DYNAMICAL SYSTEMS AND ERGODIC THEORY
BANACH CENTER PUBLICATIONS, VOLUME 23
PWN – POLISH SCIENTIFIC PUBLISHERS
WARSZAWA 1989

A TOPOLOGICAL VERSION OF A THEOREM OF MATHER ON SHADOWING IN MONOTONE TWIST MAPS

GLEN RICHARD HALL*

School of Mathematics, University of Minnesota
Minneapolis, MA, U.S.A.

This note is meant as an exposition of and advertisement for the topological approach to monotone twist mappings of the annulus. We give a topological proof of a theorem due to John Mather stating the existence of many orbits with different qualitative behaviors in area preserving monotone twist mappings of the annulus without invariant circles. Since our techniques are topological we can weaken the area preservation hypothesis.

(1) *Introduction.* This note is meant as an advertisement for the topological approach to monotone twist mappings of an annulus. We give a simple topological proof of a theorem which follows from a theorem of Mather [M1] which says that there exist orbits which "shadow" certain periodic orbits of area preserving monotone twist maps of the annulus without invariant circles. Mather uses variational techniques. Our arguments give only some index information about the orbits, but they allow a weakening of the area preservation condition. The techniques we use are similar to those in Boyland–Hall [BH].

(2) *Notation.* We let $A = \boldsymbol{R} \times [0, 1]$ denote the strip and we consider *monotone twist mappings* f: $A \to A$, i.e., mappings which satisfy

(1) f is a diffeomorphism preserving orientation and boundary components,

(2) for all $(x, y) \in A$, $f(x+1, y) = f(x, y)+(1, 0)$, i.e., f is the lift of a diffeomorphism of an annulus,

(3) $\partial(\pi_1(f))/\partial y > 0$,

where π_1 and π_2 are the projections of A onto $\boldsymbol{R}$ onto the x and y coordinates respectively.

* Partially supported by NSF and the Sloan Foundation.

For such a map we let, for $z \in A$,

$$\text{extended orbit}(z) = eo(z) = \{f^i(z)+(j, 0): i, j \in \mathbf{Z}\}.$$

We say that two points w_1, $w_2 \in eo(z)$ are *adjacent* if $\pi_1(w_1) < \pi_1(w_2)$ and $\{w: \pi_1(w_1) < \pi_1(w) < \pi_1(w_2)\} \cap eo(z) = \emptyset$.

We say $z \in A$ is a *p/q-periodic point* if

$$f.q(z) = z+(p, 0)$$

and following Boyland [Bd] we say that z (or its orbit) is *monotone* if

$$\text{for all } z_1, z_2 \in eo(z), \quad \text{if } \pi_1(z_1) < \pi_1(z_2) \text{ then } \pi_1(f(z_1)) < \pi_1(f(z_2)).$$

Remark. The theorems we state will also be true with "positive tilt" replacing "monotone twist" once the appropriate definitions have been made for, e.g., monotone orbits (see Boyland [Bd]).

We will study maps $f: A \to A$ satisfying the following:

CONDITION B. For every $\varepsilon > 0$ there exist $z_1, z_2 \in A$ with $\pi_2(z_1) < \varepsilon$ and $\pi_2(z_2) > 1-\varepsilon$ such that there exist $n_1, n_2 > 0$ with

$$\pi_2(f^{n_1}(z_1)) > 1-\varepsilon \quad \text{and} \quad \pi_2(f^{n_2}(z_2)) < \varepsilon$$

i.e., Condition B states that there exists orbits starting near the lower boundary which eventually get near the upper boundary and vice versa. This property was shown by Birkhoff [Bf] (see also Herman [Hn]) to hold for area preserving monotone twist maps without invariant circles, i.e., in "zones of instability". By variational techniques Mather, and by topological techniques Le Calvez (see Le Calvez [LC]) have shown that for area preserving monotone twist maps there are actually orbits whose α and ω limit sets are in opposite boundary components when there are no invariant circles.

Finally, if $f: A \to A$ is monotone twist, we let

$$\varrho_0 = \lim(\pi_1(f^n(x, 0)))/n \quad \text{and} \quad \varrho_1 = \lim(\pi_1(f^n(x, 1)))/n$$

be the rotation numbers of f on the boundary components (which are independent of the choice of x). Then we have

THEOREM A (Boyland [Bd]). *If $f: A \to A$ is a monotone twist map satisfying condition B and $p/q \in (\varrho_0, \varrho_1)$ then f has two distinct monotone p/q periodic orbits.*

Remarks. (1) If we replace "Condition B" with "area preserving" then this theorem is due to Aubry and Mather (see [Ma] for references). Boyland's version of this theorem is purely topological and hence allows, for example, periodic sinks.

(2) We could also attain points with irrational rotation number by the usual limit arguments (see Katok [K], Boyland [Bd]).

We can now state the main result of this note, which in the area preserving case follows from a theorem of Mather.

THEOREM B. *Let f: $A \to A$ be a monotone twist map satisfying Condition* B. *Suppose we are given bi-infinite sequences* $\{n_i\colon n_i \in \mathbf{Z}^+,\ i \in \mathbf{Z}\}$, $\{p_i/q_i\colon p_i/q_i \in (\varrho_0, \varrho_1),\ i \in \mathbf{Z}$ *and* $p_i/q_i \neq p_{i+1}/q_{i+1}$ *for all* $i\}$ *and* $\{z_i \in A\colon z_i$ *is a monotone* p_i/q_i*-periodic point of* f, $i \in \mathbf{Z}\}$. *Then there exists a sequence* $\{k_i\colon k_i \in \mathbf{Z}^+,\ i \in \mathbf{Z}\}$ *and a point* $\zeta \in A$ *such that for each* $s \geqslant 0$:

(*s) *There exist adjacent points* $w_0, w_1 \in eo(z_s)$ *with*

$$\pi_1(f^j(w_0)) < \pi_1(f^j(\zeta)) < \pi_1(f^j(w_1))$$

for

$$\sum_{i=0}^{s-1} n_i + \sum_{i=0}^{s-1} k_i \leqslant j \leqslant \sum_{i=0}^{s} n_i + \sum_{i=0}^{s-1} k_i \quad (\text{for } s = 0 \text{ use } 0 \leqslant j \leqslant n_0),$$

and for each $s < 0$:

(**s) *There exist adjacent points* $w_0, w_1 \in eo(z_s)$ *with*

$$\pi_1(f^j(w_0)) < \pi_1(f^j(\zeta)) < \pi_1(f^j(w_1))$$

for

$$-\Big(\sum_{i=-1}^{s+1} n_i + \sum_{i=-1}^{s} k_i\Big) \geqslant j \geqslant -\Big(\sum_{i=-1}^{s} n_i + \sum_{i=-1}^{s} k_i\Big)$$

$$(\text{for } s = -1 \text{ use } -k_{-1} \geqslant j \geqslant -(n_{-1} + k_{-1})).$$

Remarks. (1) We will see in the proof that if there is an $\varepsilon > 0$ such that $p_i/q_i \in (\varrho_0 + \varepsilon, \varrho_1 - \varepsilon)$ for all i then we may assume that the $|k_i - (4q_i + 2q_{i+1})|$ are uniformly bounded with a constant depending only on f.

(2) If all three of the given sequences are periodic then we may choose the point ζ to be periodic (see Boyland–Hall [BH]).

(3) *A lemma.* The proof will follow by induction on s and relies on repeated application of the lemma of this section. This lemma includes the essential use of the monotone twist condition.

Notation. For $z \in A$ we let

$$I_z = \{w \in A\colon \pi_1(w) = \pi_1(z)\},$$

$$I_z^+ = \{w \in I_z\colon \pi_2(w) \geqslant \pi_2(z)\},$$

$$I_z^- = \{w \in I_z\colon \pi_2(w) \leqslant \pi_2(z)\}.$$

For $z_1, z_2 \in A$ let

$$B_{z_1 z_2} = \{w \in A\colon \pi_1(z_1) < \pi_1(w) < \pi_1(z_2)\}.$$

We say a set $C \subseteq \text{closure}\,(B_{z_1 z_2})$ is a *positive diagonal* of $B_{z_1 z_2}$ if it satisfies

(0) C is simply connected and the closure of its interior,

(1) $\partial C \cap \partial B_{z_1 z_2} \subseteq I^-_{z_1} \cup I^+_{z_2} \cup \{z\colon \pi_2(z) = 0 \text{ or } 1\}$,

(2) $\partial C \cap I^-_{z_1} \neq \emptyset$, $\partial C \cap I^+_{z_2} \neq \emptyset$.

Similarly, C is a *negative diagonal* if it satisfies (0) and

(1′) $\partial C \cap \partial B_{z_1 z_2} \subseteq I^+_{z_1} \cup I^-_{z_2} \cup \{z\colon \pi_2(z) = 0 \text{ or } 1\}$,

(2′) $\partial C \cap I^+_{z_1} \neq \emptyset$, $\partial C \cap I^-_{z_2} \neq \emptyset$.

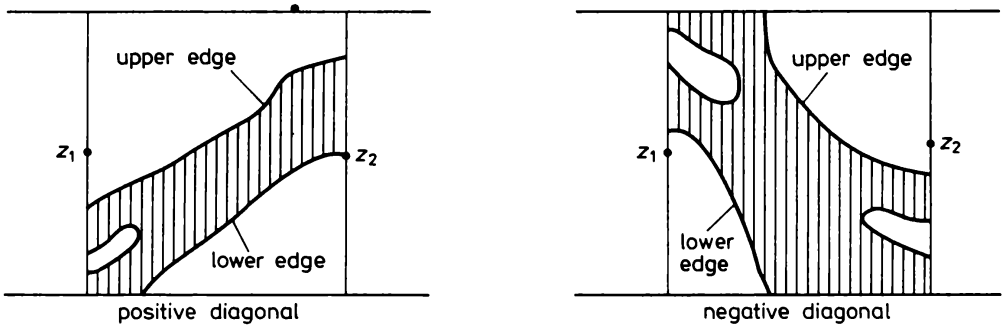

Fig. 1

The set $\partial C \cap B_{z_1 z_2}$ will have exactly two components stretching across $B_{z_1 z_2}$ (connecting $I^-_{z_1} \cup \{z\colon \pi_2(z) = 0\}$ to $I^+_{z_2} \cup \{z\colon \pi_2(z) = 1\}$ for positive diagonals) and these are called the *upper* and *lower edges* of ∂C (see Fig. 1).

LEMMA. *Suppose* $f\colon A \to A$ *is a monotone twist map,* $z_1, z_2 \in A$ *with* $\pi_1(f^i(z_1)) < \pi_1(f^i(z_2))$, $i = -1, 0, 1$, *and* $C \subseteq A$ *a positive diagonal (respectively negative diagonal) of* $B_{z_1 z_2}$. *Then* $f(C) \cap B_{f(z_1)f(z_2)}$ *(respectively* $f^{-1}(C) \cap B_{f^{-1}(z_1)f^{-1}(z_2)}$*) has a component* C_1 *which is a positive (respectively negative) diagonal.*

Moreover, if C *has upper edge contained in* $f^k(I^+_{w_0})$ *and lower edge contained in* $f^k(I^-_{w_1})$ *and* $\partial C \cap B_{z_1 z_2} \subseteq f^k(I^+_{w_0} \cup I^-_{w_1})$ *for some* w_0 *and* w_1 *with* $\pi_1(w_0) < \pi_1(w_1)$ *then we may choose* C_1 *so that its upper edge is in* $f^{k+1}(I^+_{w_0})$ *and its lower edge is in* $f^{k+1}(I^-_{w_1})$.

Remark. Of course, we could phrase the last sentence for f^{-1}.

Proof of the lemma. The first part of the lemma follows immediately from the monotone twist condition, noting that if $z \in f(I^+_{z_2})$ then $\pi_1(z) > \pi_1(f(z_2))$ and similar statements for $f(I^-_{z_1})$ and f^{-1} (see Fig. 2).

The second statement of the lemma follows from the fact that the upper edge of the component of $f(C) \cap B_{f(z_1)f(z_2)}$ with the largest y value must be in the image of the upper edge of C while the lower edge of the lowest component is in the image of the lower edge of C. The other boundary

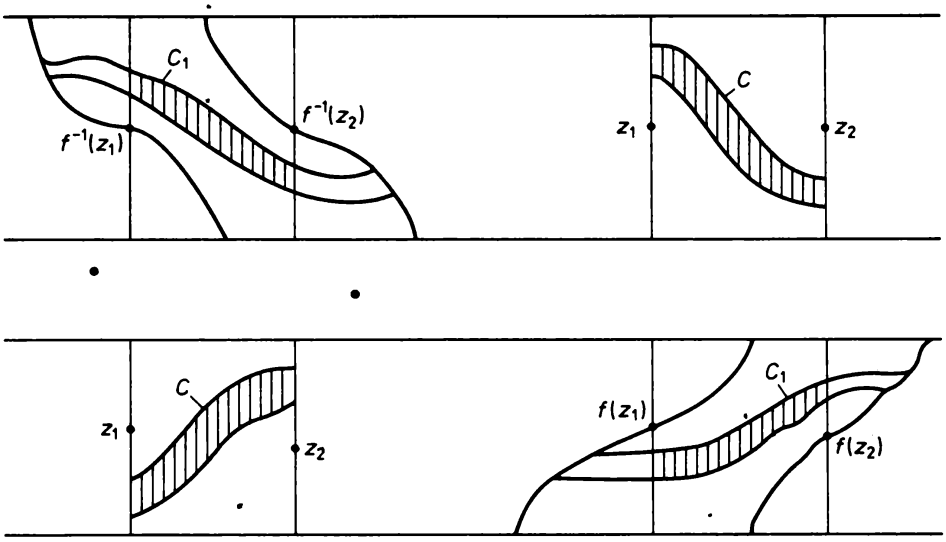

Fig. 2

components which stretch across $B_{f(z_1)f(z_2)}$ are images of loops, both of whose end points are connected to $I^-_{z_1}$ or to $I^+_{z_2}$. These loops are contained entirely in $f^k(I^+_{w_0})$ or $f^k(I^-_{w_1})$. Hence if the lower edge of a component of $f(C) \cap B_{f(z_1)f(z_2)}$ is in $f^{k+1}(I^+_{w_0})$ then the upper edge of the next component is also in $f^{k+1}(I^+_{w_0})$. Hence there is a largest component (i.e., component with points with the largest y value) for which the upper edge is in $f^{k+1}(I^+_{w_0})$ and the lower edge is in $f^{k+1}(I^-_{w_1})$-we may take this as the component C_1 (see Fig. 3). ∎

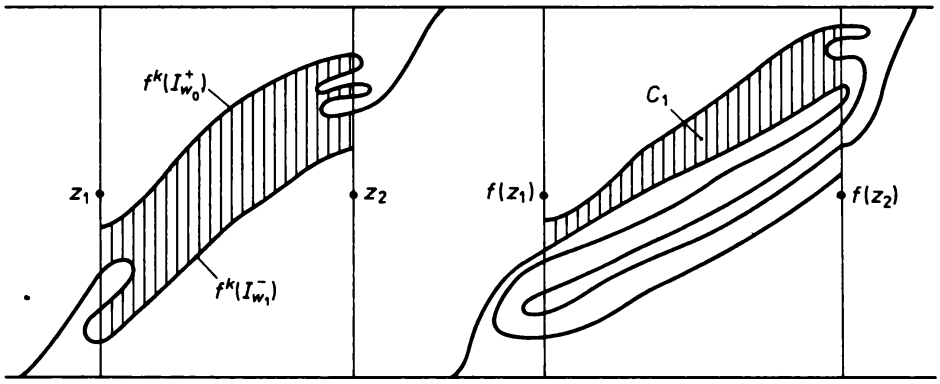

Fig. 3

(4) *Proof of Theorem* B. The proof is by induction.

We first fix $\underline{w}_0, \underline{w}_1 \in eo(z_0)$, adjacent and note that by n_0 applications of the lemma we may find a set $C_0 \subseteq A$ such that

(1) C_0 is a negative diagonal for $B_{\underline{w}_0\underline{w}_1}$,

(2) every point $\zeta \in C_0$ satisfies (*s) for $s = 0$,

(3) $f^{n_0}(C_0)$ is a positive diagonal of $B_{f^{n_0}(\underline{w}_0)f^{n_0}(\underline{w}_1)}$ with upper edge contained in $f^{n_0}(I^+_{\underline{w}_0})$, lower edge contained in $f^{n_0}(I^-_{\underline{w}_1})$ and $\partial f^{n_0}(C_0) \cap B_{f^{n_0}(\underline{w}_0)f^{n_0}(\underline{w}_1)} \subseteq f^{n_0}(I^+_{\underline{w}_0} \cup I^-_{\underline{w}_1})$ (see Fig. 4).

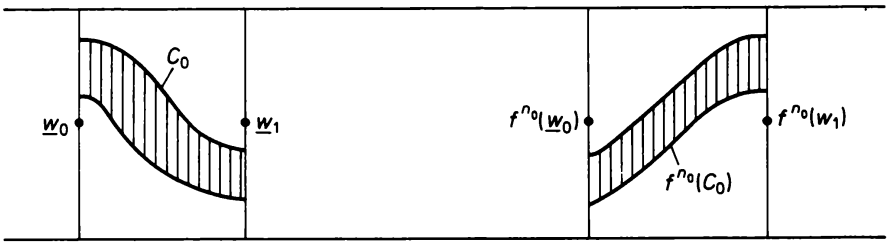

Fig. 4

Now, suppose we can find for $t > 0$ subsets $C_0 \supseteq C_1 \supseteq C_2 \ldots C_t$ so that the following conditions are satisfied:

($\#_t$) (1) C_t is a negative diagonal of $B_{\underline{w}_0\underline{w}_1}$,

(2) each $\zeta \in C_t$ satisfies (*s) for $s = 0, 1, \ldots, t$,

(3) with $J = \sum_{i=0}^{t} n_i + \sum_{i=0}^{t-1} k_i + 2(q_t + q_{t+1})$, $f^J(C_t)$ is a positive diagonal of $B_{w_0w_1}$ for w_0, w_1 adjacent points of $eo(z_t)$,

(4) the upper edge of $f^J(C_t)$ is in $f^J(I^+_{\underline{w}_0})$, the lower edge is in $f^J(I^-_{\underline{w}_1})$ and

$$\partial f^J(C_t) \cap B_{w_0w_1} \subseteq f^J(I^+_{\underline{w}_0} \cup I^-_{\underline{w}_1});$$

i.e., the points in C_t satisfy Theorem B for n_i, p_i/q_i, z_i for $i = 0, 1, \ldots, t$ plus the added geometrical conditions 1, 3 and 4 which prepare for the induction step (see Fig. 5). Note that we added $2(q_t + q_{t+1})$ in condition (3) in order to guarantee that no points of $eo(z_{t+1})$ are contained in $f^J(C_t)$ (the points of $eo(z_t)$ and $eo(z_{t+1})$ have different rotation numbers, hence move apart under iteration). We will therefore add $2(q_t + q_{t+1})$ to k_t below. Fix J as in 3 above. We must find a subset of C_t which is also a negative diagonal of $B_{\underline{w}_0\underline{w}_1}$ which satisfies ($\#_{t+1}$), i.e. which satisfies conditions (1)–(4) above with t replaced by $t+1$. The idea is to use orbits passing from near $y = 0$ to near $y = 1$ and from near $y = 1$ to near $y = 0$ to show that $f^J(C_t)$ will stretch out under application of f–in fact we can make the upper right hand part of the

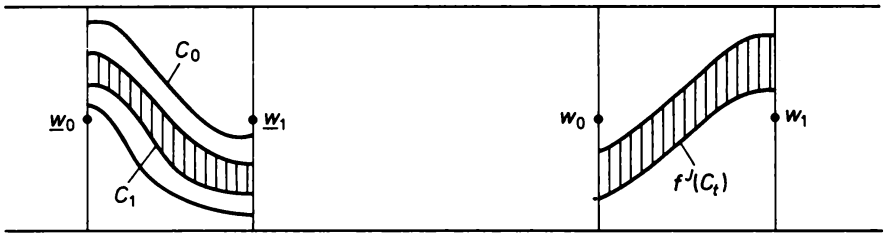

Fig. 5

image as close to $y = 1$ as we like and the lower left as close to $y = 0$ as we like. Hence the image of $f^J(C_t)$ will eventually contain a section which stretches all the way across a fundamental interval of A, and hence will have a piece which is a positive diagonal of the strip between two adjacent points of $eo(z_{t+1})$.

Fix $\varepsilon > 0$ so that

$$\varepsilon < \min\{\pi_2(z):\ z \in eo(z_t) \cup eo(z_{z+1})\}/2$$

$$\varepsilon < (1 - \max\{\pi_2(z):\ z \in eo(z_t) \cup eo(z_{t+1})\})/2.$$

Also we fix $m_1 > 0$ so that

$$\pi_1(f^{m_1}(z)) - \pi_1(f^{m_1}(w)) > 1 + \pi_1(z) - \pi_1(w),$$

whenever $\pi_2(z) > 1-\varepsilon$ and $\pi_2(w) < \varepsilon$ (i.e., the points near $y = 1$ move to the right strictly faster than those near $y = 0$). Now, fix points $p_0, p_1 \in A$ satisfying the following:

(a) $\pi_2(p_0) > 1-\varepsilon$, $\pi_2(p_1) < \varepsilon$,

(b) $\pi_2(f^{k+i}(p_0)) < \varepsilon$, $\pi_2(f^{k+i}(p_1)) > 1-\varepsilon$ for some $k > 0$ and $i = 1, \ldots, m_1$

(c) letting $K = k + m_1$ with k as in (b), for $0 \leqslant i \leqslant K$

$$\pi_1(f^i(p_0)) < \pi_1(f^i(w_1)),$$

$$\pi_1(f^i(p_1)) > \pi_1(f^i(w_0)), \quad \pi_1(f^i(p_0)) < \pi_1(f^i(p_1)),$$

and

$$\pi_1(p_0) < \pi_1(w_0), \quad \pi_1(f^k(p_0)) < \pi_1(f^k(w_0)),$$

$$\pi_1(p_1) > \pi_1(w_1), \quad \pi_1(f^k(p_1)) > \pi_1(f^k(w_1)),$$

(d) for some k_0, k_1, $0 < k_0 < k_1 < k$,

$$\pi_1(f^{k_0}(p_0)) > \pi_1(f^{k_0}(w_0)),$$

$$\pi_1(f^{k_1}(p_1)) < \pi_1(f^{k_1}(w_1)).$$

This means that p_0 goes around w_0 and p_1 goes around w_1. The existence of points $\underline{p}_0, \underline{p}_1 \in A$ satisfying (a) and (b) above follows from Condition B. Once an orbit gets close to a boundary curve it will stay close for some iterates because the boundary curves are invariant. So for any predetermined ε and m_1 we can find points satisfying (a) and (b). The constant k will depend on m_1, ε and the map. To guarantee conditions (c) and (d) we select p_0, p_1 on the extended orbits of $\underline{p}_0$, $\underline{p}_1$ respectively, e.g.: Suppose $\underline{p}_0$ satisfies (a–d) with some $\eta \in eo(z_t)$ in place of w_0. Then $w_0 = f^i(\eta)+(j, 0)$ for some i between 0 and integer j, so we let $p_0 = f^i(\underline{p}_0)+(j, 0)$.

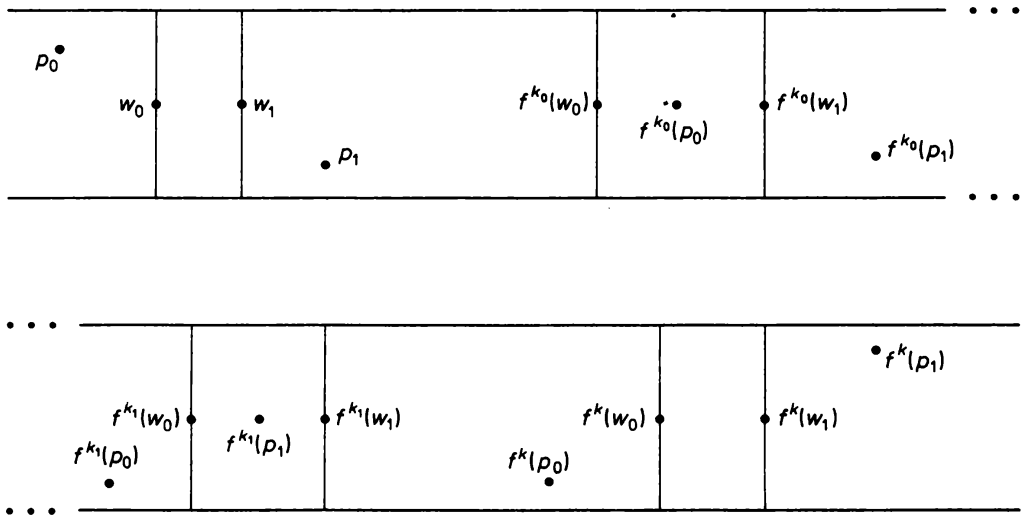

Fig. 6

Now we follow images of $f^J(C_t)$ under f and note that by repeated application of the lemma we have that $f^{J+k}(C_t)$ has a section which is a positive diagonal of $B_{f^{k(p_0)}f^{k(p_1)}}$ and contains no points of $eo(z_{t+1})$ and hence $f^{J+K}(C_t)$ has a section which is a positive diagonal of $B_{f^{K(p_0)}f^{K(p_1)}}$. But $|\pi_1(f^K(p_0))-\pi_1(f^K(p_1))| > 1$ by the choice of K so there must exist adjacent points $\tilde{w}_0, \tilde{w}_1 \in eo(z_{t+1})$ such that $f^{J+K}(C_t)$ contains a section which is a positive diagonal of $B_{\tilde{w}_0\tilde{w}_1}$. We call this set $\underline{C}_t$. By applying the lemma n_{t+1} more times we see that $f^m(\underline{C}_t) \cap B_{f^{m(\tilde{w}_0)}f^{m(\tilde{w}_1)}}$ contains a component which is a positive diagonal for $m = 1, \ldots, n_{t+1}$. Hence we may let $k_t = K+2(q_t+q_{t+1})$ and we have a set $C_{t+1} \subseteq C_t$ which satisfies $(\#_{t+1})$, (see Fig. 7).

By induction this completes the proof for $t > 0$. By an analogous argument we can obtain negative diagonals of $B_{\underline{w}_0\underline{w}_1}$, say C_{-t}, which satisfy $(**_{-t})$ of the theorem. Since positive and negative diagonals must intersect, the intersection of the C_t and C_{-t} yield the required point ζ. ∎

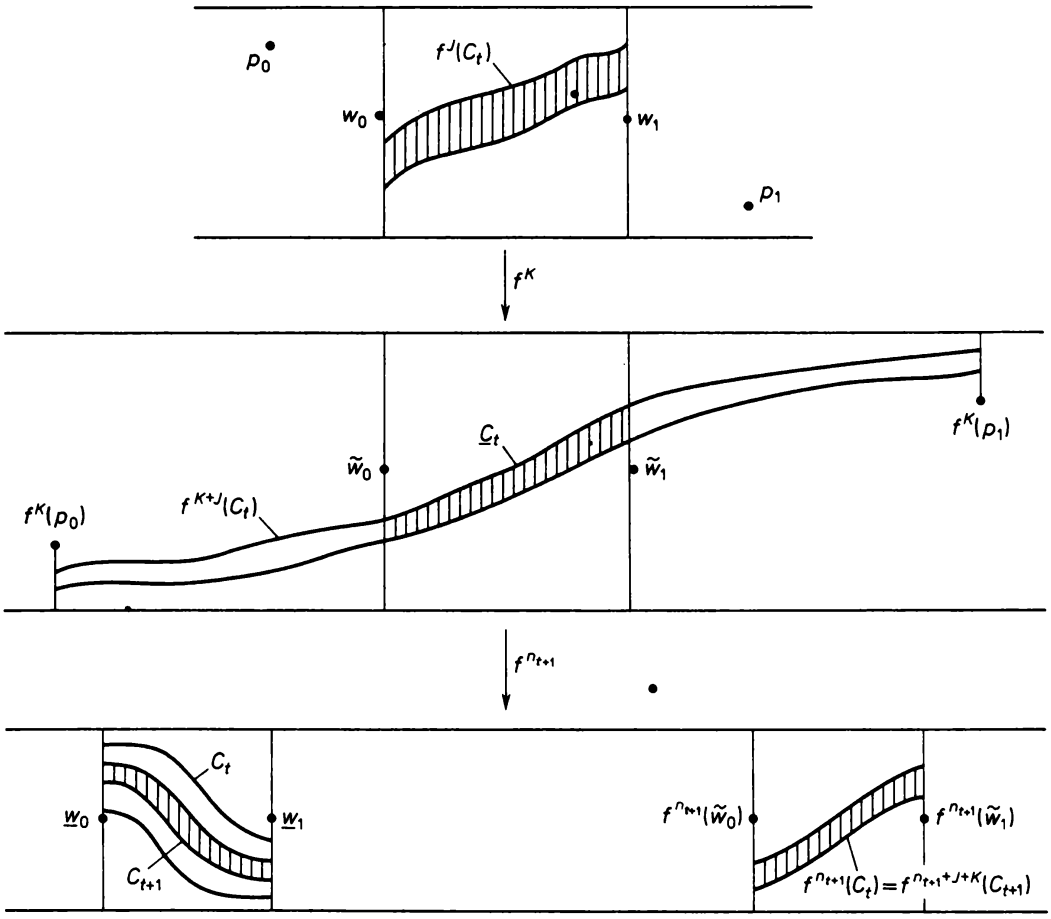

Fig. 7

Remarks. (1) Note that the size of the k_t depends only on the q_t, q_{t+1} and the number of iterates it takes for orbits to "transit" across A.

(2) Given periodic sequences we can obtain periodic orbits in the above by using the transit orbits to pick which adjacent points of $eo(z_t)$ we use, see Boyland–Hall [BH].

(5) *Concluding remarks.* The point of this note has been to show that the monotone twist condition and some assumptions on the existence of orbits connecting boundary components forces the image of vertical strips to behave in complicated but predictable ways under iteration. Hence by elementary topological arguments we can prove the existence of orbits with many different qualitative behaviors. Boyland [Bd1] has recently studied in more detail the relation between various non-monotone orbits.

The author wishes to thank the member of the School of Mathematics of the University of Minnesota for their hospitality and F. Przytycki and the referee for comments and corrections.

References

[Bf] G. D. Birkhoff, *Nouvelles recherches sur les systemes dynamiques*, Memoriae Pont. Acad. Sci. Novi Lyncaei 1 (1935), 85–216; and Collected Math. Papers, II, p. 530–661.

[Bd] P. L. Boyland, *Rotation sets and Morse decompositions in twist maps*, preprint.

[Bd1] – *Non-monotone and other periodic orbits in twist maps*, preprint.

[BH] P. L. Boyland and G. R. Hall, *Invariant circles and the order structure of periodic orbits in monotone twist maps*, to appear in *Topology*.

[Hn] M. R. Herman, *Sur les courbes invariantes par les diffeomorphismes de l'anneau*, Astérisque (1983), 103–104.

[K] A. Katok, *Some remarks on Birkhoff and Mather twist theorems*, Ergod. Th. and Dynam. Sys. 2 (1982), 195–194.

[LC] P. Le Calvez, *Proprietes dynamicques des regions d'instabilite*, preprint.

[Mal] J. N. Mather, *Talk at Oberwolfach conference*, 1985.

[Ma] – *A criterion for the non-existence of invariant circles*, Publ. Math. I.H.E.S. 63 (1986), 153–204.