Commun. Math. Phys. 67, 137–146 (1979)

Communications in Mathematical Physics

Shift Automorphisms in the Hénon Mapping

R. Devaney* and Z. Nitecki

Department of Mathematics, Tufts University, Medford, Massachusetts 02155, USA

Abstract. We investigate the global behavior of the quadratic diffeomorphism of the plane given by $H(x, y) = (1 + y - Ax^2, Bx)$. Numerical work by Hénon, Curry, and Feit indicate that, for certain values of the parameters, this mapping admits a "strange attractor". Here we show that, for A small enough, all points in the plane eventually move to infinity under iteration of H. On the other hand, when A is large enough, the nonwandering set of H is topologically conjugate to the shift automorphism on two symbols.

Several numerical studies have recently appeared [3, 4, 7, 8] on the dynamics of the diffeomorphisms of the plane

$$H(X, Y) = (1 + Y - AX^2, BX)\ .$$

Interest in these maps [12, 14, 5] has been prompted by Hénon's numerical evidence [8] for a "strange attractor" when $A = 1.4$, $B = 0.3$. Feit [4] has shown, for $A > 0$ and $0 < B < 1$, that the non-wandering set $\Omega(H)$ is contained in a compact set, and that all points outside this set escape to infinity. Curry [3] has shown that, for Hénon's values of the parameters, one of the fixed points has a topologically transverse homoclinic orbit, and hence that there is a horseshoe embedded in the dynamics of the map.

The present note is intended to clarify the behavior of the mapping H for parameter values far from those where "strange attractors" have been observed. Hénon and Feit have noted that for $B = 0.3$ and A outside a certain interval (roughly $[-0.12, 2.67]$) no attractors are observed; numerically, all points seem to escape to infinity. We exhibit, for any $B \neq 0$, a pair of A values, $A_0 < 0 < A_2$, such that the non-wandering set $\Omega(H)$ is empty for $A < A_0$, but for $A > A_2$, $\Omega(H)$ is the zero-dimensional basic set obtained from Smale's horseshoe construction [9, 11, 13]. We begin by rewriting the map in a more convenient form; then we establish Feit's result (for all $A, B \neq 0$) in a version more suited to our purposes, by

* Partially supported by NSF Grant MCS 77-00430

0010-3616/79/0067/0137/$02.00

constructing a filtration; and finally, for $A > A_2$ we exhibit the elements of the horseshoe construction in our compact set.

Hénon notes that the map H represents one canonical form for quadratic maps with constant Jacobian determinant. We will find it convenient to consider the alternate canonical form for such maps

$$F(x, y) = (A + By - x^2, x) .$$

It is easily verified that for A and B both nonzero, the linear change of coordinates

$$X = x/A \qquad Y = By/A$$

gives a topological conjugacy between H and F, with the parameter values A and B unchanged. For Hénon's map H, the parameter value $A = 0$ gives a linear map, while in our map $A = 0$ has no special significance. In fact, our results are established without the restrictions $A > 0$, $0 < B < 1$ imposed by earlier papers; we assume only that $B \neq 0$. Thus our analysis includes the orientation-preserving cases $B < 0$, and the area-preserving cases $B = \pm 1$, which were considered by Hénon in an earlier numerical study [6]. Actually, the B-values with absolute value greater than one do not exhibit new behaviour, since the inverse map

$$F^{-1}(x, y) = (y, (x - A + y^2)/B)$$

with given parameter values $A = a$, $B = b \neq 0$ is conjugate to the forward map F with $A = a/b^2$, $B = 1/b$ by the linear change of variables

$$x \to -by \qquad y \to -bx .$$

To state our result, we fix B and define three crucial A-values

$$A_0 = -(1 + |B|)^2/4$$

$$A_1 = 2(1 + |B|)^2$$

$$A_2 = (5 + 2\sqrt{5})(1 + |B|)^2/4$$

and, for any particular A-value, we define $R = R(A)$ by

$$R = (1/2)\{1 + |B| + [(1 + |B|)^2 + 4A]^{1/2}\} .$$

With this notation, our results are summarized in the following theorem.

Theorem. i) *For $A < A_0$, $\Omega(F) = \phi$.*

ii) *For $A \geqq A_0$, $\Omega(F)$ is contained in the square $S = \{(x, y) \mid |x| \leqq R, |y| \leqq R\}$.*

iii) *For $A \geqq A_1$, $\Lambda = \bigcap_{n \in \mathbb{Z}} F^n(S)$ is a topological horseshoe; for $B \neq 0$, there is a continuous semi-conjugacy of $\Omega(F) \subset \Lambda$ onto the 2-shift.*

iv) *For $A > A_2$, $\Lambda = \Omega(F)$ has a hyperbolic structure and is conjugate to the 2-shift.*

The value A_0 is, as Hénon remarked, precisely the A-value at which the first fixed point of F^2 appears. Thus, statement i) above follows from the Brouwer translation theorem [1, 2]; however, we shall give a direct proof of this fact. On the other hand, the values we give for A_1 and A_2 are somewhat larger than the experimental values given by Hénon and Feit (when $B=0.3$, $A_1=3.38$, $A_2=4.00$). In fact, they are clearly not the lowest values that yield the desired conclusions, but they yield these conclusions relatively easily.

The proofs of statements i)–iii) rely on the following technical lemmas:

Lemma 1. a) *R is real if and only if $A \geqq A_0$. In this case, R is positive and equals the larger root of*

$$R^2-(|B|+1)R-A=0\,.$$

b) $A-|B|R>R$ *if and only if* $A>A_1$.

Proof. a) is trivial; it implies that $A+|B|R=R^2-R$ and the first inequality of b) is equivalent to $R>2(1+|B|)$. Substituting this into the definition of R and solving for A gives the second inequality of b). □

We will find it convenient to denote the image of the point (x_0, y_0) by $(x_1, y_1) = F(x_0, y_0)$, and use negative subscripts for pre-images.

Lemma 2. a) *The image under F of the horizontal strip $|y_0| \leqq C$ is the region bounded by the two parabolas*

$$A-|B|C-y_1^2 \leqq x_1 \leqq A+|B|C-y_1^2\,.$$

The image under F of the vertical strip $|x_0| \leqq C$ is the horizontal strip $|y_1| \leqq C$.

b) *The inverse image of the vertical strip $|x_0| \leqq C$ is the region bounded by the two parabolas*

$$-C-A-x_{-1}^2 \leqq By_{-1} \leqq C-A-x_{-1}^2\,.$$

The inverse image of the horizontal strip $|y_0| \leqq C$ is the vertical strip $|x_{-1}| \leqq C$.

Proof. These are straightforward calculations. □

In the following, we interpret $\min(a, R)$ or $\max(a, R)$ with R complex as equal to a.

Lemma 3. a) *If $x_0 \leqq \min(-|y_0|, -R)$, then $x_1 \leqq x_0$, with equality only for $x_0=-R$, $y_0=\pm R$.*

b) *If $x_0 \geqq -|y_0|$ and $By_0 \geqq \max(0, |B|R)$, then $By_{-1} \geqq By_0$ and $|y_{-1}| \geqq |y_0|$, with equality only for $x_0=-R$, $y_0=\pm R$.*

Proof. For a), by the definition of x_1 and, in the last inequality below, our hypothesis on x_0,

$$\begin{aligned} x_1-x_0 &= A+By_0-x_0^2-x_0 \\ &\leqq A+|B|\,|y_0|-x_0^2-x_0 \\ &\leqq A-(|B|+1)x_0-x_0^2\,. \end{aligned}$$

The last expression is zero for .

$$x_0=-(1+|B|)/2 \pm [(1+|B|)^2+4A]^{1/2}/2\,.$$

If R is complex, so is x_0, and the expression above is negative for all x_0, whereas when R is real, the lesser root is $x_0=-R$, so that the expression remains negative

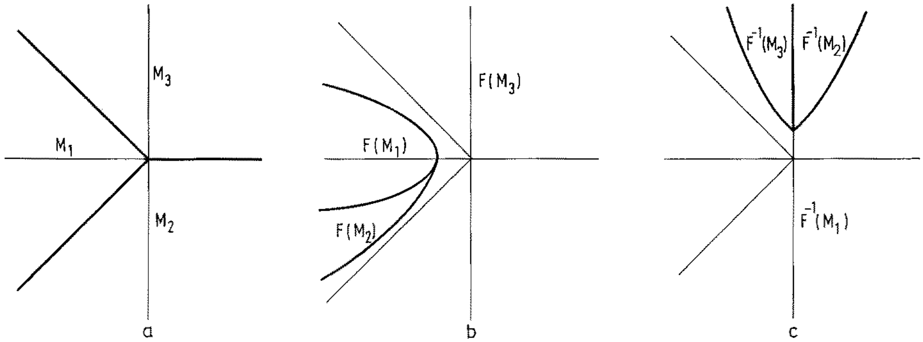

Fig. 1a–c. Filtration for $A<A_0$. Situation shown is for $B>0$; $B<0$ is obtained by reflection about the x-axis. **a** Filtration. **b** Image under F. **c** Image under F^{-1}

for $x_0<-R$. On the other hand, when $x_0=-R$ but $|y_0|<R$, the last inequality above is strict, so that equality only holds for $x_0=-|y_0|=-R$.

The proof of b) is similar. Consider

$$\begin{aligned} B(y_{-1}-y_0) &= y_0^2+x_0-A-By_0 \\ &\geqq y_0^2-(1+|B|)y_0-A\,. \end{aligned}$$

The last expression is positive provided $|y_0|>\max(0,R)$; if B is positive, this says $y_{-1}>y_0\geqq 0$, whereas for B negative, it says $y_{-1}<y_0\leqq 0$. The equality statements follow as before. □

Now, to prove statement i) of the theorem, we define a partition of the plane by

$$M_1=\{(x,y)\,|\,x\leqq -|y|\}$$

$$M_2=\{(x,y)\,|\,x\geqq -|y| \text{ and } By\leqq 0\}$$

$$M_3=\{(x,y)\,|\,x\geqq -|y| \text{ and } By\geqq 0\}$$

(see Fig. 1a).

Proposition 1. *For* $A<A_0$,

a) $F(M_1\cup M_2)\subset$ *interior* M_1.
b) x *is strictly decreasing along* F*-orbits in* M_1.
c) $F^{-1}(M_2\cup M_3)\subset$ *interior* M_3.
d) $|y|$ *is strictly increasing along* F^{-1}*-orbits in* M_3.

Proof. By Lemma 3a, if $(x_0, y_0)\in M_1$, then $y_1=x_0>x_1$. The inequality is strict because R is complex. Also, since $x_0\leqq 0$, $y_1=-|y_1|$. This shows $F(M_1)\subset$ interior M_1 and statement b). Moreover, by Lemma 2a, the x-axis maps to a parabola opening left with vertex at $(A,0)$; this lies to the left of the boundary of M_1. Furthermore, the image of the line $By=-\varepsilon$ is a parabola to the left of the previous one, so that $F(M_2)$ lies to the left of the boundary of M_1 (see Fig. 1b).

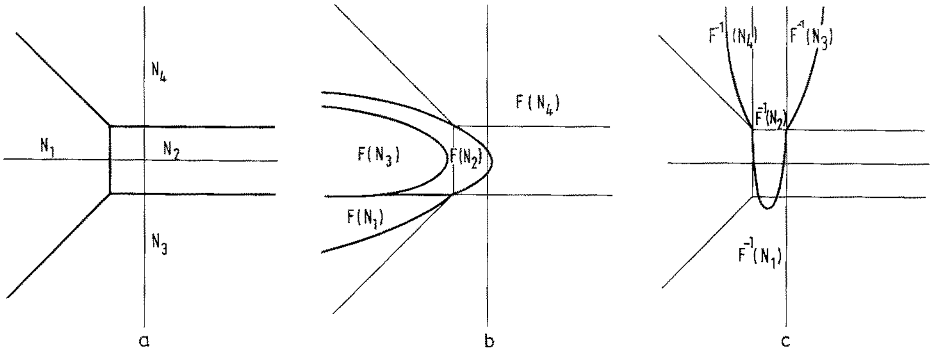

Fig. 2a–c. Partition for $A \geqq A_0$. Situation shown is for $B > 0$. **a** Partition. **b** Image under F. **c** Image under F^{-1}

The proof of c) and d) is similar. If $(x_0, y_0) \in M_3$, then by Lemma 3b,

$$|y_{-1}| > |y_0| = |x_{-1}|$$

and

$$By_{-1} > By_0 \geqq 0 .$$

An analysis similar to the above, using Lemma 2b, completes the proof of c) (see Fig. 1c). □

Proposition 1 gives a filtration for F with Liapunov functions for the extreme sets M_1, M_3, and $M_2 \cap F(M_2) = \phi$, so that $\Omega(F) = \phi$ for $A < A_0$. Statement ii) of the theorem is proven by an analogous construction, with the positive x-axis above expanded into another element of the filtration. Specifically, when R is real (i.e., $A \geqq A_0$), define four sets by Fig. 2a

$$N_1 = \{(x, y) | x \leqq \min(-|y|, R)\}$$

$$N_2 = \{(x, y) | x \geqq -R, |y| \leqq R\}$$

$$N_3 = \{(x, y) | x \geqq -|y|, By \leqq |B|R\}$$

$$N_4 = \{(x, y) | x \geqq -|y|, By \geqq |B|R\} .$$

Proposition 2. *For* $A \geqq A_0$,

a) $F(N_1) \subset N_1$.

b) $F(N_2 \cup N_3) \subset N_1 \cup N_2$.

c) *x is decreasing along F-orbits in N_1 (strictly decreasing except at the two points $x = -|y| = -R$).*

d) $F^{-1}(N_3 \cup N_4) \subset N_4$.

e) $F^{-1}(N_2) \subset N_2 \cup N_3 \cup N_4$.

f) *$|y|$ is increasing along F^{-1}-orbits in N_4 (strictly increasing except at the point $(-R, R\,\mathrm{sign}(B))$).*

Proof. The only statements whose proof differs from Proposition 1 are b) and e). b) is proven by invoking Lemma 2a with $C = R$, and by noting that the right boundary of $F(N_2)$ intersects $x = -R$ at $|y| = R$ by Lemma 1a (see Fig. 2b).

a b

Fig. 3a and b. Image of S. **a** $A_0 \leqq A < A_1$. **b** $A_1 \leqq A$

Similarly, we note that the boundary curves $-x_0 = |y_0| > R$ of N_1 map under F^{-1} to curves in the interior of N_4. This follows from Lemma 3b and proves d). Also, the line segment $x_0 = -R$, $|y_0| \leqq R$ maps to the parabolic segment

$$By_{-1} = x_{-1}^2 - R - A, \qquad |x_{-1}| \leqq R$$

which is disjoint from N_1. This proves e) (see Fig. 2c).

Proposition 1 does not, strictly speaking, give a filtration for F, since one of the two left corners of the square S maps to the boundary of N_1 under F, and to the boundary of N_4 under F^{-1}. This could be taken care of by a slight modification of N_1 and N_2, but for our purposes, it suffices to note that, by c) and f), all points in the interior of $N_1 \cup N_4$ are wandering. In fact, they escape to infinity in at least one time direction. By b) and e), this implies that $\Omega(F) \subset N_2 \cap F^{-1}(N_2)$. But from Lemma 2, we see immediately that

$$N_2 \cap F^{-1}(N_2) \subset S = \{(x, y) \mid |x| \leqq R, |y| \leqq R\}$$

thus proving statement ii) of the Theorem.

To prove statement iii) of the Theorem, we note that, by Lemma 2a, $F(S)$ is the region bounded by the parabolas

$$x_1 = A \pm |B|R - y_1^2, \qquad |y_1| \leqq R$$

and the two horizontal line segments

$$y_1 = \pm R, \; -(1 + 2|B|)R \leqq x_1 \leqq -R.$$

The latter follows from the identities

$$A + |B|R - R^2 = -R$$

$$A - |B|R - R^2 = -(1 + 2|B|)R$$

which follow immediately from Lemma 1. The vertex of the left boundary of $F(S)$ is at $(A + |B|R, 0)$; as A increases, this moves to the right. When A passes A_1, it crosses the right edge of S, by Lemma 1b (see Fig. 3). Thus, for $A > A_1$, we have the topological part of the horseshoe: the image of any horizontal line segment in S is a parabola which cuts across S in two segments. By a standard analysis [9, 11, 13] points in the invariant set $\Lambda = \bigcap_{-\infty \leqq n \leqq \infty} F^n(S)$ can be coded by a bisequence of 0's and 1's according to which of the two components of $S_{-1} = S \cap F(S)$ contain successive backward iterates and which of the two components of $S_1 = S \cap F^{-1}(S)$ contain successive forward iterates. The coding gives a continuous orbit-preserving map of Λ onto the Cantor set underlying the shift automorphism on two symbols. To see that $\Omega(F)\subset\Lambda$ maps onto the shift space, we note that every periodic sequence corresponds to a nested intersection of closed disks, and hence to at least one periodic orbit of F; but the closure of the set of these periodic orbits is a set of non-wandering points in Λ mapping onto the shift space. This establishes statement iii) of the Theorem.

To prove statement iv), we first note that the matrix of partial derivatives of F

$$JF=\begin{pmatrix}-2x & B\\ 1 & 0\end{pmatrix}$$

and of F^{-1}

$$JF^{-1}=\begin{pmatrix}1/B & 2x/B\\ 0 & 1\end{pmatrix}$$

are independent of A and y. Thus, for $B\neq 0$ fixed, the hyperbolicity of an invariant set depends only on the projection of this set onto the x-axis. We will show that orbits which stay away from a band about the x-axis (whose width depends on B) have two constant bundles of sectors

$$S_\lambda^+=\{(\xi,\eta)\,|\,|\xi|\geqq\lambda|\eta|\}$$

$$S_\lambda^-=\{(\xi,\eta)\,|\,\lambda|\xi|\leqq|\eta|\}$$

with $\lambda>1$ which are invariant, respectively, under the Jacobians JF and JF^{-1}. We will then show that the invariant set $\Lambda=\cap F^n(S)$ is disjoint from this band when $A>A_2$.

Lemma 4. *Suppose that, for some $\lambda>1$, x satisfies*

$$|x|\geqq\lambda(1+|B|)/2\,.$$

Then: a) *For any vector $(\xi_0,\eta_0)\in S_\lambda^+$, the vector $(\xi_1,\eta_1)=JF_x(\xi_0,\eta_0)$ satisfies $|\xi_1|>\lambda|\xi_0|$.*

b) *For any vector $(\xi_0,\eta_0)\in S_\lambda^-$, $(\xi_{-1},\eta_{-1})=JF_x^{-1}(\xi_0,\eta_0)$ satisfies $|\eta_{-1}|\geqq\lambda|\eta_0|$.*

Proof. By hypothesis,

$$2|x|\geqq\lambda+\lambda|B|$$

so that

1) $2|x|-|B|/\lambda>2|x|-\lambda|B|\geqq\lambda\,,$

2) $2|x|-\lambda\geqq\lambda|B|\,.$

To show a), we use the formula for JF_x, the inequality $|a+b|\geqq|a|-|b|$, the hypothesis $|\eta_0|\leqq|\xi_0|/\lambda$, and 1) in succession to conclude

$$\begin{aligned}|\xi_1|&=|-2x\xi_0+B\eta_0|\geqq 2|x|\,|\xi_0|-|B|\,|\eta_0|\\ &\geqq(2|x|-|B|/\lambda)|\xi_0|>\lambda|\xi_0|\,.\end{aligned}$$

Similarly, to show b), we use the formula for JF_x^{-1} and 2) to obtain

$$\begin{aligned}|\eta_{-1}| &= |\xi_0 + 2x\eta_0|/|B| \\ &\geqq (2|x| - \lambda)|\eta_0|/|B| \\ &\geqq \lambda|\eta_0| . \quad \square\end{aligned}$$

As a consequence, we can prove the following.

Proposition 3. *If* (x_0, y_0) *and* $(x_1, y_1) = F(x_0, y_0)$ *both satisfy*

$$|x| > \lambda(1 + |B|)/2$$

for some $\lambda > 1$, *then*

a) *for any vector* $(\xi_0, \eta_0) \in S_\lambda^+$, $(\xi_1, \eta_1) = JF_{x_0}(\xi_0, \eta_0)$ *belongs to* S_λ^+, *and*

$$|(\xi_1, \eta_1)| \geqq \lambda |(\xi_0, \eta_0)| .$$

b) *For any vector* $(\xi_1, \eta_1) \in S_\lambda^-$, $(\xi_0, \eta_0) = JF_x^{-1}(\xi_1, \eta_1)$ *belongs to* S_λ^- *and*

$$\lambda |(\xi_1, \eta_1)| \leqq |(\xi_0, \eta_0)| .$$

Proof. Our notation above is such that in both a) and b),

$$(\xi_1, \eta_1) = JF_{x_0}(\xi_0, \eta_0) .$$

In particular, $\eta_1 = \xi_0$. To show a), we invoke Lemma 4a to get

$$\lambda|\eta_1| = \lambda|\xi_0| \leqq |\xi_1|$$

so that $(\xi_1, \eta_1) \in S_\lambda^+$, while by the definition of S_λ^+

$$\lambda|\eta_0| \leqq |\xi_0| = |\eta_1| .$$

This proves a). Similarly, to establish b), we use Lemma 4b to conclude

$$|\eta_0| \geqq \lambda|\eta_1| = \lambda|\xi_0|$$

and the definition of S_λ^- to conclude

$$|\xi_0| = |\eta_1| \geqq \lambda|\xi_1| . \quad \square$$

Our last step in establishing statement iv) is to verify the hypotheses of Proposition 3 for all points of $\Lambda = \cap F^n(S)$ when $A > A_2$.

Proposition 4. *If* $A > A_2$, *there exists* $\lambda > 1$ *such that*

$$|x| \geqq \lambda(1 + |B|)/2$$

for all $(x, y) \in S \cap F^{-1}(S)$.

Proof. From Lemma 1b, we see that the pre-image $F^{-1}(S)$ lies outside the parabola

$$By_1 = R - A + x_1^2$$

which intersects the edge of S at a pair of points with x-coordinates $\pm x_*$ where

$$x_*^2 = By - R + A = A - (1 + |B|)R .$$

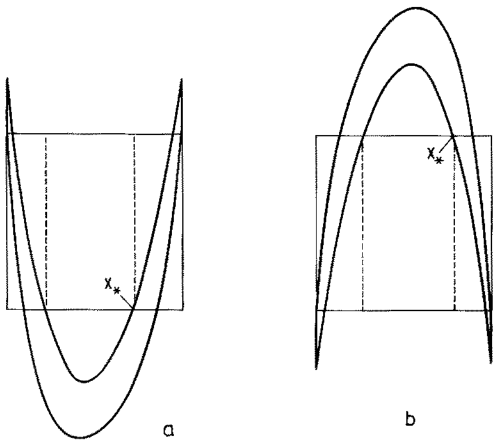

Fig. 4a and b. Preimage of S; $A>A_1$. **a** $B>0$. **b** $B<0$

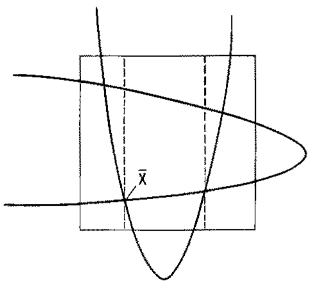

Fig. 5

It is clear that $F^{-1}(S)\cap S$ lies in the region $|x|\geqq|x_*|$ (see Fig. 4).

Let us write A in the form

$$A=k(1+|B|)^2/2\ .$$

A quick calculation from the formula for R then gives

$$R=(1+\sqrt{1+2k})(1+|B|)/2$$

so that

$$x_*^2=(k-1-\sqrt{1+2k})(1+|B|)^2/2\ .$$

For $k>0$, this quantity increases with k and the left hand factor equals 1 when $2k=5+2\sqrt{5}$ or $A=A_2$. Thus when $A>A_2$, we can take

$$\lambda^2=(k-1-\sqrt{1+2k})>1$$

to get the conclusion of the proposition, and hence hyperbolicity of λ. □

We close with the observation that the conclusion of Proposition 4 for points of Λ could be obtained with lower values of A if, instead of estimating the intersection of the edge of $F^{-1}(S)$ with the edge of S, we considered the smallest x-value, $|\bar{x}|$, among the four intersections of the edge of $F^{-1}(S)$ with the edge of $F(S)$ (see Fig. 5). This point, which is a solution of the two quadratic equations

$$By = R - A + x^2$$
$$x = A + |B|R - y^2$$

is a lower bound for $|x|$ on Λ which satisfies the hypotheses of Proposition 4 for values of A somewhat lower than A_2.

We finally remark that our theorem shows that the phenomena of the "Hénon attractor" are part of a bifurcation occurring in the creation of a horseshoe from nothing. This gives another perspective on the significance of these mappings for dynamical systems theory.

Note. After this paper was written, it came to our attention that S. Newhouse has outlined a proof of a similar result [10].

References

1. Andrea, S.A.: On homeomorphisms of the plane, and their embeddings in flows. Bull. AMS **71**, 381–383 (1965)
2. Brouwer, L.: Beweis des ebenen Translationssatzes. Math. Ann. **72**, 37–54 (1912)
3. Curry, J.: On the structure of the Hénon attractor. Preprint, National Center for Atmospheric Research (1977)
4. Feit, S.: Characteristic exponents and strange attractors. Commun. Math. Phys. **61**, 249–260 (1978)
5. Guckenheimer, J.: Bifurcations of dynamical systems. Preprint, University of California at Santa Cruz (1978)
6. Hénon, M.: Numerical study of quadratic area preserving mappings. Q. Appl. Math. **27**, 291–312 (1969)
7. Hénon, M.: A two-dimensional mapping with a strange attractor. Commun. Math. Phys. **50**, 69–77 (1976)
8. Hénon, M., Pomeau, Y.: Two strange attractors with a simple structure. In: Turbulence and the Navier-Stokes equation. Lecture notes in mathematics, Vol. 565, pp. 29–68. Berlin, Heidelberg, New York: Springer 1975
9. Moser, J.: Stable and random motions in dynamical systems. In: Annals of mathematical studies 77. Princeton: Princeton University Press 1973
10. Newhouse, S.: Lectures on dynamical systems. C.I.M.E. Summer in dynamical systems, Bressanone, Italy, 1978
11. Nitecki, Z.: Differentiable dynamics. Cambridge, MA, London: MIT Press 1971
12. Ruelle, D.: Dynamical systems with turbulent behavior. In: mathematical problems in theoretical physics. Lecture notes in physics. Vol. 80, pp. 341–360. Berlin, Heidelberg, New York: Springer 1978
13. Smale, S.: Diffeomorphisms with many periodic points. In: Differential and combinatorial topology. Princeton: Princeton University Press 1965
14. Ford, J.: Computer experiments and dynamical systems. Asterisque **49**, 75–92 (1977)

Communicated by J. L. Lebowitz

Received December 12, 1978