On Aubry Mather Sets

J. J. P. Veerman and F. M. Tangerman

April 1990

SUNY StonyBrook
Institute for Mathematical Sciences
Preprint #1990/3
ON AUBRY MATHER SETS

J.J.P. Veerman
Institute for the Mathematical Sciences,
SUNY at Stony Brook, Stony Brook, NY 11794.

F.M. Tangerman
Mathematics department,
SUNY at Stony Brook, Stony Brook, NY 11794.

ABSTRACT

Let $f$ be a two dimensional area preserving twist map. For each irrational rotation number in a certain (non trivial) interval, there is an $f$-invariant minimal set which preserves order with respect to that rotation number. For large nonlinearity these sets are, typically, Cantor sets and they are referred to as Aubry Mather sets.

We prove that under some assumptions these sets are ordered vertically according to ascending rotation number ("Monotonicity"). Furthermore, if $f$ satisfies certain conditions, the right hand points of the gaps in an irrational Cantor set lie on a single orbit ("Single Gap") and diffusion through these Aubry Mather sets can be understood as a limit of resonance overlaps ("Convergence of Turnstiles"). These conditions essentially establish the existence of a hyperbolic structure and limit the number of homoclinic minimizing orbits.

Some other results along similar lines are given, such as the continuity at irrational rotation numbers of the Lyapunov exponent on Aubry Mather sets.
I INTRODUCTION

Let f be an area preserving monotone twist map on the cylinder $A = S^1 \times \mathbb{R}$. For each number $\alpha$ in the rotation interval $I$ of $f$, Aubry [1986] and Mather [1986] have constructed $f$-invariant sets $M_\alpha$, that have the given number as rotation number. These invariant sets are constructed as the global minima of a certain action functional. The topologically minimal sets $E_\alpha$ with irrational rotation number are precisely the Aubry Mather sets.

These sets are well-defined [Mather, 1986], lie on Lipschitz graphs over $S^1$ and on them the dynamics preserves the circular ordering [Katok, 1982a]. They can be smooth invariant, homotopically non-trivial curves, so-called KAM curves. For large enough non-linearity, though, one expects them to be broken up into Cantor sets [Goroff, 1985]. In fact, the parameter value at which a set breaks up depends to a large extent on the number-theoretical properties of the rotation number in question [MacKay, 1986, Mather, 1987]. These Cantor sets are then the 'remnants' of the invariant KAM curves of the nearly integrable case.

Aubry Mather sets play an important role in the global dynamics of the map, especially in stability questions. As invariant curves, they confine the dynamics of all orbits to narrow regions. However, numerical experiments indicate, that even as Cantor sets, these sets continue to restrict vertical motion [MacKay, Meiss, and Percival, 1984]. Their attempts to understand this, led to a geometrical construction they called "turnstile". The idea
was to construct the stable and unstable manifolds in the gaps of the Cantor sets, thus capturing the area per iterate that diffuses across the set.

In this article, we prove a number of fundamental theorems. One of these (the Monotonicity theorem) states that the $E_p$ admit a vertical ordering in the cylinder. Another has been conjectured before on the basis of numerical evidence: this is the theorem that asserts that the diffusion through an Aubry Mather set can be considered as a limit of resonance overlaps. Finally, the Single Gap theorem which says that Aubry Mather sets generically have only one gap orbit in them, has not appeared in the literature, as far as we are aware.

It is often convenient to consider the rotation number as being an element of the extended rotation interval $I^+$ defined as follows. Replace each rational number $p/q$ with the set $(p/q^-, p/q, p/q^+)$ with the natural ordering between them. With this ordering, the ordering on $I$ induces an ordering on $I^+$. The topology on $I^+$ is the order topology. Notice, that $p/q$ is an isolated point. We will often use $I$ and $I^+$ interchangeably.

If $\rho$ is rational, say $p/q$, then $E_{p/q}$ will denote a minimizing $q$-periodic orbit with rotation number $p/q$. Katok [1982b] proved the existence of minimizing orbits that are homoclinic to $E_{p/q^+}$ one advancing, $E_{p/q^+}$, and one receding, $E_{p/q^-}$. In this work, the only $C^k$ generic, $k \geq 1$, properties of $f$ that we use, are the following. First of all, $E_{p/q}$ consists of a single hyperbolic periodic orbit. Second, $M_{p/q}$ consists of $E_{p/q}$ plus a single
advancing orbit, \( E_{p/q+} \), homoclinic to \( E_{p/q} \), and a single receding orbit, 
\( E_{p/q-} \), also homoclinic to \( E_{p/q} \) (in the sense that \( q \)-th iterates of points 
move between successive points in \( E_{p/q} \)). The proof of this is standard and 
an outline is given in the appendix. (In a forthcoming work [Veer and 
Tangerman, 1989] we prove uniqueness of these orbits in the case of the 
standard map with large enough non-linearity parameter.) 

Hausdorff limits (Hlim) of these sets are well-defined [Mather, 1986]:

\[
\text{Hlim}_{\alpha \to p/q+} E_{\alpha} = E_{p/q} U E_{p/q+} = \text{clos}(E_{p/q+})
\]

(1.1a)

\[
\text{Hlim}_{\alpha \to p/q-} E_{\alpha} = E_{p/q} U E_{p/q-} = \text{clos}(E_{p/q-})
\]

For \( \omega \) irrational:

\[
E_{\omega} \subseteq \text{Hlim}_{\alpha \to \omega} E_{\alpha}
\]

(1.1b)

In order to avoid notational complications, results will be stated and 
proved, where possible, in the universal covering space of the cylinder 
without further comment. For the lift of \( f \), the notation \( \tilde{F} \) will be used.
II MONOTONICITY

In this section, the monotonicity result will be proven. This result restricts the region that Aubry Mather sets with irrational rotation number can inhabit.

Let $f$ be a $C^k$ ($k \geq 1$) area preserving monotone twist map on the cylinder $A = S^1 \times \mathbb{R}$. Fix a lift $F : \mathbb{R}^2 \to \mathbb{R}^2$ of $f$. When necessary, we will use coordinates $(x, y)$ on $\mathbb{R}^2$. Note that $F$ commutes with the unit translation in the $x$ direction. Each point $p = (x_0, y_0)$ defines an orbit $(F^i(p))$ and the projections $\pi(x_i, y_i)$ on the $x$-coordinate are called $x_i$. Denote by $M_\alpha$ the set of minimizing points of rotation number $\alpha$, i.e.: $\lim x_i/i = \alpha$.

We define the local stable and unstable manifolds $W^s/u_\epsilon(x)$ and their inverse resp. forward images as the stable and unstable manifolds in the usual way (see for example Lanford, [1985]).

Let $V$ denote the foliation of $A = \mathbb{R}^2$ by vertical lines, so that $F(V)$ and $F^{-1}(V)$ are the corresponding images of $V$ under $F$. At a point $p$ in $A$, we can now define the open cone $C_p = C^-_p \cup C^+_p$ bounded by $F(V)_p$ and $F^{-1}(V)_p$ and containing $V_p$ (see figure 2.1). Here $V_p$ denotes the leaf of the foliation $V$ through $p$ and $-$ or $+$ indicates the downward respectively, the upward component. Define:

$$C^{+/-}(p) = \lim_{i \to \pm \infty} C^{+/-}_{F^i(p)}.$$
Similarly, define tangent cones $T_{p} = T_{p}^{+} U T_{p}^{-}$ as the cone in the tangent space to $p$ whose boundary is formed by the tangent lines to $C_{p}$.

The fundamental wisdom that underlies this section, is that two minimizing points, $s$ and $p$, with different rotation numbers in $I^{+}$, satisfy a geometrical inequality. Aubry's fundamental lemma [Aubry, 1986] implies, that $s$ cannot lie in $C_{p}^{-}$ if its rotation number in $I$ is greater than or equal to that of $p$. But then, of course, $s$ cannot lie in iterates of $C_{p}^{-}$ either, and the same is true for iterates of $s$.

**Lemma 2.1:** Let $p$ be a point of $E_{p/q^{+}}$ or $E_{p/q^{-}}$, the tangent to the stable and unstable manifolds $W_{s}^{s}(p)$ and $W_{u}^{u}(p)$ at $p$ is given by $\lim_{n \to \infty} \frac{t_{n}^{-}}{|t_{n}^{s}|}$, resp., $\lim_{n \to \infty} \frac{t_{n}^{u}}{|t_{n}^{u}|}$, where $t_{n}^{s} = D F^{nq_{0}}(F^{nq_{0}}(p)) \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and $t_{n}^{u} = D F^{nq_{0}}(F^{-nq_{0}}(p)) \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

**Proof:** This follows directly from the definition of the local stable and unstable manifolds plus the fact that for $n$ large enough the vector $(0, 1)$ does not lie on $W_{\epsilon}^{u}(F^{nq_{0}}(p))$ nor on $W_{\epsilon}^{s}(F^{-nq_{0}}(p))$. The latter claim is proved as follows.

The points $F^{nq_{0}}(p)$ and $F^{-nq_{0}}(p)$ are close to points $x$ and $y$ in $E_{p/q^{-}}$. By general hyperbolic theory (see also section 3),
Invariant manifolds are nearly parallel to those of x and y. But the tangents at x and y cannot be contained in $T_{C_{x}}$, because if they were, Aubry's fundamental lemma would be violated since $E_{p/q}$ and $E_{p/q+}$ accumulate onto x and y along those tangents.

The following lemma shows that the invariant manifolds emanating from a point $p \in E_{p/q}$, $E_{p/q-}$, or $E_{p/q+}$ can never be very close to vertical, and that the upper right branch is an unstable manifold. For clarity, we number the branches clockwise, starting from the vertical (see figure 2.2).

**Lemma 2.2:** If $p \in E_{p/q}$, $E_{p/q-}$, or $E_{p/q+}$, then the first clockwise branch, $W_{1}$, is an unstable one. Moreover, corresponding branches on the orbit of p map into each other.

**Proof:** The region

$$T_{C_{1}} = DF^{-1}(p)(C^{-1}) F^{-1}(p) U \quad T_{C_{p}} F^{-1}(F(p)) F(p)$$

forms a new local cone (each original cone has $V_{p}$ in its closure). It is then easy to see, that $T_{C_{n}}$, defined in the obvious way, contains $T_{C_{n-1}}$, for all $n$. By lemma 2.1, the boundary of $T_{C_{nq}}$ (q fixed) must accumulate on a stable and an unstable direction. Observe that $T_{C_{nq}}$ cannot contain any
stable or unstable directions and further that interior of $T C_{nq}$ is connected. Therefore, its boundaries can only accumulate onto $W_2$ and $W_3$.

The second statement follows from the fact that $C^-(p)$ is mapped to $C^-(f(p))$. Thus the tangents to the boundaries at $p$ and $f(p)$ are also mapped to one another. □

For each $p/q^+$ or $p/q^-$, separating curves $\gamma(p/q^+)$ or $\gamma(p/q^-)$ on the cylinder can be defined as follows. For each pair of neighboring points $p_1$ and $p_2$ in $E_{p/q}$, pick one point $s$ in $E_{p/q^+}$ (or $E_{p/q^-}$) between them. Connect $s$ to the points $p_1$ and $p_2$ along their invariant manifolds. We define $\gamma(p/q^+)$ (or $\gamma(p/q^-)$) as the closed curve obtained as the union of the segments (see figure 2.3). An orientation along the curve can be defined in such a way that the orientation has a component to the right on the points of $E_{p/q}$ for the construction of $\gamma(p/q^+)$ and to the left for the construction of $\gamma(p/q^-)$. With this orientation, there is now a slight extension of lemma 2.2:

**Corollary 2.3:** At a hyperbolic minimizing point in $E_{p/q^+}$ the first clockwise branch is unstable and oriented with a component to the right, the second is stable and oriented to the right, and so on. The orientation is opposite for $E_{p/q^-}$. 

The curves $\gamma$ are not necessarily Jordan. Clearly, $\gamma(p/q+)$ separates $A$ in an upper component containing $+\infty$, a lower component containing $-\infty$ and finitely many other components. Note, that these curves and components depend on the choice of $q$ points $s_1, \ldots, s_q$ in $E_{p/q^+}$, respectively, $E_{p/q^-}$. In the following we will make a simplifying assumption on the character of these curves (which will be proved to hold in the case of the standard map with large enough $k$ in a forthcoming work [Veeramani and Tangerman, 1989]).

Let $G$ be a pair of two neighboring points $a$ and $b$ of $E_{p/q^+}$ such that in the plane $\pi(a) < \pi(b)$. The segment $\gamma(p/q^+)$ that connects them is called $W^u(G)$ if it is an unstable segment, or $W^s(G)$ if is a stable one. Let $l_a$ and $l_b$ the left boundary $C^-_a$ resp. $C^-_b$ and $r_a$ and $r_b$ the right sides.

**Condition 2.4:** For each neighboring pair $G$ of points $a$ and $b$ in $E_{p/q^+}$ (a to left of b):

$W^u(G) \cap (l_a U l_b) = \phi$ if the connecting segment is unstable,

$W^s(G) \cap (r_a U r_b) = \phi$ if the connecting segment is stable.

Similarly for $\gamma(p/q^-)$.

We remark that some types of intersection are a priori excluded. Suppose for example that we are interested in an unstable connecting segment of $\gamma(p/q^+)$ which we parametrize, starting at the point $a$, by $\gamma(t)$. Let
$\psi(\gamma(t))$ be the angle of the tangent to $\gamma(t)$ with the positive vertical.

Counting clockwise as positive, define $\phi(\gamma(t))$ as

$$\phi(\gamma(t)) = \int_0^t \frac{d}{dt} \psi(\gamma(t)) \, dt.$$ 

If $i_a$ is an intersection in $W^u(G)$ and $i_b$ in $W^u(G)$, the remark is that $\phi(i_a) > 0$ and $\phi(i_b) > 0$.

One proves this by showing that if for example $\phi(i_a) < 0$, then the inverse image under $f$ of $W^u(G)$ has an intersection point with the same property. But we know that inverse images of $W^u(G)$ eventually land in the local unstable manifolds to $E_{p/q}$ which do not have this property.

Suppose we choose points $s_1, \ldots, s_q$ in the construction of $\gamma(p/q^+)$ such that condition 2.4 holds. The figure consisting of $W^u(G)$, $l_a$, and $l_b$ (see figure 2.4) then separates the plane in two components (similarly for $W^s(G)$, $r_a$, and $r_b$). Only one of these components contains $+\infty$. The other one is called 'below ab'. We now define 'below $\gamma(p/q^+)$' as follows: a point $x$ is 'below $\gamma(p/q^+)$' if

$$x \in C_a \text{ for some } a \in E_{p/q^+}$$

or

$$x \in \text{ 'below ab' for some neighboring pair } a \text{ and } b \text{ in } E_{p/q^+}.$$ 

'Above $\gamma(p/q^+)$' is the complement of $(\gamma \cup \text{ 'below $\gamma(p/q^+)$'})$. One gives a similar definition for 'above $\gamma(p/q^-)$' and 'below $\gamma(p/q^-)$'. We will use the symbols '<' and '>' for 'below' and 'above', respectively.
Remark: Note, that the definition of above and below $\gamma(p/q \pm)$ is not symmetric. The same holds for $\gamma(p/q \mp)$. Note further, that in the case that $\gamma$ is a Jordan curve, this notion coincides with the standard interpretation interpretation of 'above' and 'below'.

**Theorem 2.5:** For any $\gamma(p/q\pm)$, $\gamma(p/q\mp)$ for which condition 2.4 holds, we have: (i) If $\alpha > p/q$, then $M_\alpha > \gamma(p/q\pm)$ and $M_\alpha > \gamma(p/q\mp)$ and (ii) if $\alpha < p/q$, then $M_\alpha < \gamma(p/q\mp)$ and $M_\alpha < \gamma(p/q\pm)$.

**Proof:** We only prove the first half of the first statement. The statement will follow from a contradiction by supposing that there is a point $x \in M_\alpha$ such that $x$ lies on the 'wrong' side of $\gamma(p/q\pm)$.

Suppose without loss of generality that the segment that connects the neighboring points $a_0$ and $b_0$ of a 'gap' $C_\alpha$, is unstable. Let $x \in M_\alpha$ with $\alpha > p/q$. As noted before, $x$ cannot lie in $C_{a_0}^-$ or $C_{b_0}^-$. It remains to be proved that $x$ is not contained in the region $S_0$ (possibly consisting of more than one component) bounded by $W^u(G_0)$, $r_{a_0}$, and $l_{b_0}$.

Iterate by $f^{-1}$. Then $S_0$ is mapped into the region bounded by $W^u(G_1)$, $r_{a_1}$, and $v_{b_1}$. By Aubry's fundamental lemma, the point $x \in S_0$ cannot be mapped to a cone. Therefore it must land in $S_1$ which is the region bounded
by $W^u(G_1)$, $r_{a_1}$, and $l_{b_1}$. We can continue this, inductively defining $S_n$ containing $f^{-n}(x)$, until, for some $n$, $S_n$ lies in an $\epsilon$-neighborhood of a hyperbolic periodic point.

But this neighborhood can be chosen so small that $f^q$ restricted to it is very nearly linear. By lemma 2.2, we know the orientations of the local invariant manifolds (see figure 2.5). Orbits of points in $S_n$ under $f^q$ lie on hyperbolae. Any order preserving orbit in $S_n$ with rotation number greater than $p/q$ must satisfy

$$\pi(f^q(y)) > \pi(y).$$

These requirements are incompatible and thus $x$ cannot map to $S_n$. □

Notice, that we can compare two irrational sets as well, since there are always pairs of $\gamma(p/q^+)$ and $\gamma(p/q^-)$ that separate them. So, theorem 2.6 follows immediately.

**Theorem 2.6** "Monotonicity": If for all $p/q^+$ and $p/q^-$ in $I$, $\gamma(p/q^+)$ and $\gamma(p/q^-)$ can be constructed that satisfy condition 2.4, then $\alpha > \beta$ implies $E_\alpha$ lies above $E_\beta$. 
III HYPERBOLICITY

Here, we prove that the invariant minimizing sets close enough to rationals are hyperbolic (and thus for irrational rotation numbers Cantor sets). One expects these sets to be hyperbolic as soon they break up (see Li and Bak [1986]). The rest of the section is devoted to a corollary stating that such sets have Hausdorff dimension zero.

We start with some generalities concerning the hyperbolic sets that we are interested in. Again, we assume $f$ to be generic, so that equations 1.1 hold. The set $H_N = \bigcup_{\rho \in N} E_{\rho}$ is compact if $N$ is a closed interval in $\mathbb{R}^+$. According to Lanford [1985], a compact invariant set $H$ is a (uniformly) hyperbolic set, if the tangent space of each point $x$ of the set is spanned by stable and unstable spaces and if the following holds. The tangent vectors in the stable space must be contracted exponentially (as $\mu^n$, $\mu < 1$ uniformly on the set) under $Df^N(x)$, and the same holds for vectors in the unstable spaces under $Df^{-N}(x)$. These requirements imply that the local stable and unstable manifolds $W^S_\varepsilon(x)$ and $W^U_\varepsilon(x)$, tangent to the stable and unstable spaces, are continuous as functions of $x \in H$, and that their diameter is uniformly bounded away from zero [Lanford, 1985]. That, in turn, implies, that there is a $\delta_0 > 0$, so that for any pair $x$ and $y$ in $H$ whose distance is less than $\delta_0$, $W^S_\varepsilon(x)$ and $W^U_\varepsilon(y)$ have a unique intersection point [Lanford, 1985].
In the following, we will establish the genericity of hyperbolicity. To do that, we use a cone field criterion as also described in Le Calvez [1987].

**Theorem 3.1** "Hyperbolicity": Let \( h \) be a hyperbolic set for \( f \), then there exists a compact neighborhood \( H \) of \( h \) so that \( \cap_{i=0}^{\infty} f^i(H) \) is also a hyperbolic set for \( f \).

**Proof**: Since \( h \) is compact and hyperbolic, one can construct a cone field \( \{ C_x \}_{x \in h} \) which is mapped strictly into itself by \( Df \). One does this by constructing a norm on the tangent bundle restricted to \( h \) such that \( Df \) is expanding on the unstable bundle (choose unit vector \( e_u(x) \)) and contracting on the stable bundle (choose unit vector \( e_s(x) \)), see lemma 2.1 of Nitecki [1971]. Then choose the cone field \( C_x \) as follows: a vector

\[ v = ae_u(x) + be_s(x) \text{ is in } C_x \text{ if } |a| \geq |b|. \]

By continuity of \( Df \), we can extend this cone field \( C \) to a cone field \( \{ C'_x \}_{x \in H} \) defined on a sufficiently small neighborhood \( H \) of \( h \) such that \( Df \) maps the cone field \( C' \) on \( f^{-1}(H) \cap H \) strictly into \( C' \) on \( H \). Consequently, any invariant compact set in \( \cap_{i=0}^{\infty} f^i(H) \) is also a hyperbolic set. \( \square \)
Corollary 3.2 (see Le Calvez [1987]): For generic $f$, there exists an open neighborhood $U$ of the rational rotation numbers such that the collection of minimizing sets with rotation number in $U$ forms a hyperbolic set.

Proof: Take $H = E_{p/q}^{-} U E_{p/q}^{+} U E_{p/q}$. and pick $H$ as above. For generic $f$, the set $H$ is hyperbolic. \qed

We let $\lambda(\rho)$ denote the Lyapunov coefficient $\geq 1$ for an order preserving minimal set $E_\rho$. Since these sets are uniquely ergodic with invariant probability measure $\mu(\rho)$ (see Mather [1986]), $\lambda(\rho)$ is well-defined and constant $\mu$ almost everywhere.

Proposition 3.3: Let $M_\alpha$ be a hyperbolic minimizing set with irrational rotation number, then $\lambda(\rho)$ is continuous at $\rho = \alpha$.

Proof: Let $h$ be a hyperbolic set for $f$ with one-dimensional unstable bundle $E^u$. Assume $E^u$ is orientable. Choose a continuous nowhere zero section $v$ of $E^u$ and consider the function

$$D(x) = \frac{|Df \cdot v(x)|}{|v(f(x))|}.$$ 

One observes that $D$ is continuous on $h$. For an ergodic probability measure $\mu$ on $h$, its Lyapunov coefficient $\lambda(\mu)$ equals

$$\lambda(\mu) = \exp \int \ln D \, d\mu.$$
Since $M_\alpha$ is hyperbolic, we have by theorem 3.1 that $\overline{\lim}_{i \to -\infty} f^i(H)$ is also hyperbolic where $H$ is a sufficiently small neighborhood of $M_\rho$. One knows that (Mather [1986]), for $\alpha$ irrational, $\overline{\lim}_{i \to -\infty} f^i(H)$ contains nearby minimizing sets $M_\rho$ with invariant probability measures $\mu(\rho)$ and that $\lim_{\rho \to \alpha} \mu(\rho) = \mu(\alpha)$, in the weak topology. Consequently $\lim_{\rho \to \alpha} \lambda(\rho) = \lambda(\alpha)$. 

**Remark:** The function $D(x)$ is not canonical. If one chooses a different section $v'(x) = \phi(x) v(x)$, then

$$D'(x) = \frac{\phi(x) D(x)}{\phi(\xi(x))}.$$  

However, the Lyapunov coefficient is insensitive to this: one easily checks that

$$\int \ln D' \ d\mu = \int \ln D \ d\mu.$$  

Finally, a simple result that follows from hyperbolicity. Let $x$ be a point in $E_\alpha$ where both forward and backward images of the gap $G$ accumulate. In a neighborhood of $x$ one can connect these tiny gaps by local stable and unstable manifolds which are almost straight segments that make a positive angle with each other. From the accumulation of these different gaps, one concludes the following.

**Remark:** The set $E_\alpha$ cannot be imbedded in a $C^1$ curve.
IV SINGLE GAP

Consider the projections of the Aubry Mather sets on the x-axis. By a 'gap' G in $E_{\rho}$, we mean [Katok, 1982] a pair of points in $E_{\rho}$, whose projections bound an interval that contains no point of the projection of $E_{\rho}$. The length $|G|$ of the gap is simply the length of that interval. The meaning of $F^i(G)$ is then also clear. The main result of this section is that, under certain assumptions, $E_{\alpha}$ has only one gap orbit.

Before we embark on the general discourse, we first formulate the Single Intersection hypothesis, which will be needed in theorem 4.3. Denote the finite pieces of invariant manifolds to $E_{\rho}$ that connect the endpoints of a gap J in $E_{\rho}$ by $W^s(J)$ and $W^u(J)$. We will say that $f$ satisfies the Single Intersection hypothesis if all $E_{p/q}$, $E_{p/q^-}$, and $E_{p/q^+}$ are unique (true for generic $f$, see section 1), and if, for a gap J in $E_{p/q^+}$ or $E_{p/q^-}$, $W^u(J) \cap W^s(J)$ contains single point (which then has to be the minimax), see figure 4.1.

**Lemma 4.1:** If $E_{\alpha}$ is hyperbolic, then it has at most finitely many gap orbits.

**Proof:** From the generalities mentioned in section 2, one can deduce, that for a hyperbolic set, there is a $\delta_0 > 0$ with the property, that if $x$ and $y$ are points in $E_{\alpha}$, then
if \[ d(f^i(x), f^i(y)) < \delta \] for all \( i \in \mathbb{Z} \),
then \[ x = y. \]

So, each gap orbit must have a gap of length greater than \( \delta \). \( \square \)

Remark: A different proof of this fact was given by MacKay [1987].

**Proposition 4.2:** For \( f \) generic, there is a neighborhood \( N \) of \( p/q \), such that if \( \alpha \in N \), then \( E_\alpha \) has one gap orbit.

**Proof:** If \( f \) generic, then \( E_{p/q^+} \) and \( E_{p/q^-} \) consist of a single homoclinic orbit (see section 1), and thus have one single gap orbit. According to proposition 1.3, for \( \alpha \) close enough to \( p/q \) (without loss of generality \( \alpha > p/q \)), \( E_\alpha \) is hyperbolic. Pick any gap in \( E_{p/q^+} \). There is an \( m > 0 \), such that all gaps in \( E_{p/q^+} \) with length greater the \( \delta/3 \) are contained in \( \{ F^i(G) \}_{i=-m}^{i=m} \).

By equation 1.1, we can pick \( \alpha \) so close to \( p/q \), that \( d(E_\alpha, E_{p/q^+}) < \delta/3. \)

Then, by uniform continuity of \( \{ f^i \}_{i=-m}^{i=m} \), every gap with length greater than \( \delta \) is shadowed by a gap in \( E_{p/q} \) and vice versa. \( \square \)

This result is somewhat unsatisfactory, since one would like to have a statement for \( \alpha \) fixed. In studying the global stability of these systems, numerical work indicates that the curve with rotation number \( (1+\sqrt{5})/2 \) (or a related Diophantine number, see MacKay [1986]) is the last one to break up.
By assumption c, we can choose an interval $K^+$ in $I^+$ of rotation numbers $\rho$ such that the set $E = \bigcup E^\rho$ (union over $K^+$) is uniformly hyperbolic. By assumption b, we can choose the interval $K^+$ so that in addition we have:

$$\text{Hdist}(E^\rho, E^\alpha) < \delta << \delta_0,$$

where $\delta_0$ is a lower bound for the diameter of the local invariant manifolds (see section 3) to $E^\rho$ with $\rho \in K^+$. Thus each $E^\rho_{p/q^+}, E^\rho_{p/q^-}$ in $E$ is contained in the local stable and unstable manifolds to $E^\rho_{p/q}$. Also each gap orbit in $E$ must have a gap which is larger than $\delta_0$. Denote the two 'big' independent gaps in $E^\alpha_{\alpha}$ by $G^\alpha_{\alpha}$ and $H^\alpha_{\alpha}$. According to equation 4.1, we can uniquely define by taking $G^\rho_{\rho}$ and $H^\rho_{\rho}$ to be approximating gaps in $E^\rho_{\rho}$. If $\rho$ is rational, rational+, or rational-, then, of course, we have that there exists an $m(\rho)$ with:

$$\varepsilon^{m(\rho)}(G^\rho_{\rho}) = H^\rho_{\rho}.$$  

By the continuity of $f$, it is clear that $m(\rho)$ has to become unbounded as $f \to \alpha$.

The question we address now, is, how does $m(\rho)$ change as a function of $\rho$ in $K^+$? From relation 1.1b and the continuity of $f$, it transpires that if $\omega \in K^+$ is irrational, then either $m(\omega) = \infty$ or $m(\rho)$ is continuous at $\omega$.

As a consequence, $m(\rho)$ can make finite jumps only at rational values of $\rho$. This situation is depicted in figure 4.2, where $m(p/q^-)$ is not equal to $m(p/q^+)$. Because $m(p/q)$ is unique, we have:
As a consequence, one is especially interested in the gap structure of this set, being, as it were, a 'bottleneck' for the dynamics of $f$.

We are now in a position to prove the main result.

**Theorem 4.3** "Single Gap": Let $\alpha$ be irrational. If

a) $f$ satisfies Single Intersection,

and

b) $\varlimsup_{\rho \to \alpha} E_\rho = E_\alpha$,

and

c) $E_\alpha$ is hyperbolic,

then $E_\alpha$ has only one gap orbit.

Remark: To prove the result for a single $\alpha$, it is enough to require that $f$ satisfy a local variant of the Single Intersection hypothesis.

Remark: Numerical work suggests that a) holds for the standard map. The more general case is commented upon after the proof. As stated before, c) has been proved only in a restricted setting [Goroff, 1985], but appears to hold more generally. One suspects that b) is generically true, see Bangert [1986].

**Proof**: We will assume from now on that there are two independent gaps in $E_\alpha$ and eventually deduce from that a contradiction with the Single Intersection hypothesis.
\[ m(p/q^+) = m(p/q^-) + kq. \quad (4.3) \]

Without loss of generality, we take \( m(p/q^-) > 0 \).

We will now argue that \( k \) is negative. Suppose, then, that \( k \) is positive. Relations 4.1 and 4.2 together with the well ordered character of \( E_p \) imply that \( f^m(p/q^-)(G_{p/q^+}) \) lies in a local unstable manifold. Upon iterating this \( m(p/q^-) \) times back to the original gap, as in figure 4.2a, one encounters a contradiction (namely, that points have left the local unstable manifold under the application of \( f^{-1} \)). So \( k \) is negative.

If \( m(p/q^+) > 0 \), then (recall that \( m(p/q^-) \) is positive) (4.3) implies that \( |m(p/q^+)| < |m(p/q^-)| \). So, in order to allow \( m(\rho) \) to become unbounded, we need:

\[ m(p/q^-) > 0 \quad \text{and} \quad m(p/q^+) < 0. \]

This is the situation sketched in figure 4.2b, and it is here that the final contradiction with Single Intersection arises. It can be seen as follows that in this case \( W^s(H_{p/q^-}) \) and \( W^u(H_{p/q^+}) \) intersect at least two times. Under forward iterates, the number of intersections involving the local stable manifolds along with the relative orientations (use corollary 2.3 to determine the orientations) is conserved. One concludes that the point marked \( p \) in the figure is mapped under \( f^m(p/q^-) \) (with \( m(p/q^-) > 0 \)) to the point labelled \( p' \), with the orientation as indicated. So

\[ p' \in W^u(f^m(p/q^-)(G_{p/q^-})) \cap W^s(f^m(p/q^-)(G_{p/q^+)}) = W^u(H_{p/q^-}) \cap W^s(f^{-kq}(H_{p/q^+})), \]
where $k$ is negative. Let $a'$ be the image under $f^m(p/q^-)$. Since $W^u(C_{p/q^+})$
does not intersect the $a$-$p$ (by Single Intersection) its image under $f^m(p/q^-)$
does not intersect $a'$-$p'$. Therefore, if $W^u(H_{p/q^-})$ intersects $p'$, it is
caught in an unstable 'lobe' of $f^m(p/q^-)(H_{p/q^+})$ and must intersect the local
stable manifold to $E_{p/q}$ again in order to leave the 'lobe'. By uniform
hyperbolicity $W^u(H_{p/q^-})$ then must intersect $f^m(p/q^-)(H_{p/q^+})$ another time.
The intersections with $W^s(H_{p/q^-})$ follow by the hyperbolicity of the two thus
constructed points (and hence the existence of their local unstable
manifolds).

\[\square\]

From the last paragraph of this proof, it is clear, that it is
sufficient to replace condition a in the theorem by the requirement
\[a^*) \quad W^s(H_{p^-}) \cap W^u(H_{p^-})\] is bounded away (uniformly in $K^+$) from the
endpoints of $H_{p^-}$.

This requirement appears to be borne out (Percival, private communication)
by extensive numerical experiments for the standard map. (If this were not
so, one would have ever longer and thinner lobes formed by $W^u(H_{p/q^-})$ as the
$H_{p/q^-}$ accumulate on the gap $H_{a^-}$.) We suspect that Single Gap is a persistent
property for an open neighborhood of maps around the standard map but can
only prove that for large values of the non-linearity parameter [Veeran and
Tangerman, 1989].
V TURNSTILES

Let $\alpha$ be an irrational rotation number in $I$ and assume that $E_\alpha$ is hyperbolic. This section is dedicated to proving that the leakage of orbits through an Aubry Mather set can be understood in terms of overlap criteria. It is somewhat speculative in nature, since we have to assume all the conditions that are required for theorem 4.3 to be true. Nevertheless, as stated, there is good reason to believe that the result holds for an open set of maps containing the standard map. This confidence is partly based on numerical results by various authors, especially MacKay, Meiss and Percival [1984]. We will thus proceed to elaborate on some of the consequences of the theorem.

We define turnstiles as follows [MacKay, Meiss, and Percival, 1984]. Let $G$ denote a gap in $E_\alpha$. Connect the endpoint of $f^n(G)$ with a straight line segment $\lambda_n^s$. Similarly, connect the endpoints of $f^{-n}(G)$ with a straight line segment $\lambda_n^u$. Clearly, $f^{+n}(\lambda_n^u)$ and $f^{-n}(\lambda_n^s)$ connects the endpoints of the gap $G$.

Note, that by hyperbolicity the endpoints of every gap $G$ are connected by a branch of stable manifold, and by a branch of the unstable manifold (the future and past iterates of $G$ collapse the gap). We will denote finite branches that connect a gap $J$ by $W^s(J)$ and $W^u(J)$.

**Proposition 5.1:** $\lim_{i \to \infty} f^{-1}(\lambda_i^s) = W^s(G)$
ii): \( \lim_{i \to \infty} f^i(u_i) = w^u(G) \)

**Proof:** It suffices to prove i only. Because the gaps \( f^i(G) \) do not overlap, the sum of their lengths is less or equal to one. So there is an \( N \) such that the length \( |f^{N+1}(G)| < \varepsilon \) for all \( i \geq 0 \). Then, by uniform hyperbolicity, for each \( i \) the left and the right endpoint of \( f^{N+1}(G) \) have to lie on the same local stable manifold. But then we also have \( \lim f^{-i}(\lambda_{N+1}^s) = w^s(f^N(G)) \), because \( \lambda_{N+1}^s \) is transversal to the local stable manifold of \( f^{-1} \).

As in section 3, we can define a hyperbolic set \( H_N \) that contains \( E_\rho \) with \( \rho \) in \( N \), a neighborhood of \( \alpha \). We now take \( (p_i/q_i) \) to be a sequence in \( N \) with \( \alpha \) as its limit. Define \( G_i^+/\cdot \) as the gaps in \( E_{p_i/q_i}^+ \) and \( E_{p_i/q_i}^- \) that have limit (as \( i \to \infty \)) \( G \).

The main result of this section is:

**Theorem 5.2** "Convergence of Turnstiles": Let \( f \) satisfy the same conditions as in theorem 4.3, then:

i): \( \lim_{i \to \infty} w^s(G_i^+) = w^s(G) \),

ii): \( \lim_{i \to \infty} w^s(G_i^-) = w^s(G) \),

iii): \( \lim_{i \to \infty} w^u(G_i^+) = w^u(G) \),
iv): \( \lim_{i \to \infty} W^u(G_i^-) = W^u(G) \),

**Proof:** It suffices to prove the first statement only. Since we are dealing with gaps \( G_i^+ \) and \( H_i^s \) only, we will drop the unnecessary superscripts \(^+\) and \(^s\) on them.

By the Single Gap theorem (4.3), we can choose an integer \( N \) such that
\[
\frac{j_{i+N}^+}{j_{i-N}^-} | f^j(G) | > 1 - \epsilon/2.
\]
By the equations 1.1 and the continuity of \( f \), it follows that, for \( i \) sufficiently big:
\[
\frac{j_{i+N}^+}{j_{i-N}^-} | f^j(G_i) | > 1 - \epsilon.
\]
So, if \( H_i = f^{+2N}(G_i) \), its length and that of its forward images is smaller than \( \epsilon \). Thus, by uniform hyperbolicity, its endpoints lie on the same local stable manifold. Let \( H = f^{+2N}(G) \). Then, by taking 2N (N fixed) inverse iterates, one concludes that i) is true, if and only if \( W^s(H_i) \) converges to \( W^s(H) \). But that follows directly from the fact that uniform hyperbolicity implies that local stable and unstable manifolds vary continuously as function of their base-point. So the theorem is proved. \( \Box \)

This theorem immediately implies, that the diffusion through an Aubry Mather set can be understood as a limit of "resonance overlaps". The way to see this, is to construct curves \( \gamma(p_i/q_i^+) \), with \( p_i/q_i \uparrow \alpha \), as in section 2,
except that now we take $s_i$ to be the left endpoint of the gaps $G_i$, defined as in the proof of the previous theorem. The other points needed in the construction are taken to be the $q_i-1$ images of $s_i$ (see figure 5.1). Iterate the area $B_i$ below $\gamma(p_i/q_i^+)$ once, and it is clear that one can define a region $I_i^+$ with:

$$f^{-1}(I_i^+) > \gamma(p_i/q_i^+) \quad \text{and} \quad I_i^+ < \gamma(p_i/q_i^+).$$

Similarly, a region $O_i^+$ can be defined by:

$$O_i^+ < \gamma(p_i/q_i^+) \quad \text{and} \quad f^{+1}(O_i^+) > \gamma(p_i/q_i^+).$$

Theorem 5.2 immediately implies that $I_i^+$ converges in the Hausdorff limit to $I$ and $O_i^+$ to 0, and $I$ and 0 are the areas enclosed by $W^s(G)$ and $W^u(G)$. The corresponding statement holds also for $\gamma(r_i^+)$, if we define $O_i^-$ and $I_i^-$ in a similar vein and if $s_i$ now accumulates to the right endpoint of the gap $G$.

By Convergence of Turnstiles, it follows, that the 'resonance overlaps'

$O_i^+ \cap O_i^-$ and $I_i^+ \cap I_i^-$ limit on 0 and I, respectively. We summarize this loosely with the following corollary:

**Corollary 5.3:** Under the conditions of theorem 4.3, the diffusion through an Aubry Mather set is a limit of resonance overlaps, if $E_\alpha$ is hyperbolic.
If one assumes that $E_\alpha$ is hyperbolic whenever it is a Cantor set, it is also clear that the above proves the following (geometric) criterion for the non-existence of an invariant circle with rotation number $\alpha$ (compare [Mather, 1986], whose result is more general but less geometric).

**Corollary 5.4:** If $E_\alpha$ is hyperbolic whenever it is a Cantor set, then, under the conditions of theorem 4.3, $\lim_{t \to \pm \infty} \cap O_1^+ \cap O_1^-$ converges and has positive area if and only if $E_\alpha$ is not (contained in) an invariant circle.
VI CONCLUDING REMARKS

We have argued, that one of our main results, the Single Gap theorem, holds for \( f \) satisfying a number of conditions (see theorem 4.3).

Nevertheless, it is not hard to find a counter-example. Suppose \( f \) is a twist map for which the Single Gap theorem holds and let [MacKay, private communication] \( g = f^2 \), then we have the equality \( E_\alpha(f) = E_{2\alpha}(g) \) (as sets).

If \( G \) and \( f(G) \) are gaps in \( E_\alpha(f) \), they can never be mapped into each other by \( g \). Moreover, if \( E_\alpha(f) \) is hyperbolic, then so is \( E_{2\alpha}(g) \), and under small perturbations of \( g' \) of \( g \), \( E_{2\alpha}(g') \) retains the same number of gaporbits.

(Note that \( g + \epsilon \), for \( \epsilon \) small enough, does not satisfy requirement \( a^* \) at the end of section 4.) One can conclude, therefore, that theorem 4.3 certainly will not hold generically.

This counter-example is also of interest in connection with the Monotonicity theorem. Note, that in our definition of 'above' and 'below', we have assumed, that \( E_{p/q} \), \( E_{p/q^-} \), and \( E_{p/q^+} \) are unique for all \( p/q \in \mathbb{I} \) (true for generic \( f \)). This does not hold for the map \( g \). As we let \( \epsilon \) run through zero, we can witness, that the minimizing sets jump. And so, our definition 2.4 also yields curves \( \gamma(p/q^+ \) and \( \gamma(p/q^-) \) that jump. A similar comment is valid for the Convergence of Turnstiles theorem.

The hyperbolicity of the Aubry Mather sets in a Birkhoff zone is a very powerful tool (see section 3). We expect it to hold in a general situation.
It would be useful to have a proof for a more general case than the one discussed in Goroff [1985].
APPENDIX

The lift $F$ of the twist map has a generating function $h$ satisfying:

$$F(x,y) = (x',y') \text{ iff } y = -\delta_1 h(x,x')$$

$$y' = \delta_2 h(x,x')$$

**Proposition A.1:** For $C^k$-generic ($k > 1$) $h$, the global minima $E_{p/q}$, $E_{p/q^-}$, $E_{p/q^+}$ are unique.

**Proof:** The Kupka Smale theorem for area preserving maps (see [Robinson, 1970]) implies in this context that generically there are a finite number of periodic orbits of given type. Moreover, if they are hyperbolic, two fundamental domains of a stable and an unstable invariant manifold associated with these orbits intersect finitely many times, and the intersections are transversal. This implies that there are finitely many orbits of type $E_{p/q^-}$ or $E_{p/q^+}$.

For $p/q$ fixed, let $(\xi_i)$ be a (globally) minimizing orbit of type $E_{p/q^+}$.

Consider the collection $C$ of sequences $\{x_i\}_{i=1}^{\infty}$ of type $E_{p/q^+}$. If $\{\eta_i\}$ is an orbit, then each interval $(\xi_i, \xi_{i+q})$ close enough to a point in $E_{p/q}$ contains exactly one of the $\eta_j$ and none of the $\xi_j$. We have

$$\sum (h(\eta_{i}, \eta_{i+1}) - h(\xi_{i}, \xi_{i+1})) \geq 0$$

We change
\[ h' = h(x, x') = h(x, x') + \phi(x), \]

where \( \phi(x) \) is a 'bump' function which is positive on the interval \((\xi_i, \xi_{i+q})\)
for some fixed \( i \) but has vanishing first and second derivatives on \( \xi_i \) and \( \xi_{i+q} \). Now \( (\xi_i) \) is a unique minimizing orbit. This proves that the property of having a unique global minimum of type \( E_{p/q^+} \) is dense.

For \( p/q \) fixed, suppose that fundamental domains of stable and unstable manifolds to \( E_{p/q} \) intersect finitely many times (open and dense). Then there are finitely many orbits of type \( E_{p/q^+} \). Since the intersections are transversal, their number is conserved under small perturbations. The value of the above sum then also changes continuously. Therefore, uniqueness is an open property.

\( \square \)
ACKNOWLEDGEMENTS

One of the authors (JJPV) has the pleasure to thank the Institute for theoretical Physics of the City University of Amsterdam and the Mathematics Department of Warwick University for their hospitality. He acknowledges useful discussions with Hans Capel, Ko van der Weele, Robert MacKay, and Maciej Wojtkowski. His research was supported by a postdoctoral fellowship at the Autonomous University in Madrid and by DOE grant DE-AC-02-83-ER-13044 at the Rockefeller University. The other (FMT) carried out his research at Queens College in Flushing (NY).

Both of us thank John Mather for a very illuminating conversation.
REFERENCES


A. Fathi, appendix to chapter one in : M. Herman, "Sur les courbes invariantes par les diffeomorphismes de l'anneau", Asterisque 103-104.


A. Olvera, C. Simo, "An Obstruction Method...", Central University Barcelona
J. Palis, W. de Melo, "Geometric Theory of Dynamical Systems" Springer
Verlag, 1982.
J.J.P. Veerman, F.M. Tangerman, "Scaling of Aubry Mather Cantor Sets",
J.J.P. Veerman, F.M. Tangerman, "Intersection properties of Invariant
FIGURE CAPTIONS

Figure 2.1: The cone $C_p$.

Figure 2.2: The clockwise enumeration of the invariant manifolds.

Figure 2.3: Construction of the curve $\gamma(p/q^+)$.

Figure 2.4: Horizontally shaded are interior components, vertically shaded are interior lobes.

Figure 2.5: The region $S_n$ in a small neighborhood of a hyperbolic periodic point.

Figure 4.1: The Single Intersection hypothesis.

Figure 4.2: The local invariant manifolds with rational rotation number of the gap $G$ and their $m(p/q^-)$-th image, drawn if, a): $m(p/q^+) > 0$, and b): $m(p/q^+) < 0$.

Figure 5.1: Resonance overlaps.
Figure 2.1
Figure 2.3
figure 4.1
Figure 4.2
StonyBrook
Institute for Mathematical Sciences
SUNY, Stony Brook, New York 11794-3651
telephone: (516) 632-7318
email: IMS@math.sunysb.edu

AVAILABLE AND FORTHCOMING PREPRINTS:

1990/1. B. BIELEFELD (editor), Conformal Dynamics Problem List.

/2. A. M. BLOKH & M. YU. LYUBICH, Measurable Dynamics of S-Unimodal Maps of the Interval.


/5. J. MILNOR, Dynamics in One Complex Variable: Introductory Lectures.

Publication of this preprint series is made possible in part
by a grant from the Paul and Gabriella Rosenbaum Foundation