact $\tau^s(\Lambda)\tau^u(\Lambda)>1$ implies there is a persistent tangency of $W^s(\Lambda)$ and $W^u(\Lambda)$. See [N1] and [N5, pp. 103–104].

In this section and the next, we show that whenever a family nondegenerately creates a homoclinic intersections, then there are created hyperbolic invariant sets $\Lambda(t)$ which have $\tau^s(\Lambda(t))$ arbitrarily large and a persistent tangency. The idea of the construction is as follows. In this section the parameter value t_n^* is chosen carefully so that there is a hyperbolic invariant set $\Lambda_n(t_n^*)$ in B_n with $\tau^s(\Lambda_n(t_n^*))$ arbitrarily large. The parameter value must be chosen so $F_t^n(B_n)$ comes out the bottom of the box B_n enough so the maximal invariant set in B_n has a hyperbolic structure, $\Lambda_n(t) = \bigcap_n F_t^n(B_n)$. On the other hand t must be chosen so that most of $F_t^n(B_n)$ lies in B_n so $\Lambda_n(t_n^*)$ has large stable thickness. In terms of the quantities in Fig. 9 below, the gap g must be large enough to make $\Lambda_n(t_n^*)$ hyperbolic while the gap g' must be small enough to make $\tau^s(\Lambda_n(t_n^*))$ large. However for $t=t_n^*$ there might not be any homoclinic tangencies because the local extreme points of $W^u(\Lambda_n(t_n^*))$ relative to $W^s(\Lambda_n(t_n^*))$ lie away from $W^s(\Lambda_n(t_n^*))$. In the next section, t_n^* is decreased to t_n^{**} to

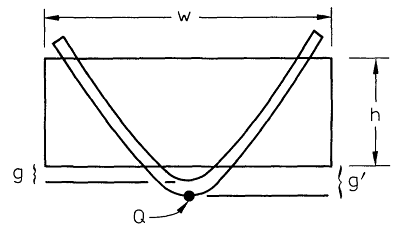


Fig. 9. The gaps g and g' are indicated as well as the height h and width w. The point Q is $(x_n(t), y_n(t))$

cause the extreme points of $W^u(\Lambda_n(t_n^*))$ to intersect the $W^s(\Lambda_n(t_n^*))$ and so to cause a persistent tangency. The value $t=t_n^{**}$ is chosen carefully so that $F_t^n(B_n)$ does not come out the bottom of B_n but there is still a smaller hyperbolic set $\Lambda_n'(t)$ in B_n with $\tau^s(\Lambda_n'(t))$ large. There is a C^1 curve γ along which $W^u(\Lambda_n'(t))$ has extrema relative to $W^s(\Lambda_n'(t))$. The value $t=t_n^{**}$ is chosen so that $\gamma \cap W^u(\Lambda_n'(t))$ and $\gamma \cap W^s(\Lambda_n'(t))$ overlap along γ and the product thickness is larger than one. Thus there is a tangency of $W^u(\Lambda_n'(t))$ and $W^s(\Lambda_n'(t))$ at some point of γ. This tangency persists for t near t_n^{**}. More details of the construction will follow.

It is easier to use coordinates in which F_t is linear. We need to have C^2 coordinates because we look at the curvature of $W^u(P_t, F_t)$ near q_0. If F_t is C^∞ and the eigenvalues at P_t satisfy nonresonance conditions, then Sternberg linearization gives C^2 coordinates in which F_t is linear near P_t. Even without the nonresonance conditions, Newhouse showed that if F_t is C^3 it is possible to find coordinates on U that are C^1 everywhere, are C^2 off $W^u(P_t, F_t)=\{(0, y)\}$, and such that in these coordinates F_t is linear, $F_t(x, y)=(\mu x, \lambda y)$. See [N4, pp. 124–126]. The proof uses the fact that $\mu\lambda<1$. The main idea is that with these assumptions there are C^2 line bundles in the stable and unstable directions off $W^u(P_t, F_t)$. We use these coordinates in this and the next section.

7.1. Proposition. [N4, Lemma 7.] *For suitably chosen $t>t_n+\varepsilon_n$, the hyperbolic basic set $\Lambda_n(t)$ created in $F_t^n(B_n)\cap B_n$ has arbitrarily large $\tau^s(\Lambda_n(t))$.*

Proof. The idea of the proof is that because $\mu\lambda<1$, $F_t^n(B_n)$ is much thinner than it is long. Therefore by carefully choosing the dimensions of B_n, for value t for which $F_t^n(B_n)\cap B_n$ first has two components, the length of the bend $F_t^n(B_n)-B_n$ is much smaller than the length of the components $F_t^n(B_n)\cap B_n$. This ratio is roughly the same as the ratio of the height of the component of $F_t^n(B_n)\cap B_n$ to the height of the middle gap in $B_n-F_t^{-n}(B_n)$, which in turn can be shown to be roughly the stable thickness of the hyperbolic set for F_t^n in B_n, $\Lambda_n(t)=\bigcap_{j=-\infty}^{\infty} F_t^{jn}(B_n)$. See [N4, pp. 121–138] for the proof. Since this ratio is arbitrarily large for large n, $\tau^s(\Lambda_n(t))$ is arbitrarily large for suitably chosen t and large n. Also $\Lambda_n(t)$ has a dense orbit because it is conjugate to a shift on two symbols. Therefore $\Lambda_n(t)$ is a basic set.

Let g be the gap by which $F_t^n(B_n)$ clears the bottom of B_n at the bottom bend

$$g=\inf\{\sup\{\text{distance}(F_t^n(x_0, y), B_n) : (x_0, y)\in B_n\ x_0 \text{ fixed}\} : x_0 \text{ varies with } (x_0, y')\in B_n\}.$$

See Fig. 9. Let g' be the distance the lowest point on the bend is from the bottom of B_n, $g'=\sup\{\sup\{\text{distance}(F_t^n(x_0,y),B_n):(x_0,y)\in B_n\ x_0 \text{ fixed}\}:x_0 \text{ varies with } (x_0,y')\in B_n\}$.

The dimensions of the box B_n and bounds on g and g' (hence choice of t) have to be chosen to satisfy four conditions:

(7.2a) the height h of B_n must be big enough for $F_t^n(B_n)$ to stretch across B_n from top to bottom and back to the top again,

(7.2b) the width w must be large enough so that $F_t^n(B_n)$ comes out the top of B_n and not the sides,

(7.2c) the bound on g' must be small enough to insure that for x' fixed the length of a component of $F_t^n\{(x',y):(x',y)\in B_n\}\cap B_n$ must be long relative to the components of $F_t^n\{(x',y)\in B_n\}-B_n$ so that $\tau^s(\Lambda_n(t))$ is arbitrarily large,

(7.2d) the gap g must be large enough to make the slope of $\partial F_t^n/\partial y(q)$ large enough for q in $B_n\cap F_t^n(B_n)$ to prove the set $\Lambda_n(t)$ has a hyperbolic structure.

Let $m=n-k$ and $H_t=F_t^k$ be the map from $F_t^m(B_n)\ V_1$ to V_0. Thus F_t^n from B_n to V_0 is given by $F_t^n(x,y)=H_t(\lambda^m x,\mu^m y)$. For $t=t_0$, $H_t(0,y_1)=(x_0,0)$ corresponds to the nondegenerate tangency. Let

$$d_1=|\partial H_t/\partial y(0,y_1)|=|\partial H_{1t}/\partial y(0,y_1)|,$$
$$d_2=|\partial H_t/\partial x(0,y_1)| \quad \text{and} \tag{7.3}$$
$$y=a(x-x_n(t))^2+y_n(t)$$

be such that for x' fixed $\{H(x',y):(x',y)\in F_t^m(B_n)\}$ is C^2 near $\{(x,y):y=a(x-x_n(t))^2+y_n(t)\}$. Here $(x_n(t),y_n(t))$ is the point where y has the smallest value on $F_t^n(B_n)$. Letting α be an arbitrarily small positive constant which is different for different quantities, the choices which work to satisfy (7.2) are

$$w=\frac{(4+\alpha)}{ad_1}\lambda^{-m},\qquad h=\frac{w\lambda^{-m}}{d_1-\alpha}=\frac{(4+\alpha)}{ad_1^2}\lambda^{-2m},$$
$$g\geqq a^{-1}(\mu\lambda^{-1})^{2m-2\alpha m},\qquad g'\leqq w10d_2\mu^m=h10d_2(d_1-\alpha)(\mu\lambda)^m. \tag{7.4}$$

The choices for w and h are the same as [N4, p. 126 (1)–(4)] but the bounds for g and g' are different and more specific. In [N4], g and g' are assumed to be $\alpha w\lambda^{-m}$, where α is a small unspecified constant. Our lower bound on g makes it clear how large it must be in order to show that $\Lambda_n(t)$ has a hyperbolic structure. This bound is important in the next section to know, for a nearby value of t, how much of B_n must be removed to get a hyperbolic set $\Lambda_n'(t)$. As a final remark, the only place the lower bound of g is used is in the proof of Lemma 7.5 which shows there are invariant sectors in the unstable direction. The rest of this section sketches the proof that the choices (7.4) imply (7.2).

Proof of 7.2a. For x' fixed, on either side of the minimum the curve $F_t^n\{(x',y\in B_n)\}$ is monotone so its length is less than the sum of the change in the x value plus the change in the y value. Since $F_t^n(x,y)=H_t(\mu^m x,\lambda^m y)$, to show $F_t^n(B_n)$ can stretch across B_n it suffices for

$$(d_1-\alpha)\lambda^m h\geqq 2(w/2)+2(h+g')$$

or

$$h \geqq [(d_1 - \alpha)\lambda^m - 2]^{-1} w(1 + 10 d_2 \mu^m).$$

Thus the value $h = w(d_1 - \alpha)^{-1}\lambda^{-m}$ in (6.4) suffices when the α for λ is different and slightly larger than the α for g'. □

Proof of 7.2b. To show $F_t^n(B_n)$ comes out the top of B_n rather than the sides, it suffices to show that $a(w/2)^2$ is larger than $h + g' = w\lambda^{-m}[(d_1 - \alpha)^{-1} + 10 d_2(\mu\lambda)^m]$. Thus it suffices for

$$w \geqq \lambda^{-m} 4 a^{-1}[(d_1 - \alpha)^{-1} + 10 d_2(\mu\lambda)^m].$$

Thus $w = (4 + \alpha)\lambda^{-m} a^{-1} d_1^{-1}$ as in (7.4) is adequate to satisfy (7.2b). □

Proof of 7c. We need to show for x' fixed the length of a component of $F_t^n\{(x', y) \in B_n\} \cap B_n$ is long compared with the length of the gap $F_t^n\{(x'y) \in B_n\} - B_n$. The length of the gap is less than the sum of the change in its x coordinate plus the change in its y coordinate. Thus it is less than

$$g' + (g'/a)^{1/2} \leqq w 10 d_2 \mu^m + (w 10 d_2 \mu^m a^{-1})^{1/2}.$$

The length of a component of $F_t^n\{(x', y) \in B_n\} \cap B_n$ is greater than the change in its x-coordinate. Thus it is greater than

$$(h/a)^{1/2} - (g'/a)^{1/2} > w^{1/2}\lambda^{-m/2}(d_1 - \alpha)^{-1/2} a^{-1/2} - (w 10 d_2 \mu^m a^{-1})^{1/2}.$$

Combining, the ratio is greater than

$$\frac{w^{1/2}\lambda^{-m/2}[a(d_1 - \alpha)]^{-1/2} - w^{1/2}\mu^{m/2}[10 d_2 a^{-1}]^{1/2}}{w 10 d_2 \mu^m + w^{1/2}\mu^{m/2}[10 d_2 a^{-1}]^{1/2}} \geqq \frac{w^{1/2}\lambda^{-m/2}[a d_1]^{-1/2} 2^{-1}}{w^{1/2}\mu^{m/2} 4 [d_2 a^{-1}]^{1/2}}$$
$$= (\mu\lambda)^{-m/2} 8^{-1} d_1^{-1/2} d_2^{-1/2}.$$

Since $\mu\lambda < 1$, this ratio can be made as large as desired for large n. The argument in [N4, pp. 134–136] shows this estimate implies the stable thickness $\tau^s(\Lambda_n(t))$ is as large as desired for large n. □

Proof of 7.2d. To show $\Lambda_n(t)$ has a hyperbolic structure, it is sufficient to find invariant unstable sectors $S(z) \subset T_z M$ for z in $F_t^n(B_n) \cap B_n$ which are invariant and expanded by $DF_t^{kn}(z)$ for large k, $|DF_t^{kn}(z)v| > v$ for v in $S(z)$. Unfortunately the vectors in the unstable sectors are not expanded for $k = 1$, but there is a power $j = k(z)n$ depending on the point z which expands vectors. By a compactness argument there is one power kn which works for all points in $\Lambda_n(t)$ proving the existence of an expanding invariant subbundle. The existence of the stable bundle is similar and even slightly easier because $\mu < \lambda^{-1}$.

The first step is to define the sectors $S(z)$. For $z = (x, y)$ in $F_t^n(B_n) \cap B_n$ and $z_{-1} = F_t^{-n}(z)$, the slope of $(\partial F_t^n/\partial y)(z_{-1})$ is about $2a(x - x_n(t))$. Define $\xi(z)$, $\xi^+(z)$, and $\eta > 1$ by

$$\xi(z) = 2a(x - x_n(t)),$$
$$\xi^+(z) = |\xi(z)| = 2[a(y - y_n(t))]^{1/2},$$
$$1 < \eta < 2^{1/6} = \min\{2^{1/6}, 4/3, 2^{3/4}\}.$$

The sectors are defined by

$$S(z)=\{v=(v_1,v_2)\in T_zM:\eta^{-1}\leqq v_2/[v_1\xi(z)]\leqq\eta\}.$$

Thus for v in $S(z)$ the slope of v lies between $\eta^{-1}\xi(z)$ and $\eta\xi(z)$. This allows for either positive or negative slope depending on the sign of $\xi(z)$, i.e. the sign of $x-x_n(t)$.

7.5. Lemma. *The sectors are invariant under DF_t^n, i.e. if z and $F_t^n(z)$ are in $B_n\cap F_t^n(B_n)$, then $DF_t^n(z)S(z)\subset S(F_t^n(z))$.*

Proof. If $v=(v_1,v_2)\in S(z)$, then $(v_1',v_2')\equiv DF_t^m(z)(v_1,v_2)=(\mu^m v_1,\lambda^m v_2)$ is nearly vertical:

$$\begin{aligned}|v_2'/v_1'|&=|\lambda^m v_2/\mu^m v_1|\geqslant(\lambda\mu^{-1})^m\eta^{-1}\xi^+(z)\\&\geqslant(\lambda\mu^{-1})^m\eta^{-1}2(ag)^{1/2}\geqslant(\lambda\mu^{-1})^m\eta^{-1}2(\lambda\mu^{-1})^{m-\alpha m}\\&\geqslant\eta^{-1}2(\lambda\mu^{-1})^{\alpha m},\end{aligned}$$

which is arbitrarily large for large n because $\lambda\mu^{-1}>1$. Then $v''=(v_1'',v_2'')$ $=DF_t^n(z)(v_1,v_2)=DH_t(F_t^m(z))(v_1',v_1')$ has slope about $\xi(F_t^n(z))$ which is between $\eta^{-1}\xi(F_t^n(z))$ and $\eta\xi(F_t^n(z))$. Thus v'' lies in $S(F_t^n(z))$. This proves Lemma 7.5. □

Next we show for z bounded away from the bend at the bottom, the vectors in $S(z)$ are immediately expanded.

7.6. Lemma. *If $z=(x,y)$ is in B_n with $y-y_n(t)=\gamma h$ and v is in $S(z)$, then $\|DF_t^n(z)v\|\geqq 4\gamma^{1/2}\eta^{-1}(1-\alpha)\|v\|$. If $\gamma\geqq 1/12$, then $\|DF_t^n(z)v\|>\|v\|$.*

Proof. Let $v''=DF_t^n(z)v$. Then

$$\begin{aligned}\|v''\|&\geqslant(d_1-\alpha)\|v'\|\geqslant d_1(1-\alpha)|v_2'|=d_1(1-\alpha)\lambda^m|v_2|\\&\geqslant d_1(1-\alpha)\lambda^m\eta^{-1}\xi^+(z)|v_1|\geqslant d_1(1-\alpha)\\&\geqslant d_1(1-\alpha)\lambda^m\eta^{-1}\xi^+(z)[1+\eta^2\xi^+(z)^2]^{-1/2}\|v\|\\&\geqslant d_1(1-\alpha)\lambda^m\eta^{-1}\xi^+(z)\|v\|\\&\geqslant d_1(1-\alpha)\lambda^m\eta^{-1}2[a\gamma h]^{1/2}\\&\geqslant 4\gamma^{1/2}\eta^{-1}(1-\alpha)\|v\|.\end{aligned}$$

For $\gamma\geqq 1/12$, $4\gamma^{1/2}\eta^{-1}(1-\alpha)>1$ by the choice of η. This proves (7.6). □

Even for the larger gap g allowed in [N4], it is not possible for all points in B_n to have $y-y_n(t)\geqq h/12$. For z with $y-y_n(t)\leqq h/12$, the derivative $DF_t^n(z)$ contracts vectors in $S(z)$, but the next few iterates $z_j=F_t^{nj}(z)$ for $1\leqq j\leqq k$ have $y_j-y_n(t)\geqq h/2$, and so $DF_t^n(z_j)$ expands vectors in $S(z_j)$. Claim 7.7 below shows that the total effect of $DF_t^{n(k+1)}(z)=DF_t^n(z_k)\dots DF_t^n(z_1)DF_t^n(z)$ is to expand vectors in $S(z)$. More precisely let y_0 be the value of y on the bottom of B_n and $H_1=\{z\in B_n:y-y_0\geqq h/2\}$. For $z=(x,y)$ with $y-y_0\leqq h/8$, let $k=k(z)$ be the integer such that $F_t^{nj}(z)\in H_1$ for $1\leqq j\leqq k(z)$ and $F_t^{n(k+1)}(z)\notin H_1$.

7.7. Lemma. *If $y-y_0\leqq h/12$, then for $k=k(z)$ as above and v in $S(z)$*

$$\|DF_t^{nk+n}(z)v\|\geqq(1-\alpha)\eta^{-1}2\|v\|>\|v\|.$$

Proof. Let $z=(x, y)$ be in B_n with $y-y_0=\gamma h$. To estimate the number of iterates z stays in H_1, $k(z)$, we need to show $z_1=(x_1, y_1)=F_t^n(z)$ is very near the top of B_n. Also we need to compare the maximal and minimal stretches of vectors in $S(\zeta)$ for ζ in H_1.

The partial derivative $|\partial F_t^n/\partial y| \leqq (1+\alpha)d_1\lambda^m$. The slope of the image $F_t^n(B_n)$ is always less than the slope at the top, s,

$$s \leqq 2[a(h+g')]^{1/2} < 2a^{1/2}h^{1/2}(1+\alpha) < 4(1+\alpha)d_1^{-1}\lambda^{-m},$$

because g'/h is small. Let y_{top} be the value of y on the top of B_n. Then

$$y_{\text{top}} - y_1 \leqq \gamma h(1+\alpha)d_1\lambda^m s \leqq \gamma h(1+\alpha)4 < h/2,$$

because $\gamma \leqq 1/12 < 1/8$. Thus z_1 is in H_1 and $y_{\text{top}} - y_1$ is proportional to γh.

Let $\lambda_{\max}$ and $\lambda_{\max}$ be the maximal and minimal stretch of vectors in $S(\zeta)$ for ζ in H_1. Similarly let $\xi_{\max}$ and $\xi_{\max}$ be the maximal and minimal value of $\xi^+(\zeta)$ for ζ in H_1. Then

$$\xi_{\max} \leqq 2[ah+ag']^{1/2} \leqq 2a^{1/2}h^{1/2}(1+\alpha) \leqq 4d_1^{-1}\lambda^{-m}(1+\alpha),$$

and

$$\xi_{\max} \geqq 2[ah/2]^{1/2} \geqq 2^{1/2}a^{1/2}h^{1/2} \geqq 2^{1/2}2d_1^{-1}\lambda^{-m}.$$

Then

$$\lambda_{\max} \leqq \lambda^m(d_1+\alpha)\eta\xi_{\max} \leqq 4\eta(1+\alpha),$$

and

$$\lambda_{\min} \geqq \lambda^m(d_1-\alpha)\eta^{-1}\xi_{\min} \geqq 2^{1/2}2\eta^{-1}(1-\alpha).$$

By the coice of η, $\lambda_{\min} \geqq (\lambda_{\max}^{1/2})2^{1/4}$ and $\lambda_{\min}^k \geqq \lambda_{\max}^{k/2}2^{1/4}$ because $k(z) \geqq 1$ (z_1 is in H_1).

Using the fact z_1 is a distance at most $\gamma h(1+\alpha)d_1\lambda^m$ down the strip $F_t^n(B_n)$ and that F_t^n stretches lengths along $F_t^n(B_n)$ at most $\lambda_{\max}$, it follws that the value $k=k(z)$ satisfies

$$\gamma h(1+\alpha)d_1\lambda^m\lambda_{\max}^k s \geqq h/2,$$

$$h\gamma 4(1+\alpha)\lambda_{\max} \geqq h/2,$$

$$\lambda_{\max}^k \geqq (1-\alpha)/(8\gamma),$$

and

$$\lambda_{\max}^k \geqq (1-\alpha)/(\gamma^{1/2}2^{5/4}).$$

Thus for v in $S(z)$ and $y-y_n(t)=\gamma h \leqq h/12$

$$\begin{aligned}\|DF_t^{nk+n}(z)v\| \cdot \|v\|^{-1} &\geqq \lambda^m(d_1-\alpha)\eta^{-1}\xi^+(z)\lambda_{\min}^k \\ &\geqq (1-\alpha)\lambda^m d_1\eta^{-1}2[a\gamma h]^{1/2}\gamma^{-1/2}2^{-5/4} \\ &\geqq (1-\alpha)\eta^{-1}2^{3/4} \\ &> 1\end{aligned}$$

by the choice of η. This completes the proof of (7.7). □

7.8. Proposition. *For g as given in* (7.4), *the maximal invariant set in B_n for F_t^n, $\Lambda_n(t)$, has a hyperbolic structure.*

Proof. The sector $S(z)$ is mapped inside the sector $S(F_t^n(z))$ by Lemma 7.5. Therefore there is an invariant bundle E_z^u for z in $\Lambda_n(t)$. By Lemmas 7.6 and 7.7, the vectors v in E_z^u are eventually expanded. Since this is true for each z in $\Lambda_n(t)$, there are $C>0$ and $\lambda_u>1$ such that for v in E_z^u and $j\geqq 0$, $\|DF_t^{nj}(z)v\|\geqq C\lambda_u^j\|v\|$. (This follows because all the Lyapunov characteristic exponents are positive.)

To show there is a contracting bundle, let $S^*(z)$ be the complementary sector to $S(z)$. It is overflowing by $DF_t^n(z)$ and so is invariant under $DF_t^{-n}(z): DF_t^{-n}(z)S^*(z) \subset S^*(F_t^{-n}(z))$. Therefore there is an invariant bundle E_z^s for z in $\Lambda_n(t)$. By the hypothesis that $\mu\lambda<1$, it follows that $\det(DF_t^n(z))<1$. Since $DF_t^n(z)$ is expanding on the invariant bundle E^u, it has to be contracting on the complementary invariant bundle E^s. This completes the proof of Proposition 7.8 and (7.2d). □

7.9. Remark. If the gap g is smaller than allowed in (7.4) and is so small that a vector v in a narrow sector $S(z)$ near the bottom of B_n has an image v'' which is horizontal near the top of B_n, then $DF_t^n(z_1)$ would contract v'' again. Therefore for such a small gap, the maximal invariant set is probably not hyperbolic.

8. Thick Stable Manifolds and Tangencies: Proof of Theorem D

Proposition 7.1 proves that there are values $t=t_n^*$ when F_t^n has a hyperbolic basic set $\Lambda_n(t_n^*)\subset B_n$ with $\tau^s(\Lambda_n(t_n^*))$ arbitrarily large. There are also other values $t=t_n$, where $W^u(P_t,F_t)$ has a nondegenerate tangency with $W^s(P_t,F_t)$. The following result shows that there are parameter values where both phenomena occur simultaneously. Since being a wild hyperbolic set is open, Theorem D follows.

8.1. Proposition. *For large enough n, there exists $t=t_n^{**}$ for which* (a) *F_t^n has a hyperbolic basic set $\Lambda_n'(t_n^{**})\subset B_n$,* (b) *$F_t$ has a hyperbolic basic set $\Lambda(t_n^{**})\supset\Lambda_n'(t_n^{**})\cup\{P_t\}$.* (c) *$\tau^s(\Lambda(t_n^{**})\tau^u(\Lambda(t_n^{**}))>1$, and* (d) *$\Lambda(t_n^{**})$ has a persistent tangency of $W^s(\Lambda(t_n^{**}))$ and $W^u(\Lambda(t_n^{**})$,* i.e. *$\Lambda(t_n^{**})$ is a wild hyperbolic set.*

Proof. The idea of the proof is to pull back the parameter value from t_n^* to $t=t_n^{**}$ so that $F_t^n(B_n)$ does not come out the bottom of B_n and there is a near tangency of $W^s(p_n)$ and $W^u(p_n)$, where $p_n=p_n(t)$ is a hyperbolic fixed point of F_t^n in B_n. Done correctly, p_n is still contained in a (smaller) hyperbolic basic set $\Lambda_n'(t)$ with $\tau^s(\Lambda_n'(t))\geqq\tau^s(\Lambda_n(t_n))-\varepsilon$.

Before this is done we need a preliminary step. It is seen by the work of Silnikov that there is a parameter value $t=t_1$ near t_0 such that $W^u(P_t,F_t)$ has some point of transverse intersection with $W^s(P_t,F_t)$ and other points where there is a nondegenerate tangency. See [HG], [GS] or [N4]. See Fig. 11b. Thus we can assume that P_t is already an element of a hyperbolic set $\Lambda_1(t_1)$.

Take T such that $T\tau^u(\Lambda_1(t_1))>4$. The thickness varies continuously for C^2 changes of F_t, [N4, Proposition 6.2], so there is an interval I about t_1 such that for t in I, $T\tau^u(\Lambda_1(t))>2$. By the results of Sect. 7, for large enough n there is a t_n^* in I and another hyperbolic set for F_t^n, $\Lambda_n(t_n^*)\subset B_n$, with $\tau^s(\Lambda_n(t_n^*))>T$. There is a fixed point of F_t^n in $\Lambda_n(t_n^*)$, $p_n(t_n^*)$. Take $t=t_n^{**}$ with $t_1<t_n^{**}<t_n$ such that $F_t^n(B_n)$ does not

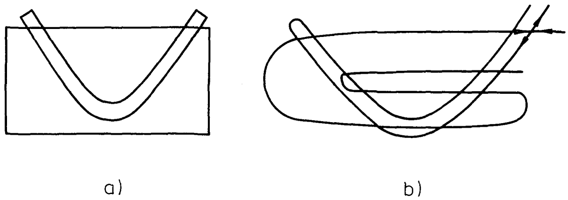


Fig. 10 a. The image $F_t^n(B_n)$ is shown with B_n for $t=t_n^{**}$: **b** The position of the stable and unstable manifolds for p_n are shown for $t=t_n^{**}$

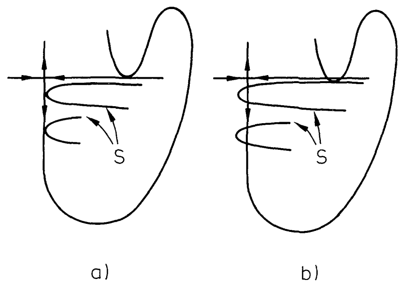


Fig. 11a and b. The stable and unstable manifolds for P_t are shown. The extra pieces of the stable manifold are labelled S. **a** $t=t_0$; **b** $t=t_1$

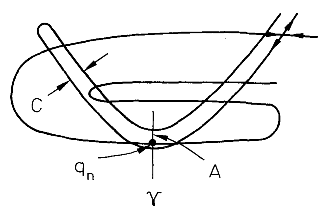


Fig. 12. "A" labels the overlap of $\gamma \cap W^s(\Lambda_n'(t_n^{**}))$ and $\gamma \cap W^u(\Lambda_n'(t_n^{**}))$ which is of size about $wd_2\mu^m/4$. The width "C" between pieces of $W^u(p_n(t_n^{**}))$ is about $wd_2\mu^m$

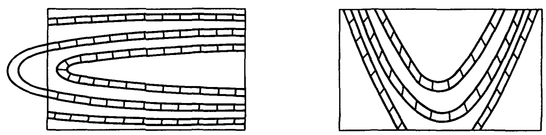
come all the way out the bottom of B_n and in fact $W^u(p_n(t), F_t^n)$ comes above but within $wd_2\mu^m/4$ of a tangency with $W^s(p_n(t), F_t^n)$. See Fig. 12. Note the first tangency of $W^s(p_n(s), F_s^n)$ and $W^u(p_n(s), F_s^n)$ occurs for $t_n^{**} < s < t_n$, when there are still other points of transverse intersection of $W^u(p_n(s), F_s^n)$ and $W^s(p_n(s), F_s^n)$. We show below that there is a new smaller hyperbolic set for F_t^n with $t=t_n^{**}$, $\Lambda_n'(t_n^{**})$, with

(8.2a) $$\tau^s(\Lambda_n'(t_n^{**})) > T,$$

(8.2b) the Cantor set fo stable and unstable manifolds for $\Lambda_n'(t_n^{**})$ overlap near the bottom of $F_t^n(B_n)$ for $t=t_n^{**}$,

and

(8.2c) the stable and unstable manifolds of $\Lambda_n'(t_n^{**})$ and $\Lambda_1(t_n^{**})$ intersect transversally.



a) b)

Fig. 13. (**a**) The shaded region is $B(-J)$; (**b**) The shaded region is $B(J)$

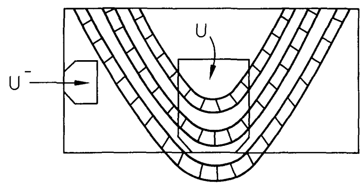


Fig. 14. U^- is $F_t^{-n}(U)$

Using (8.2c), [N4, Lemma 8] proves there is a larger hyperbolic set $\Lambda(t_n^{**}) \supset \Lambda_1(t_n^{**}) \cup \Lambda_n'(t_n^{**})$. Then $\tau^s(\Lambda(t_n^{**}))\tau^u(\Lambda(t_n^{**})) \geqq \tau^s(\Lambda_n'(t_n^{**}))\tau^u(\Lambda_1(t_n^{**})) > 1$.

To use the thickness to show there is a persistent tangency, we need a C^1 curve γ on which the minima of W^u relative to W^s occurs. Thus γ is a curve of potential tangencies. If the manifolds $W^s(\Lambda(t_n^{**}))$ could be extended to a C^2 foliation, the existence of γ would be easy: we could take C^2 coordinates such that these manifolds were straight horizontal lines and look at the minima of $W^s(\Lambda(t_n^{**}))$ in these coordinates. The argument [N4, Lemma 9 and pp. 119–120] modifies this argument so it works with only a foliation whith C^1 tangent vectors. The sketch of this proof is as follows. The Cantor set of unstable manifolds, $W^s(\Lambda(t_n^{**}))$ can be extended near q_n [a point on $W^s(p_n)$ near the minimum of $W^u(p_n)$, see Fig. 12] to a foliation $\mathscr{F}^u$ with C^1 tangent vectors. Similarly $W^s(\Lambda(t_n^{**}))$ can be extended near q_n to $\mathscr{F}^s$. The foliations $\mathscr{F}^u$ and $\mathscr{F}^s$ will have nondegenerate tangencies and so intersect along a C^1 curve γ. It then follows that for a small enough curve γ_ε around q_n, $\tau(W^s(\Lambda(t_n^{**}))\cap\gamma_\varepsilon)\tau(W^u(\Lambda(t_n^{**}))\cap\gamma_\varepsilon) > 1$, since the thickness can be measured along any C^1 curve. Therefore these two Cantor sets intersect and $W^s(\Lambda(t_n^{**}))$ has a persistent tangency with $W^u(\Lambda(t_n^{**}))$ proving (8.1d).

Fix $t = t_n^{**}$, which is chosen as above. We end this section by showing there is a hyperbolic basic set $\Lambda_n'(t)$ for F_t^n satisfying (8.2a–c). This clarifies the last paragraph of the proof of [N4, Lemma 7] on p. 138.

Let $B(-J) = \bigcap_{j=0}^{-J} F_t^{nj}(B_n)$ and $B(J) = \bigcap_{j=0}^{J} F_t^{nj}(B_n)$. The set $B(J)$ is made up of "vertical" strips of width about $wd_2^J\mu^{mJ}$ and $B(-J)$ of "horizontal" strips of width about $wC\lambda^{-m-mJ}$. See Fig. 13.

The idea is to remove the part of B_n near the bottom where the slope of $\partial F_t^n/\partial y$ is less than δ_n, where $\delta_n = 2[awd_2\mu^m/4]^{1/2}\eta^{-1}$. If any part of a component of $B(-J)\cap B(J)$ is removed, then the whole component should be removed. Let the set removed be denoted by U, and let $B_n' = B_n - U$. Let $B'(J) = \bigcap_{j=0}^{J} F_t^{nj}(B_n')$, $B'(-J)$

$= \bigcap_{j=0}^{-J} F_t^{nj}(B'_n)$, and $B'(-J,J)=B'(-J) \cap B'(J)$. Then $B'_n(-J,J)$ is made up of boxes. In $B'_n(-J,J)$, $\partial F_t/\partial y$ has slope greater than δ_n. Therefore the maximal invariant set in B'_n, $\Lambda'_n(t)=\bigcap_j F_t^{nj}(B'_n)=\bigcap_j F_t^{nj} B'_n(-J,J)$, has a hyperbolic structure by the proof of Proposition 7.8. [Note the tangencies in $F_t^n(U)$ are also removed.] For J large enough, the height of the hook of a strip in $B(J)$ below the intersection with B'_n is less than $g'=w10d_2\mu^m=h10d_2(d_1-\alpha)(\mu\lambda)^m$. The length left in B'_n is approximately as before. Therefore the argument as before shows that $\tau^s(\Lambda'_n(t)>T$, condition (8.2a). The Cantor sets of stable and unstable manifolds overlap because $p_n(t)$ is in $\Lambda'_n(t)$ and the choice of t_n^{**}, condition (8.2b).

Condition (8.2c) follows by looking at F_t^n on $W^u(p_n)$. Each iterate extends a local unstable manifold further up until it crosses $W^s(P_t)$ transversally. See Figs. 6, 11, and 12. Similarly the manifolds $W^s(p_n)$ cut across $W^u(P_t)$ which cups around $\Lambda'_n(t_n^{**})$. See Fig. 6 and [N4, p. 134]. This completes the proof of (8.2), Proposition 8.1, and Theorem D.

8.3. Remark. For maps of the real line, $f(x)$, [or their corresponding graphs $F(x,y) = (y, f(y))$] it is also possible to create hyperbolic sets with τ^s as large as desired. However for maps of the interval $\tau^u=0$, so the product $\tau^s\tau^u=0$ is never greater than one. Therefore it is impossible to create persistent tangencies for maps of the interval. See [V] for a discussion of the comparison of maps of the interval and maps of the plane.

9. Infinitely Many Sinks: Theorem E

By the assumptions of the theorem, there is an interval of parameter values J such that for each t in J there is a nondegenerate tangency of $W^s(\Lambda(t), F_t)$ and $W^u(\Lambda(t), F_t)$. Because $\Lambda(t)$ is a hyperbolic set with a dense orbit, the manifolds $W^s(P_t, F_t)$ and $W^u(P_t, F_t)$ are dense in $W^s(\Lambda(t), F_t)$ and $W^u(\Lambda(t), F_t)$, respectively. Because $W^s(\Lambda(t), F_t)$ and $W^u(\Lambda(t), F_t)$ have a tangency for each t in J and because this tangency changes location by assumption (iv) in the definition of nondegenerate creation of homoclinic intersection, it follows that there is a dense set of parameter values $J' \subset J$ such that for t in J' the manifolds $W^s(P_t, F_t)$ and $W^u(P_t, F_t)$ have nondegenerate tangencies. By Proposition 3.3 and the argument of Sect. 5, the set $J_1 \subset J$ with at least one sink is dense. It is also open. Since $J' \cap J_1$ is dense, repeating the argument proves that the set J_2 with at least two sinks is dense and open. By induction, the set J_k with at least k sinks is dense and open. Therefore $J_\infty = \bigcap \{J_k : k=1,2,\dots\}$ is a residual subset of J in the sense of Baire category. For t in J_∞, F_t has infinitely many sinks.

Acknowledgement. I would like to thank R. Devaney and P. Blanchard for their hospitality at Boston University and their encouragement to write this paper. I would like to thank J. Franks for useful conversations in the refinement of the arguments.

References

[ACHM] Aronson, D.G., Chory, M.A., Hall, G.R., McGehee, R.P.: Bifurcation from an invariant circle for two-parameter families of maps of the plane: a computer-assisted study. Commun. Math. Phys. **83**, 303–354 (1982)

[CJ] Curry, J.H., Johnson, J.R.: On the rate of approach to homoclinic tangency. Preprint, University of Minnesota, 1982

[GS] Gavrilov, N.K., Silnikov, L.P.: On the three dimensional dynamical systems close to a system with a structurally unstable homoclinic curve. I. Math. USSR Sbornik **17**, 467–485 (1972); II. Math. USSR Sbornik **19**, 139–156 (1973)

[GH] Guckenheimer, J., Holmes, P.: Nonlinear oscillations, dynamical systems and bifurcation of vector fields. Applied Math. Sci. **42**. Berlin, Heidelberg, New York: Springer 1983 1983

[H] Hénon, M.: A two dimensional mapping with a strange attractor. Commun. Math. Phys. **50**, 69–78 (1976)

[HP] Hirsch, M., Pugh, C.: Stable manifolds and hyperbolic sets. Proc. Am. Math. Soc. Symp. Pure Math. **14**, 133–163 (1970)

[HM] Holmes, P., Marsden, J.: A partial differential equation with infinitely many periodic orbits. Archive Rational Mech. Anal. **76**, 135–166 (1981)

[L] Levi, M.: Qualitative analysis of periodically forced relaxation oscillations. Memoirs Amer. Math. Soc. **32**, 244 (1981)

[Mis] Misiurewicz, M.: Strange attractors for the Lozi mappings, nonlinear dynamics. Helleman, R. (ed.). Annals of New York Academy of Science **357** (1980)

[M] Moser, J.: Stable and random motions in dynamical systems. Ann. of Math. Studies **77**, Princeton, NJ: Princeton University Press 1973

[N1] Newhouse, S.: Nondensity of axiom $A(a)$ on S^2. Proc. A.M.S. Symp. in Pure Math. **14**, 191–203 (1970)

[N2] Newhouse, S.: Diffeomorphisms with infinitely many sinks. Topology **13**, 9–18 (1974)

[N3] Newhouse, S.: Quasi-elliptic periodic points in conservative dynamical systems. Am. J. Math. **99**, 1061–1087 (1977)

[N4] Newhouse, S.: The abundance of wild hyperbolic sets and non-smooth stable sets for diffeomorphisms. Publ. Math. IHES **50**, 101–151 (1979)

[N5] Newhouse, S.: Lectures on dynamical systems. Progress in Math. **8**, 1–114. Boston: Birkhäuser 1980

[V] Van Strien, S.: On the bifurcations creating horseshoes. Lecture Notes in Mathematics, Vol. 898. Rand, D., Young, L.-S. (eds.), pp. 316–351. Berlin, Heidelberg, New York: Springer 1981

Communicated by O. E. Lanford

Received February 7, 1983; in revised form May 16, 1983