Journal of Statistical Physics, Vol. 67, Nos. 3/4, 1992

A Rigorous Partial Justification of Greene's Criterion

Corrado Falcolini[1,2] **and Rafael de la Llave**[1]

Received October 4, 1991

We prove several theorems that lend support to Greene's criterion for the existence or not of invariant circles in twist maps. In particular, we show that some of the implications of the criterion are correct when the Aubry–Mather sets are smooth invariant circles or uniformly hyperbolic. We also suggest a simple modification that can work in the case that the Aubry–Mather sets have nonzero Lyapunov exponents. The latter is based on a closing lemma for sets with nonzero Lyapunov exponents, which may have several other applications.

KEY WORDS: Greene's criterion; breakdown of tori; twist maps; Aubry–Mather sets; periodic orbits; shadowing lemmas; Lyapunov exponents.

1. INTRODUCTION

In a remarkable paper, Greene[9] proposed a criterion for the existence of nontrivial invariant circles in twist mappings. Using it, he was able to compute the critical value at which golden circles ceased to exist with an accuracy that even today is unsurpassed and that, at the time of its appearance, was almost impossible to believe.

The purpose of this paper is to present some mathematically rigorous results that serve as a partial justification of Greene's criterion.

We recall that, given any number ω, Aubry–Mather theory establishes the existence of at least one set on which the motion is semiconjugate to a rotation of angle ω in the circle. Such sets enjoy several remarkable properties; among them, they are either Cantor sets or Lipschitz circles (we refer to ref. 21 for a review). From the practical point of view, it is quite

[1] Department of Mathematics, University of Texas at Austin, Austin, Texas 78712.
[2] Permanent address: Dipartimento di Matematica, Universitá degli Studi di Roma "Tor Vergata," 00133 Rome, Italy.

0022-4715/92/0500-0609$06.50/0

important to distinguish between these two possibilities, since an invariant circle is a complete barrier to long-scale transport and Cantor sets are not.

Greene's criterion asserts that an invariant circle exists if and only if a certain limit is positive. We show that, if the circle exists and is analytic or sufficiently differentiable, the number is 0 and, moreover, the limit is reached exponentially fast. On the other hand, if we have hyperbolic sets, the limit is a number bigger than one. Using techniques from Pesin theory, we can show that if there is a possibly not smooth invariant circle or a Cantor set with zero Lyapunov exponent, the lim sup of a related number is bounded by one, and that if there is a positive Lyapunov exponent, we can get a subsequence converging to a positive number.

The practical importance of Greene's criterion is that the limit is computed on periodic orbits which are quite easy to compute.

There is considerable evidence that Greene's criterion is correct, at least in some cases.

First there is the agreement between the quantitative values obtained rigorously. Recently the methods of *computer-assisted proofs* have been applied to the problem of computing the range of applicability of the KAM theorem. For particular examples, there are positive results on values for which the theorem does apply[5,6,16,35] as well as results on values for which the conclusions of the theorem are false.[13,19–20] Notice that the values in ref. 13 and 16 differ by about 7% and that the value obtained by Greene's method is in the allowed interval. The value of ref. 13 agrees to several decimal places with the value of ref. 9.

Ref. 30 and 31 introduce another method that not only establishes the nonexistence of invariant circles, but also that the invariant set of golden mean rotation is hyperbolic. Even if the implementation of the criterion in ref. 30 is not completely rigorous because it ignores the effects of roundoff error, the authors have performed a very careful analysis that makes the results of the paper quite close to a proof. It seems that the algorithm proposed is within the reach of computer-assisted proofs. The agreement of these numerical results with Greene's value is quite remarkable and lends support to the conjecture that, for the standard family, as soon as the invariant circle disappears, it becomes a hyperbolic invariant set.

Besides the rigorous numerical results indicated above, there are arguments based on the renormalization group that lend credence to the Greene's method. There is considerable evidence that the phenomenon of breakup of invariant tori can be described for a large class of families by a renoemalization group picture.[26,27] (Indeed, the arguments for the existence of a fixed point and the linearization of the spectrum of these two papers are quite close to being a proof.) This renormalization group picture implies that all dynamical quantities have *bulk properties* and that to compute the parameter value at which a transition occurs we can use as indicator whatever property is more convenient to measure. (This is quite similar to the fact that we can measure the boiling point of water by examining electric or magnetic properties, density, etc.)

The scaling properties predicted by the renormalization group for periodic orbits can be displayed quite dramatically in *fractal diagrams*[37] and can be used to improve the numerical effectiveness of Greene's method (ref. 26, §4.6.2, §4.6.3).

We should nevertheless point out that the renormalization group picture gets considerably more complicated when the families are slightly different from the standard one,[41,42,14] which can be explained by saying that the dynamics of the renormalization operator has basins in which the dynamics is controlled by a more complicated landmark than a simple fixed point, as exhibited in refs. 26 and 27. This matter merits further investigation.

In this paper, we present some rigorous results which are independent of the renormalization group picture, but rather use standard techniques from KAM and from hyperbolic perturbation theories.

We consider Greene's method as part of a long tradition in mathematics of using periodic orbits, the simplest landmark that organizes the long-term behavior as the skeleton on which to study dynamical properties. Perhaps the forerunner of this approach was Poincaré (see, e.g., ref. 34, Vol. I, p. 82):

> There is more: here is a fact that I have not been able to demonstrate rigorously, but that seems to me nonetheless very reasonable. Given equations of the form defined in no. 13 [Hamiltonian equation] and a particular solution of these equations one can always find a periodic solution (whose period, it is true, can be very long) such that the difference between the two solutions may be as small as one wishes for as long as one wishes. Moreover, what makes these periodic solutions so valuable is that they are, so to speak, the only opening by which we may try to penetrate into a place which until now is reputed to be inaccessible.

2. NOTATION AND STATEMENT OF RESULTS

Let $f\colon \mathbf{T}^1 \times \mathbf{R} \to \mathbf{T}^1 \times \mathbf{R}$ be an analytic, area-preserving map.

Let $x \in \mathbf{T}^1 \times \mathbf{R}$ satisfy $f^N(x) = x$. We say that it is a periodic orbit of type M/N, $N \in \mathbf{N}$, $M \in \mathbf{Z}$ if, denoting by $\tilde{f}$ and $\tilde{x}$ the lifts of f and x to the universal cover of $\mathbf{T}^1 \times \mathbf{R}$, we have $\tilde{f}^N(\tilde{x}) = \tilde{x} + (M, 0)$. We denote the orbit of a periodic point by $o(x)$.

For such an orbit Greene defined the "residue" by

$$R(x) = \tfrac{1}{4}[\mathrm{Tr}(Df^N(x)) - 2] \qquad (2.1)$$

Greene defined the "mean residue" to be $[R(x)]^{1/N}$ and observed numerically that, if M_i/N_i were the continued convergents of an irrational numer ω and x_i are points of type M_i/N_i, then $[R(x_i)]^{1/N_i} \to \rho(\omega)$ and that $\rho(\omega) > 1$ when there is no invariant circle and that $\rho(\omega) < 1$ when there is an invariant circle.

The practical importance of this criterion lies in the fact that there are quite efficient methods for the computation of periodic orbits. Moreover, by computing the residues of a significan number of periodic orbits, we can get an idea of the set of rotation numbers for which there is an invariant circle.

We notice that

$$R(x) = \tfrac{1}{4}[\mathrm{Tr}(Df(f^{N-1}(x)) \cdots Df(x)) - 2]$$

so that, using the invariance of the trace under cyclic permutations, the residue is the same for all points in an orbit. We also emphasize that the residue is invariant under a C^1 change of coordinates.

Notice that the residue of a periodic orbit can be easily related to the eigenvalues of the derivative of the return map. Hence, it is natural that the Lyapunov exponents come into play when the residue grows exponentially fast. If one eigenvalue is λ, by the preservation of area, the other one should be $1/\lambda$ and the trace is $\lambda + 1/\lambda$. If $\lambda = \exp(\gamma N)$, the mean residue should be $\approx e^{\gamma}$ if $N\gamma$ is large enough. In particular, if γ_n are the Lyapunov exponents of orbits of period N_n and $\gamma_n \to \gamma$, $N_n \to \infty$, the mean residue converges to e^{γ} if $\gamma > 0$. If $\gamma = 0$, the limit of the residue could be any number between 0 and 1, depending on the relative rates of convergence of N_n and γ_N.

We recall that a number ω is called Diophantine if, for every $p, q \in \mathbf{N}$ we have

$$|\omega - p/q| \geqslant K\,|q|^{-\nu} \tag{2.2}$$

These numbers play an important role in KAM theory. We also recall that the convergents of the continued-fraction expansion of a number ω satisfy $|\omega - p/q| \leqslant K/q^2$, so the best exponent ν we can hope to have in (2.2) is 2. The numbers for which it is possibly to satisfy (2.2) with $\nu = 2$ are called *constant type numbers* and, even if they have measure zero, they include all quadratic irrationals and, in particular, are dense. If we take any $\nu > 2$, the set of Diophantine numbers with this exponent has full measure.

Theorem 2.1. Assume that f as above admits a topologically nontrivial analytic invariant circle and that the motion on it is analytically conjugate to a rotation ω such that

$$\lim \frac{1}{N} \sup_{q \leqslant N} \log \left|\omega - \frac{p}{q}\right| = 0 \tag{2.3}$$

Then, for every $k \in \mathbf{N}$, we can find $C_k > 0$, depending on f and on the circle, such that for every N, M such that $|\omega - M/N| \leqslant 1/N$ and any periodic point x of type M/N, we have

$$|R(x)| \leqslant C_k \left| \omega - \frac{M}{N} \right|^k N$$

In particular, if $|\omega - M_i/N_i| \leqslant K/(N_i)^2$ (e.g., if M_i/N_i are the continued-fraction convergents to ω), then $\limsup |R(x_i)|^{1/N_i} \leqslant 1$.

Remark. The same method of proof establishes that if ω is Diophantine and the circle and the map are C^r, then, if x is a periodic orbit of type M/N, we have $R(x) \leqslant C_k |\omega - M/N|^k$ for all $k \leqslant k^*(r)$, where $k^*(r)$ depends on the exponent v in (2.2), but $k^*(r) \to \infty$ as $r \to \infty$.

For Diophantine numbers, the previous result can be improved from the residue being smaller than any power to being exponentially small.

Theorem 2.2. Let f be as before, ω as in (2.2). Assume that

$$\sup_{|\mathrm{Im}\, \varphi| \leqslant \delta} |f(A, \varphi)| \leqslant \Gamma \leqslant \infty, \qquad \sup_{|\mathrm{Im}\, \varphi| \leqslant \delta} |f^{-1}(A, \varphi)| \leqslant \Gamma \leqslant \infty$$

and that there is a mapping $K: \mathbf{T}^1 \to \mathbf{T}^1 \times \mathbf{R}$ with $f(K(\varphi)) = K(\varphi + \omega)$ and that $\sup_{|\mathrm{Im}\, \varphi| \leqslant \delta} |K(\varphi)| \leqslant \Gamma$.

Then, there exists a constant $D > 0$—depending on the Diophantine properties of the number ω and the analyticity properties of the tori—such that for every periodic orbit x of type M/N with $|\omega - M/N| \leqslant 1/N$

$$|R(x)| \leqslant D \exp(-D\delta\, |\omega - M/N|^{-1/1+v}) \tag{2.4}$$

Remark. Notice that if M, N saturate the bounds in the Diophantine inequality for ω, we have $|\omega - M/N| \leqslant K/N^v$. Hence $|R| \leqslant D \exp(-DN^{v/1+v})$.

Remark. The fact that the residues converge exponentially fast to zero when there is an analytic invariant circle is one of the predictions of the renormalization group analysis. Notice that, when one knows that the convergence of a sequence to its limit is exponentially fast, it is possible to use Aitken extrapolation (ref. 38, §5.10) to compute the limit more effectively. This leads to more effective implementations of Greene's method. This idea is suggested in ref. 26, §4.6.3. This renormalization group analysis suggests that the exponent of $|\omega - M/N|$ computed in (2.4) is not optimal. The problem of computing the optimal exponent in (2.4) is very similar to optimizing the exponent in the Nekhoroshev theorem.

Remark. The conclusion of Theorem 2.2 suggests that there is a relation between the exponent of decrease of the residue and the analyticity domain of the circle. Unfortunately, the statement we have proved is not enough to conclude that. Notice that the coefficient also depends on Γ, which depends on the analyticity properties of the circle. The main reason to conjecture that such a relation should exist is that both of them scale with the renormalization group in the same way.

We now proceed to state our results for the case in which the Aubry–Mather sets are hyperbolic.

Theorem 2.3. Assume that Γ is a hyperbolic Aubry–Mather set of rotation number ω and that $\{x_n\}$ is a sequence of periodic points of type M_n/N_n such that $o(x_n)$ converges to Γ. Then, for sufficiently large n, $|R(x_n)|^{1/N_n} > \lambda > 1$. Actually, if the hyperbolic set has Lyapunov exponent γ, then $\lim_n R(x_n)^{1/N_n} = e^\gamma$.

Theorem 2.4. Let f be a C^2 twist mapping as above and let Γ be an Aubry–Mather Cantor set with rotation number $\omega \notin \mathbf{Q}$. If $f|_\Gamma$ has a Lyapunov exponent γ, then:

(a) For any sequence x_n of periodic orbits of type M_n/N_n converging to Γ, $\limsup_n \|Df^{N_n}(x_n)\|^{1/N_n} \leqslant e^\gamma$.

(b) If $\gamma > 0$, there exists a sequence of periodic points x_n of type M_n/N_n converging to Γ such that $\lim_n R(x_n)^{1/N_n} = e^\gamma$.

Remark. In principle, when we use Lyapunov exponents, we should specify with respect to which ergodic measure we take them. Nevertheless, as we will discuss in the proof of Theorems 2.3 and 2.4, for Aubry–Mather sets with irrational rotation number there is only one invariant measure with support in the set, so that the notation is unambiguous.

Remark. Claim (a) is much easier to prove than claim (b). In fact, claim (a) is an abstract statement about uniquely ergodic systems (notice that it does not claim that periodic orbits exist) and, given the remarks about the relation of mean residue and Lyapunov exponents we made before, it only requires to show that $\limsup \gamma(x_n) \leqslant \gamma$. Claim (b), on the other hand, uses methods of Pesin theory and establishes the existence of periodic orbits. It is based on proving a shadowing lemma for sets with nonzero Lyapunov exponents that also controls the Lyapunov exponents.

Remark. Notice that in the case where Γ is a Cantor set, there could be several Aubry–Mather sets with the same rotation number (see, e.g., ref. 24); hence, to verify experimentally the claim, one would have to verify the convergence of the approximating periodic orbits to the Cantor set. In practice, if one uses the algorithm of the critical lines of ref. 9, this is not a problem since the periodic orbits are found on a vertical line. If one uses other algorithms, one has to verify that independently. In the case that Γ is an invariant circle, it is not difficult to show, using the twist condition, that there can be no other invariant sets with the same rotation number, so that part (a) of Theorem 2.4 can be improved to:

Let M_n/N_n converge to ω and let x_n be periodic orbits of type M_n/N_n. If there is an invariant circle of rotation number ω, then $\limsup \|Df^{N_n}(x_n)\|^{1/N_n} \leqslant 1$.

We remark that a sketch of a method of proof of Theorems 2.1 and 2.2 has been available for a long time. In particular, it was suggested by John Mather as early as 1982 (see, e.g., ref. 26, p. 1.3.2.4). Nevertheless, we thought it would be worth publishing a detailed account of these arguments, since fairly quantitative results are needed in subsequent numerical work by the authors.[7]

The method we present here is optimized for computability and it does not require the performance of successive changes of variables. It is also written in such a way that it readily generalizes to higher number of variables or to the case when the values of some of the parameters are complex. The latter is used essentially in ref. 7.

The proof of Theorem 2.3 is a standard result of the perturbation of hyperbolic structures.

Theorem 2.4 is a basic result about the approximation of non-uniformly hyperbolic dynamical systems by periodic orbits. Except for the quantitative results of the Lyapunov exponent, part (b) is the main lemma in ref. 43. Related results appear in ref. 25. The proof we present here is based on a shadowing lemma for partially hyperbolic systems, which has other applications. The method of proof is inspired by the treatment of hyperbolic sets in ref. 18.

Upper bounds for the residues similar to ours as well as a consideration of the uniformly hyperbolic cases have been proved in ref. 28 by different methods.

3. PROOF OF THE RESULTS

3.1. Proof of Theorem 2.1 and Theorem 2.2

The basic idea in the proof of Theorem 2.1 is to show that given $k \in \mathbb{N}$, we can find a complex neighborhood U_k of the invariant circle Γ, an integrable mapping I_k, and a constant C_k in such a way that

$$\|f - I_k\| \leqslant C_k \operatorname{dist}(x, \Gamma)^k$$

Then, an elementary perturbation argument would allow us to estimate the trace of the derivatives of orbits that stay close to the invariant circle. It will be a corollary of Moser's twist mapping theorem that the maximum distance of a periodic orbit to the invariant circle can be estimated—in the appropriate coordinates—by the difference between the rotation numbers of the orbit and the circle.

The construction of an integrable system will be done by finding an approximate integral.

It will simplify the notation to choose an appropriate system of coordinates.

Proposition 3.1. Let $f: \mathbf{T}^1 \times \mathbf{R}$ be as in Theorem 2.1 and Γ be an invariant circle, $f|_\Gamma$ analytically conjugate to a rotation ω. Then, we can find a globally canonical analytic mapping h defined in a neighborhood of Γ, with an analytic inverse in a neighborhood of Γ, and such that

$$h \circ f \circ h^{-1}(A, \varphi) = (A + A^2 u(A, \varphi), \varphi + \omega + A v(A, \varphi))$$

with u, v analytic,

$$\frac{\partial A u}{\partial A} \geqslant \alpha > 0 \qquad \text{for} \quad |A| \leqslant \varepsilon, \quad \varphi \in \mathbf{T}^1$$

Proof. By Birkhoff's theorem,[8,20] we know that Γ is the graph of an analytic function $\gamma: \mathbf{T}^1 \to \mathbf{R}$.

The transformation $h_1: \mathbf{T}^1 \times \mathbf{R} \circlearrowleft$ defined by

$$h_1(A, \varphi) = (A + \gamma(\varphi), \varphi)$$

is globaly symplectic and sends the circle $\mathbf{T}^1 \times \{0\}$ into the graph of γ. Hence, $h \circ f \circ h_1^{-1}$ leaves invariant the circle $\mathbf{T}^1 \times \{0\}$. Hence

$$h_1 \circ f \circ h_1^{-1}(A, \varphi) = (A u_1(A, \varphi), v_1(A, \varphi))$$

Since the motion on this circle is conjugate to a rotation, there exists an analytic $\delta: \mathbf{T}^1 \to \mathbf{T}^1$ with an analytic inverse [hence $\delta'(\varphi) \neq 0$] such that $v_1(0, \delta(\varphi)) = \delta(\varphi + \omega)$.

The transformation $h_2(A, \varphi) = (A/\delta'(\varphi), \delta(\varphi))$ is globally canonical and $h_2^{-1} \circ h_1 \circ f \circ h_1^{-1} \circ h_2$ is of the form

$$(A, \varphi) \to (A', \varphi') \equiv (A u_2(A, \varphi), \varphi + \omega + A v_2(A, \varphi))$$

Since the map preserves volume,

$$\det \begin{pmatrix} \partial A'/\partial A & \partial A'/\partial \varphi \\ \partial \varphi'/\partial A & \partial \varphi'/\partial \varphi \end{pmatrix} = 1$$

and since

$$\left.\frac{\partial\varphi'}{\partial\varphi}\right|_{A=0}=1, \qquad \left.\frac{\partial A'}{\partial\varphi}\right|_{A=0}=0$$

we should have

$$\left.\frac{\partial A'}{\partial A}\right|_{A=0}=1$$

That is the form of the map claimed in Proposition 3.1.

It is a simple calculation to show that the transformations h_1, h_2 preserve the positive twist condition. Hence, the last inequality in the claim is established. ▌

Lemma 3.2. Let f be as in Proposition 3.1, ω Diophantine.

Given any $k \in \mathbf{N}$, we can find analytic functions $H_0(\varphi),..., H_k(\varphi)$ so that $H=\sum_{i=0}^{k} A^i H_i(\varphi)$ satisfies $|H \circ f - H| \leqslant C_{k+1} A^{k+1}$.

Proof. We will derive a hierarchy of equatons and show that we can solve them recursively.

We observe that

$$H \circ f(A, \varphi)=\sum (A+A^2 u(A, \varphi))^i\, H_i(\varphi+\omega+Av(A, \varphi)) \tag{3.1}$$

Moreover, if we expand $H_i(\varphi+\omega+Av(A, \varphi))$ using Taylor's formula in A, we obtain

$$H_i(\varphi+\omega+Av(A, \varphi))=\sum_{i=0}^{N} H_i^j(\varphi)\, A^{i+j}+\mathcal{O}(A^{i+N+1}) \tag{3.2}$$

where $H_i^0(\varphi)=H_i(\varphi+\omega)$, $H_i^1(\varphi)=H_i'(\varphi+\omega)\, v(0, \varphi)$. For higher j, H_i^j is an expression involving derivatives of H_i and of v. We observe that the derivatives entering in H_i^j are of order up to j and that the derivatives of H_i enter linearly.

If we substitute (3.2) into (3.1), we obtain

$$H \circ f(A, \varphi)=\sum_{i=0}^{N} A^i(H_i(\varphi+\omega)+H_{i-1}(\varphi+\omega)\, u(0, \varphi)+L_i(\varphi)) + \mathcal{O}(A^{N+1}) \tag{3.3}$$

where L_i is an expression involving $H_0,..., H_{i-2}$ and their derivatives of order up to i as well as derivatives of H_{i-1}. We emphasize that H_{i-1} only enters in L_i in the form of derivatives, so that, if H_{i-1} changes by a constant, L_i remains unaltered.

If we equate the term of A^i in (3.3) with that in the expansion for H, we are led for $i>0$ to a hierarchy of equations of the form

$$H_i(\varphi+\omega)+u(0,\varphi)\,H_{i-1}(\varphi+\omega)+L_i(\varphi)=H_i(\varphi) \tag{3.4}$$

We recall the following:

Proposition 3.3. Let $\eta\colon \mathbf{T}^1\to\mathbf{R}$ be an analytic function, $\int_{\mathbf{T}^1}\eta=0$. Let ω be a Diophantine number. Then, there exists $H\colon \mathbf{T}^1\to\mathbf{R}$ analytic satisfying

$$H(\varphi+\omega)-H(\varphi)=\eta(\varphi)$$

Moreover, H is unique up to an additive constant. In particular, all the derivatives of H are uniquely determined.

Proof. The proof of Proposition 3.3 is quite well known and is obtained just matching the Fourier coefficients. Details can be found, among other places, in ref. 1, §12, and ref. 39, §32. ∎

Using Proposition 3.3, it is possible to solve all the equations in (3.4).

We assume inductively that $H_0,\ldots,H_{i-2}$ are determined and that H_{i-1} is determined up to an additive constant. Since L_i depends only on $H_0,\ldots,H_{i-2}$ and the derivatives of H_{i-1}, we see that L_i is determined. Using the twist condition, we have $\int u(0,\varphi)\neq 0$, so that it is possible to determine uniquely the additive constant in H_{i-1} by imposing

$$\int u(0,\varphi)\,H_{i-1}(\varphi)+\int L_i(\varphi)=0$$

Using Proposition 3.3, H_i is determined up to an additive constant, so that we recover the induction hypothesis with $i-1$ replaced by i.

The first step of the induction reduces to an obvious identity. ∎

Notice that, if H is a conserved quantity, so is any function of H.

Observe also that the curves $H=h$, for small $|h|$, are homotopically nontrivial since H is a small perturbation of A.

We can define

$$\tilde{H}(h)=\int_{H=h} A\,d\varphi$$

The function $H^* = \tilde{H}(H)$ will be conserved up to $O(A^{k+1})$ and it has the property that

$$\int A \, d\varphi = h$$

$H^* = h$.

We can now define a canonical transformation in such a way that H^* becomes the action variable.

In effect, if we can find an S in such a way that

$$\begin{aligned} H^* &= A - \frac{\partial S(H^*, \varphi)}{\partial \varphi} \\ \varphi' &= \varphi + \frac{\partial S}{\partial H^*}(H^*, \varphi) \end{aligned} \tag{3.5}$$

then the transformation $(A, \varphi) \to (H^*, \varphi')$ will be canonical.

Using the first equation of (3.5), we can determine S up to the addition of a function of H^*.

We can determine this additive function in such a way that $\varphi'(A, 0) = 0$.

Expressed in the coordinates (H, φ'), the mapping f has the form

$$(H, \varphi') \xrightarrow{\tilde{f}} (H, \varphi' + \omega + H\Delta(H)) + R(H, \varphi')$$

where $|R| \leqslant C_N H^N$.

We emphasize that, since all the changes of variables are analytic, the estimates on the remainder remain true in a complex neighborhood of $(\mathbf{T}^1 \times \{0\})$ of the form $\{|\text{Im}\, \varphi'| \leqslant \xi,\ |A| \leqslant \xi\}$. As a consequence, $\|DR\| \leqslant CH^{N-1}$. Hence,

$$D\tilde{f}(H, \varphi') = \begin{pmatrix} 1 & \Gamma(H) \\ 0 & 1 \end{pmatrix} + O(H^{N-1})$$

We notice that the trace of the derivative of a periodic point—hence the residue—can be computed in any system of coordinates.

Since

$$D\tilde{f}^N(H, \varphi') = D\tilde{f}(\tilde{f}^{N-1}(H, \varphi'))\, D\tilde{f}(\tilde{f}^{N-2}(H, \varphi')) \cdots D\tilde{f}(H, \varphi')$$

we will find it useful to estimate eigenvalues of products of matrices close to upper triangular.

Lemma 3.4. Let $\{A_i\}_{i=1}^N$ be a set of 2×2 matrices of the form

$$A_i = \begin{pmatrix} 1 & a_i \\ 0 & 1 \end{pmatrix}$$

with $\sup_{1 \leqslant i \leqslant N} |a_i| \leqslant A$.

Let $\{B_i\}_{i=1}^N$ satisfy

$$\sup_{\substack{1 \leqslant i \leqslant N \\ j,k=1,2}} |(B_i)_{jk} - (A_i)_{jk}| \leqslant \varepsilon \qquad \text{with } \varepsilon \leqslant A$$

Then $B = B_1, \ldots, B_N$ satisfies

$$|\mathrm{Tr}\, B - 2| \leqslant 2[(1 + 3\sqrt{A}\sqrt{\varepsilon})^N - 1]$$

Proof. Given any norm on 2-vectors, if we define $\|C\| = \sup_{v \in \mathbf{R}^2} \|Cv\|/\|v\|$, clearly all eigenvalues of C have modulus not bigger than $\|C\|$. Hence, for a 2×2 matrix C, $\mathrm{Tr}\, C \leqslant 2\, \|C\|$.

If we define $\|v\| = |v_1|\, \delta + |v_2|$, then

$$\left\| \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \right\| \leqslant \max(|C_{11}| + \delta^{-1}\, |C_{21}|, |C_{21}|, |C_{21}|\, \delta + |C_{22}|)$$

In particular, for matrices such as those in the hypothesis of Lemma 3.4 and for $\delta \leqslant 1$

$$\begin{aligned} \|A_i\| &\leqslant 1 + |a_i|\, \delta \leqslant 1 + A\delta \\ \|A_i - B_i\| &\leqslant \varepsilon \max((1 + \delta^{-1}), (1 + \delta)) = \varepsilon(1 + \delta^{-1}) \end{aligned} \tag{3.6}$$

We can write

$$\begin{aligned} B &\equiv B_1 \cdots B_N \\ &= (A_1 + (B_1 - A_1))(A_2 + (B_2 - A_2)) \cdots (A_N + (B_N - A_N)) \end{aligned}$$

Expanding and grouping by the same factors $(B_i - A_i)$, we find

$$\begin{aligned} B = {} & A_1 \cdots A_N \\ & + \sum_i A_1 \cdots A_{i-1}(B_i - A_i)\, A_{i+1} \cdots A_N \\ & + \sum_{i,j} A_1 \cdots A_{i-1}(B_i - A_i)\, A_{i+1} \cdots A_{j-1}(B_j - A_j)\, A_{j+1} \cdots A_N \\ & + \cdots \\ & + (B_1 - A_1) \cdots (B_N - A_N) \end{aligned}$$

The trace of the first term is 2 and the trace of the other terms can be bounded by twice the norm. Using the estimates of the norms in (3.6) and bounding the norms of the products by the product of the norms of the factors, we can bound the residue by

$$2\binom{N}{1}(1+A\delta)^{N-1}(1+\delta^{-1}\varepsilon+2\binom{N}{2}(1+A\delta)^{N-2}[(1+\delta^{-1})\varepsilon]^2$$

$$+\cdots+2\binom{N}{2}(1+\delta^{-1}\varepsilon)^N$$

$$=2\{[1+A\delta+(1+\delta^{-1})\varepsilon]^N-1\}$$

If we choose $\delta=(\varepsilon/A)^{1/2}$, which is smaller than 1, the upper bound for the residue we just computed becomes

$$2[(1+\sqrt{A}\sqrt{\varepsilon}+\varepsilon+\sqrt{A}\sqrt{\varepsilon})^N-1]$$

Since $\varepsilon\leqslant A$, $\varepsilon\leqslant\sqrt{A}\sqrt{\varepsilon}$ and we obtain the bound in the claim of the lemma. ▌

The next ingredient in the proof is an argument that says that periodic orbits of rotation number close to that of Γ are contained in a small strip near $H=0$.

Notice that, even if it is not difficult to show that most of the points should be close enough (otherwise the twist would force the rotation to be much bigger), we want the much stronger property that all the points of the orbit are close to the invariant circle.

Lemma 3.5. If $|\omega-M/N|$ is small enough, all orbits of type M/N are contained in the strip

$$|H|\leqslant\left|\omega-\frac{M}{N}\right|K$$

where K depends only on the system and on the circle.

Proof. By Moser's twist theorem, we can find invariant circles whose rotation numbers $\omega\in\Omega$.

Moreover, $\mu(\Omega\cap[-\Delta_0+\omega,\Delta+\omega])/2\Delta\to 1$, where μ denotes the Lebesgue measure.

It follows that if M/N is close enough to ω, there are going to be points ω' of Ω in $[M/N, M/N+|\omega-M/N|]$.

Furthermore, since the mapping that associates to a rotation number the invariant circle of this rotation number is Lipschitz, the circle of rotation number ω' is contained in

$$|H| \leqslant K\,|\omega - \omega'| \leqslant 2K \left|\omega - \frac{M}{N}\right|$$

By the twist property, the orbit of rotation number M/N has to be contained between the circles of rotation number ω' and ω. ▮

Remark. Notice that the dependence of K on the system and on the circle is rather weak. It is, roughly, the Lipschitz constant in the mapping that associates to a rotation number a KAM circle when we topologize the circles with the C^0 norm. In particular, it can be chosen uniformly in a sufficient C^5 neighborhood of the integrable case. If we know that a map has a sufficiently differentiable circle, it can be chosen uniformly in a C^5 neighborhood.

Using Lemma 3.4 and 3.5, it follows that, for every k,

$$R_{M/N} \leqslant 2\left[\left(1 + C_k K \left|\omega - \frac{M}{N}\right|^k\right)^N - 1\right] \tag{3.7}$$

where K and C_k are the constants respectively in Lemmas 3.5 and 3.2.

If $|\omega - M/N|^k\, C_k KN$ is sufficiently small and N is sufficiently large, we can bound the RHS of (3.7) by

$$8C_k K \left|\omega - \frac{M}{N}\right|^k N \tag{3.8}$$

Since C_k is an arbitrary constant, multiplying it by 8 does not change anything, so that we can denote it by the same letter. This finishes the proof of Theorem 2.1. ▮

Remark. The method carried out above can be generalized to higher dimensions. First, the normal form given by Proposition 3.1 can be carried out with the only modification that, rather than using the determinant of the transformation being 1, we have to use the preservation of the symplectic form. Moreover, it is possible to use an analogue of (3.4) to compute as many independent approximate conserved quantities as the dimension of the tori. We point out that an alternative approach to compute similar normal forms can be found in ref. 40 based on the use of generating functions and successive transformations. Even if, from the point of view of theoretical calculations, both methods could be used, the method explained here lends itself to quite efficient computer implementations, so that it should be possible to obtain good estimates of the residues in concrete cases as well as estimates of the times of escape from neighborhoods of the tori in higher dimensions.

Notice that, for any k, (3.8) produces a valid estimate of the residue. The proof of Theorem 2.2 will consist only in estimating explicitly the C_k so that for given N, M we can choose the k that gives the best bound.

We recall that we had to solve for H_i in

$$H\circ f=\sum_{i=0}^{k}(A+A^2u(A,\varphi))^i\,H_i(\varphi+\omega+Av(A,\varphi))$$
$$=\sum_{i=0}^{k}A^iH_i(\varphi)+\mathcal{O}(A^{k+1}) \tag{3.9}$$

If we write $\theta=\varphi+\omega+Av(A,\varphi)$, we can write (3.9) as

$$H\circ f=\sum_{i=0}^{k}(A+A^2\tilde{u}(A,\theta))^i\,H_i(\theta))$$
$$=\sum_{i=0}^{k}A^iH_i(\theta-\omega-A\tilde{v}(A,\theta))+\mathcal{O}(A^{k+1}) \tag{3.10}$$

where $\tilde{u}$, $\tilde{v}$ are analytic functions whose domain of analyticity depends only on the properties of u, v. Also, $\tilde{u}(0,\phi)>0$.

We will assume that they are defined in a domain of the form $\{\theta \mid |\mathrm{Im}(\theta)|\leqslant\delta\}$ and that their absolute values there are bounded by a constant K.

If $H\colon \mathbf{T}^1\mapsto\mathbf{C}$ is an analytic function, we will denote

$$\|H\|_\eta=\sup_{|\mathrm{Im}\,\theta|\leqslant\eta}|H(\theta)|$$

As before, we can solve (3.10) recursively. Expanding both sides in powers of A and equating the factors of A^i, we obtain

$$H_i(\theta)+H_{i-1}(\theta)\,u(0,\phi)+L_i^1(\theta)=H_i(\theta-\omega)+L_i^2(\theta) \tag{3.11}$$

where, as before, L_i^1, L_i^2 are expressions involving $H_0,\ldots,H_{i-2}$ and their derivatives and the derivatives of H_{i-1}.

The procedure to solve (3.11) is very similar to the one that we used in the proof of Theorem 2.1.

We assume inductively that $H_0,\ldots,H_{i-2}$ are determined completely and that H_{i-1} is determined up to an additive constant. Then, we determine the additive constant in H_{i-1} in such a way that

$$\int (H_{i-1}u(0, \varphi) + L_i^1(\varphi) - L_i^2(\varphi))\, d\varphi = 0$$

Then, using Proposition 3.3, we can determine H_i up to a constant.

We denote $\bar{H}_i = \int H_i(\theta)\, d\theta$ and $\tilde{H}_i(\theta) = H_i(\theta) - \bar{H}_i$.

Lemma 3.6. Equations (3.11) can be solved recursively. For ω Diophantine, if $\delta - k\eta > 0$, we have

$$\begin{aligned} \|\tilde{H}_i\|_{\delta - i\eta} &\leqslant E(D)^i \\ |\bar{H}_i| &\leqslant E(D)^i \end{aligned} \tag{3.12}$$

where D is a number of the form $D = \tilde{K}\eta^{-1-\nu}$ and $\tilde{K}$ depends on the system but can be taken uniformly in a $\|\cdot\|_\delta$ neighborhood. Similarly for E.

Proof. The quantitative statements in (3.12) will be obtained by estimating all the steps in the above construction.

We recall that, by definition,

$$L_i^2(\varphi) = \sum_{j=1}^{i} \frac{1}{j!} \left(\frac{\partial}{\partial A}\right)^j H_{i-j}(\theta - \omega - A\tilde{v}(A, \theta)) \bigg|_{A=0}$$

If we denote

$$K = \sup_{\substack{|A| \leqslant \delta \\ |\mathrm{Im}(\theta)| \leqslant \delta}} |\tilde{v}(A, \theta)|$$

we can bound

$$\sup_{\substack{|A| \leqslant \eta/2K \\ |\mathrm{Im}(\theta)| \leqslant \delta - (i-1/2)\eta}} |H_{i-j}(\theta - \omega - Au(a, \theta))|$$
$$\leqslant \|H_{i-j}\|_{\delta - (i-1/2)\eta + 1/2\eta} \leqslant \|H_{i-j}\|_{\delta - j\eta}$$

Using Cauchy estimates in the variable A, we obtain that

$$\sup_{|\mathrm{Im}(\theta)| \leqslant \delta - (i-1/2)\eta} \left| \frac{1}{j!} \left(\frac{\partial}{\partial A}\right)^j H_{i-j}(\theta - \omega - A\tilde{v}(A, \theta)) \bigg|_{A=0} \right|$$
$$\leqslant \|H_{i-j}\|_{\delta - j\eta}\, \eta^{-j}(2K)^j$$

Hence, if we substitute the induction hypothesis, we obtain that

$$\|L_i^2\|_{\delta - i\eta + \eta/2} \leqslant \sum_{j=1}^{i} \left(\frac{2K}{\eta}\right)^j ED^{i-j}$$
$$= D^{i-1} E \frac{2K}{\eta} \sum_{j=1}^{i} \left(\frac{2K}{D\eta}\right)^{j-1} \leqslant D^{i-1} E \frac{4K}{\eta}$$

Similarly, we can obtain bounds for L_i^1. We observe that

$$L_i^1(\theta) = \sum_{j=2}^{i} H_{i-j} \frac{1}{j!} \left(\frac{\partial}{\partial A}\right)^j (A + Au(A, \theta))^j \bigg|_{A=0}$$

We can estimate the derivatives using Cauchy estimates to obtain

$$\left| \frac{1}{j!} \left(\frac{\partial}{\partial A}\right)^j (A + Au(A, \theta))^j \bigg|_{A=0} \right| \leqslant \left(\frac{K+1}{\eta}\right)^j$$

when $|\mathrm{Im}(\theta)| \leqslant \delta$. Hence,

$$\|L_i^1\|_{\delta - i\eta + 1/2\eta} \leqslant \sum j = 2^i D^{i-j} \left(\frac{K+1}{\eta}\right)^j \leqslant D^{i-1} \frac{2K+1}{\eta}$$

Since $\|\tilde{H}_{i-1} i\|_{\delta - i\eta} \leqslant D^{i-1}$, we see that we can determine $\bar{H}_{i-1}$ and that it satisfies

$$|\bar{H}_{i-1}| \leqslant \tilde{K}/\eta$$

where $\tilde{K}$ depends only on the suprema of $\tilde{u}$ and $\tilde{v}$ and can be chosen uniformly, as η is arbitrarily small.

We can now apply a quantitative version of Proposition 3.3 that is proved in the same references quoted before.

Lemma 3.7. Let ω satisfy (2.2). Then, for every $L\colon \mathbf{T}^1 \mapsto \mathbf{C}$ analytic, satisfying $\int L(\theta)\, d\theta = 0$, we can find a unique $H\colon \mathbf{T} \mapsto \mathbf{C}$ satisfying

$$H(\theta) - H(\theta + \omega) = L(\theta)$$

$$\int H(\theta)\, d\theta = 0$$

Moreover, for any $\eta > 0$ we have

$$\|H\|_{\delta - \eta} \leqslant C\eta^{-\nu} \|L\|_\delta$$

Applying Lemma 3.7, we obtain $\|\tilde{H}_i\|_{\delta - i\eta} \leqslant D^{i-1} \tilde{K} \eta^{-1-\nu}$.
So, we see that the induction hypothesis are recovered.

To conclude the proof of Theorem 2.2, we just observe that if we perform k operations, we can take η as big as $\delta/2k$ and still satisfy the condition that $\delta - k\eta > 0$.

We also observe that the same argument that we used to bound $L_i^1(\theta) + H_{i-1}u(0, \theta) - L_i^2(\theta)$ serves to bound the ith derivative with respect to A of $H \circ f - H$ in a complex neighborhood for A.

In the notation of Theorem 2.1, we have established that $C_k \leqslant \tilde{K}(k/\delta)^{k(1+\nu)}$.

An elementary computation of maxima shows that for a positive number B,

$$\max_k \left(\frac{k}{\delta}\right)^{k(1+\nu)} B^k$$

is reached when

$$\log k = -\frac{1}{1+\nu}(\log B + 1 + \nu) + \log \delta$$

and takes the value

$$\exp[-(1+\nu)]\, B^{-1/(1+\nu)}\, \delta e^{-1}$$

This establishes the desired result when we take $B = |\omega - M/N|\, K$. ∎

4. PROOF OF THEOREM 2.3

The proof of Theorem 2.3 is a perturbation theory for hyperbolic structures.

We recall the following result.

Definition 4.1. We say that a closed set $\Omega \subset M$ is a hyperbolic set for $f\colon M \to M$ if $f\Omega = \Omega$. We can find $C > 0$, $\lambda < 1$, and a splitting $T_x M = E_x^s \oplus E_x^u$ such that

$$\|Df^n(x)v\| \leqslant C\lambda^n \|v\| \qquad \text{if} \quad n \geqslant 0, \quad v \in E_x^s$$

$$\|Df^n(x)v\| \leqslant C\lambda^n \|v\| \qquad \text{if} \quad n \leqslant 0, \quad v \in E_x^u$$

Remark. It follows from the definition that the subspaces E_x^s, E_x^u are uniquely determined and that $Df(x)(E_x^s) = E_{f(x)}^s$, $Df(x)(E_x^u) = E_{f(x)}^u$.

Moreover, the mapping $x \to E_x^s$, $x \to E_x^u$ are continuous.

The following result is stated and proved in ref. 17.

Lemma 4.2. Let Ω be a closed hyperbolic set, and Ω' be an invariant set contained in a sufficiently small neighborhood of Ω. Then, $\Omega \cup \Omega'$ is a hyperbolic set and it is possible to extend the bundles E^s_x, E^u_x to Ω' in such a way that Definition 4.1 is satisfied for some other λ. If the neighborhood of Ω containing Ω' is small enough, then λ can be chosen as close as desired to that on Ω.

For a periodic point x of period N, the remark after Definition 4.1 implies that

$$\|Df^{Ni}(x) \mid_{E^s_x}\| \leqslant C\lambda^{Ni}$$

$$\|Df^{-Ni}(x) \mid_{E^u_x}\| \leqslant C\lambda^{Ni}$$

This implies, by the spectral radius formula, that all eigenvalues of $Df^N(x) \mid_{E^s_x}$ have modulus less than λ^N and that all eigenvalues of $Df^N(x) \mid_{E^u_x}$ have modulus bigger than λ^{-N}.

Therefore, $|R(x)| \geqslant -\lambda^N - 2 - \lambda^{-N}$.

To prove that the Lyapunov exponents of the periodic orbit converge to those of the set, we refer to the proof of a similar statement in the proof of Theorem 2.4. ∎

4.1. Proof of Theorem 2.4

The two claims of Theorem 2.4 are general results about limit properties of Lyapunov exponents. Claim (a), which is much easier, is a statement about Lyapunov exponents of uniquely ergodic measures.

We recall that a mapping defined on a set is called uniquely ergodic if it leaves invariant only one measure.

It is well known (see, e.g., ref. 44, p. 138) that an irrational rotation on the circle leaves invariant the standard Lebesgue measure and no other, so it is a uniquely ergodic system.

Since the motion on an Aubry–Mather Cantor set is semiconjugate to a rotation, that is, we can find a continuous $h: \Gamma \mapsto \mathbf{T}$ such that $h \circ f \mid_\Gamma = h \circ R_\omega$, we see that the only measure defined on Γ, invariant under f, is the pullback under h of the Lebesgue measure on the circle. We will denote such a measure by δ_Γ.

The fact that Aubry–Mather sets are uniquely ergodic justifies speaking of the Lyapunov exponent of the set without specifying explicitly the ergodic invariant measure with respect to which they are considered.

If $o(x_n) \equiv \{x_n, f(x_n),..., f^{N_n - 1}(x_n)\}$ is an orbit of period N_n, the measure that assigns weight $1/N_n$ to each of the points in the orbit is invariant under f. We will denote such a measure by $\delta_{o(x_n)}$.

Given a sequence of orbits $\{o(x_n)\}_{n=0}^{\infty}$ converging to Γ, by the Banach–Alaoglu theorem, we can extract a subsequence $\{o(x_{n_i})\}$ such that the measures $\delta_{o(x_{n_i})}$ converge to a measure δ_∞. Since each of the measures is invariant,

$$f^*\delta_{o(x_{n_i})} = \delta_{o(x_{n_i})}$$

and the pullback is continuous in the weak-* topology, we conclude that the $f^*\delta_\infty = \delta_\infty$. On the other hand, it is easy to see that δ_∞ has support in Γ. By the unique ergodicity discussed before, we conclude that $\delta_\infty = \delta_\Gamma$.

We also recall that the largest Lyapunov exponents of an ergodic measure are upper semicontinuous with respect to the ergodic invariant measures.

This can be easily seen by noticing that the largest Lyapunov exponent is computed by appealing to the *subadditive* ergodic theorem (see, e.g., ref. 36, p. 30). If we denote by $\gamma(f, \mu)$ the Lyapunov exponent of a measure μ ergodic for f, we have

$$\gamma(f, \mu) = \lim \frac{1}{n} \int \log \|Df^n(x)\| \, d\mu(x) = \inf_n \frac{1}{b} \int \log \|Df^n(x)\| \, d\mu(x) \tag{4.1}$$

If we denote

$$\gamma_n(f, \mu) = \frac{1}{n} \int \log \|Df^n(x)\| \, d\mu(x)$$

we have $\gamma_n(f, \mu_i) \geqslant \gamma(f, \mu_i)$. Taking $\limsup_i$ on both sides of the inequality, we obtain $\limsup_i \gamma_n(f, \mu_i) \geqslant \limsup_i \gamma(f, \mu_i)$. Since $\log \|Df^n(x)\|$ is a continuous function and μ_i converges weakly to μ, we have $\gamma_n(f, \mu) \geqslant \limsup_i \gamma(f, \mu_i)$. Taking $\inf_n$ on both sides, we obtain the desired result. This finishes the proof of claim (a) of Theorem 2.4. We emphasize that the proof works word for word for any set on which the motion is uniquely ergodic.

The proof of claim (b) is much more complicated. It will be a trivial consequence of the following theorem, which we state in full generality, since it can be applied in other contexts.

Theorem 4.3. Let $f: M \mapsto M$ be a C^2 diffeomorphism leaving invariant the ergodic measure μ. Assume that, with respect to this measure, f has no zero Lyapunov exponents. Then, for almost every point x_0 in the support of μ, it is possible to find a sequence $\{x_n\}_{n=0}^{\infty}$ of periodic points which converge to x_0. Moreover, the sequence of orbits can be chosen in such a way that the Lyapunov exponents of x_n converge to the Lyapunov exponents of x_0.

Remark. Results similar to Theorem 4.3 appear in ref. 44 (see Theorem 4.1) and ref. 25. They are usually called *ergodic closing lemmas*.

Proof. The proof we present here, like the proofs above, will rely on a shadowing lemma for partially hyperbolic orbits.

The argument will start by proving a constructive version of a shadowing lemma and then we will show that partially hyperbolic systems satisfy the hypothesis. We emphasize that the version of the shadowing lemma we prove does not assume any global hyperbolic properties of the dynamical system, but only hyperbolicity properties of the pseudo-orbit considered. Such statements are useful in other contexts. For example, they are useful when one wants to verify rigorously that near a computed periodic orbit there is a true orbit. In that case, even if one has the approximate orbit quite explicitly, one does not have much control about the global properties of the dynamical system.

We will prove the shadowing lemma by systematically analyzing sequences of orbits. We will adopt the convention of denoting sequences in boldface and their components by the same letter with a subindex.

Definition 4.4. Let M be a manifold and $f: M \mapsto M$ be a diffeomorphism. We say that a sequence $\{x_n\}_{n=-\infty}^{\infty}$ is an ε-pseudo-orbit if $d(x_i, f(x_{i-1})) \leqslant \varepsilon$. This is equivalent to saying that we can find mappings g_i defined in a neighborhood U_i of x_i in such a way that $g_i(x_i) = x_{i-1}$, $\|f - g_i\|_{C^0} \leqslant \varepsilon$.

Definition 4.5. We say that an ε-pseudo-orbit is ε-pseudohyperbolic if we can find a decomposition $T_{x_i} = E^s_{x_i} \oplus E^u_{x_i}$ and mappings g_i defined in neighborhoods U_i of x_i and such that:

(i) $g_i(x_i) = x_{i-1}$.

(ii) $\|f - g_i\|_{C^1} \leqslant \varepsilon$.

(iii) The following hold:

$$\|Dg_{i+n}(x_{i+n})\, Dg_{i+n-1}(x_{i+n-1}) \cdots Dg_i(x_i)\| \leqslant C\lambda^n \|v\| \qquad \text{if} \quad n > 0, \quad v \in E^s_i$$

$$\|Dg^{-1}_{i-n}(x_{i-n})\, Dg^{-1}_{i-n+1}(x_{i-n+1}) \cdots Dg^{-1}_i(x_i)\| \leqslant C\lambda^n \|v\| \qquad \text{if} \quad n > 0, \quad v \in E^u_{i-n}$$

We will refer to ε, C, and λ above as the parameters of hyperbolicity. If $\mathbf{x} \equiv \{x_n\}_{n=-\infty}^{\infty}$ is a sequence, we can pick neighborhoods U_i around

x_i and choose coordinate systems Φ_i: $U_i \mapsto \mathbf{R}^d$ in such a way that the coordinate mappings are uniformly C^∞ and that $\Phi_i(x_i) = 0$. [A geometrically natural way of doing this is using the exponential mapping of Riemannian geometry $\Phi_i(y) = \exp_{x_i}^{-1}(y)$.]

If we define $\tilde{g}_i \equiv \Phi_{i+1} \circ g_i \circ \Phi_i^{-1}$, $\tilde{f}_i \equiv \Phi_{i+1} \circ f \circ \Phi_i^{-1}$, they are mappings mapping a neighborhood of $0 \in \mathbf{R}^d$ to another neighborhood of $0 \in \mathbf{R}^d$. Moreover, $\tilde{g}_i(0) = 0$.

Following ref. 18, we consider the space $\Xi = \{\mathbf{y} \in (\mathbf{R}^d)^N \mid \sup_i |y_i| < \infty\}$. Clearly, Ξ is a Banach space under the norm $\|\mathbf{y}\| \equiv \sup_i |y_i|$. Notice that, for some $\delta > 0$, $\|\mathbf{y}\| \leqslant \delta$ implies that $y_i \in \Phi_i(U_i)$.

On a sufficiently small neighborhood of 0, we can define the operators $\mathscr{T}_f$ by $\mathscr{T}_f(\mathbf{y})_i = f_{i-1}(y_{i-1})$.

Notice that $\mathscr{T}_f(\mathbf{y}) = \mathbf{y}$ if and only if $\{\Phi_i^{-1}(y_i)\}_{i=-\infty}^{\infty}$ is an orbit for f and that $\mathbf{y}$ is an ε-pseudo-orbit if and only if $K^{-1}\varepsilon\ \|\mathscr{T}(\mathbf{y}) - \mathbf{y}\| \leqslant K\varepsilon$, where K is a bound on the derivatives of Φ_i and Φ_i^{-1}.

Proposition 4.6. If f is uniformly differentiable, then $\mathscr{T}$ is differentiable in a neighborhood of the origin and we have

$$[D\mathscr{T}(\mathbf{x})\mathbf{a}]_n = Df_{n-1}(x_{n-1})\, a_{n-1}$$

If f is uniformly C^2, then $\mathscr{T}$ is C^2 and we can bound $\|D^2\mathscr{T}(\mathbf{x})\|$ uniformly in a neighborhood of the origin.

Proof. To establish the first claim, we just have to bound

$$\|\mathscr{T}(\mathbf{x} + \mathbf{a}) - \mathscr{T}(\mathbf{x}) - D\mathscr{T}(\mathbf{x})\mathbf{a}\| \tag{4.2}$$

and show that it converges to zero with $\|\mathbf{a}\|$ faster than $\|\mathbf{a}\|$.

We recall that f is uniformly differentiable if one can find an increasing function η: $\mathbf{R}^+ \mapsto \mathbf{R}^+$ with $\eta(0) = 0$ and $\lim_{t\to 0} \eta(t)/t = 0$ such that $|f(x+a) - f(x) - Df(x)a| \leqslant \eta(|a|)$. If the function f that we used to construct $\mathscr{T}_f$ is uniformly differentiable—this is automatic if the manifold is compact or if f has uniformly continuous first derivatives—using the fact that the mappings Φ_i and their inverses have uniformly continuous derivatives, we conclude that for some η: $\mathbf{R}^+ \to \mathbf{R}^+$ increasing and $\eta(0) = 0$, we have

$$|f_i(x+a) - f_i(x) - Df_i(x)\, a_{i-1}| \leqslant \eta(|a_{i-1}|)$$

Using the definition of the norm, the quantity (4.2) that we have to estimate is just

$$\sup_n \|f_{n-1}(x_{n-1} + a_{n-1}) - f_{n-1}(x_{n-1}) - Df_{n-1}(x_{n-1})\, a_{n-1}\|$$

Using the uniform differentiability, we obtain that this can be bounded by

$$\sup_n \eta(|a_{n-1}|) \leqslant \eta(\sup_n |a_{n-1}|) = \eta(\|\mathbf{a}\|)$$

which is what we wanted to establish.

The argument for the second derivative is very similar and we leave the details to the reader. ∎

The following lemma provides us with a characterization of the hyperbolicity of orbits by properties of the derivative of the operator $\mathscr{T}_f$ at $\mathbf{x}$.

Their usefulness comes from the fact that they allow us to prove properties that are true for whole orbits—uniformly on the time—by doing soft analysis on the operator $\mathscr{T}_f$. They are nonautonomous versions of the characterizations in ref. 22 and the proofs are actually quite similar. We point out that property (i) will not be used in this paper, but we included it because it fits nicely in the circle of ideas discussed here. Since the spectral theory on Banach spaces is much more natural on complex spaces, we will consider χ the natural complexification of Ξ. We leave to the reader the elementary task of checking that, when the problem considered has real data, the results are real.

Lemma 4.7. Let $\mathbf{x}$ be a fixed point of $\mathscr{T}_f$ as before. Then:

(i) The spectrum of $D\mathscr{T}_f(\mathbf{x})$ is invariant under rotations, i.e.,

$$z \in \operatorname{spec}(D\mathscr{T}_f(\mathbf{x})) \Rightarrow \forall\theta \in \mathbf{R}, \qquad e^{i\theta} \in \operatorname{spec}(D\mathscr{T}_f(\mathbf{x}))$$

(ii) Assume that for $0 < \mu_- < \mu_+$

$$\operatorname{spec}(D\mathscr{T}_f(\mathbf{x})) \cap \{z \in \mathbf{C} \mid \mu_- \leqslant |z| \leqslant \mu_+\} = \varnothing$$

Then, we can find a sequence of subspaces $E_i^{[>]}$, $E_i^{[<]}$ in such a way that

(a) $\mathbf{R}^d = E_i^{[>]} \oplus E_i^{[<]}$
$\operatorname{angle}(E_i^{[>]}, E_i^{[<]}) \geqslant \alpha > 0$

(b) $\|Df_{i+m}(x_{i+m}),..., Df_i(x_i) \mid_{E_i^{[<]}}\| \leqslant C\mu_+^m$
$\|Df_{i-m}^{-1}(x_{i-m}),..., Df_i^{-1}(x_i) \mid_{E_i^{[>]}}\| \leqslant C\mu_-^{-m}$

(c) $Df_i(x_i)\, E_i^{[<]} = E_{i+1}^{[<]}$
$Df_i(x_i)\, E_i^{[>]} = E_{i+1}^{[>]}$

(iii) Conversely, if we can find $E_i^{[>]}$, $E_i^{[<]}$, C, μ_+, and μ_- satisfying (iia)–(iic) as before, then

$$\operatorname{spec}(D\mathscr{T}_f(\mathbf{x})) \cap \{z \in \mathbf{C} \mid \mu_- \leqslant |z| \leqslant \mu_+\} = \varnothing$$

Proof. To prove (i), we recall that a number $z \in \mathbf{C}$ is in the spectrum of $(D\mathcal{T}_f(\mathbf{x}))$ if and only if there exists a sequence $\{\mathbf{v}_n\}_{n=0}^{\infty}$, $\|\mathbf{v}_n\| = 1$, $\lim_n \|D\mathcal{T}_f(\mathbf{x})\, \mathbf{v}_n - z\mathbf{v}_n\| = 0$.

So, to prove the theorem, it suffices to show that if we have $\mathbf{v} \in \chi$ such that $\|D\mathcal{T}_f(\mathbf{x})\mathbf{v} - z\mathbf{v}\| \leqslant \varepsilon$, we can find $\mathbf{w}$ with

$$\|D\mathcal{T}_f(\mathbf{x})\mathbf{w} - ze^{i\theta}\mathbf{w}\| \leqslant \varepsilon$$

Expressed in components, the hypothesis means that

$$|Df_i(x_i)\, v_i - zv_{i+1}| \leqslant \varepsilon, \qquad \sup |v_i| = 1$$

If we set $w_n = e^{-ni\theta} v_n$ we have

$$\sup_n |w_n| = \sup_n |v_n| = 1$$

$$\sup |Df_n(x_n)\, w_n - ze^{i\theta} w_{n+1}| = \sup_n |e^{in\theta}[Df_n(x_n)\, v_n - zv_{n+1}]| = \varepsilon$$

This finishes the proof of (i).

To prove (ii), we observe that, by the spectral theorem on Banach spaces, we can find a $\chi = \chi^{[<]} \oplus \chi^{[>]}$ invariant under $D\mathcal{T}_f(\mathbf{x})$ in such a way that

$$\begin{aligned} \operatorname{spec}(D\mathcal{T}_f(\mathbf{x})\,|_{\chi^{[<]}}) &\subset \{z \in \mathbf{C} \mid |z| \leqslant \mu_-\} \\ \operatorname{spec}(D\mathcal{T}_f(\mathbf{x})^{-1}\,|_{\chi^{[>]}}) &\subset \{z \in \mathbf{C} \mid |z| \leqslant \mu_+^{-1}\} \end{aligned} \tag{4.3}$$

We also recall that another corollary of the spectral theorem is that

$$\begin{aligned} \mathbf{v} \in \chi^{[<]} &\Leftrightarrow \lim_{n \to \infty} \|[D\mathcal{T}_f(\mathbf{x})]^n\, \mathbf{v}\|^{1/n} \leqslant \mu_- \\ &\Leftrightarrow \forall n\ \|[D\mathcal{T}_f(\mathbf{x})]^n\, \mathbf{v}\| \leqslant C\mu_-^n\ \|\mathbf{v}\| \\ \mathbf{v} \in \chi^{[>]} &\Leftrightarrow \lim_{n \to \infty} \|[D\mathcal{T}_f(\mathbf{x})]^{-n}\, \mathbf{v}\|^{1/n} \leqslant \mu_+^{-1} \\ &\Leftrightarrow \forall n \in \mathbf{N} \|[D\mathcal{T}_f(\mathbf{x})]^{-n}\, \mathbf{v}\| \leqslant C\mu_-\ \|\mathbf{v}\| \end{aligned} \tag{4.4}$$

where, in the second characterization, we understand implicitly that $\mathbf{v} \in \operatorname{Dom}(D\mathcal{T}_f(\mathbf{x}))^{-n}$.

Both characterizations say roughly that $\mathbf{v} \in \chi^{[<]}$ if iterates of $D\mathcal{T}_f$ decrease faster than an exponential rate μ_-. Notice that the two precise definitions of "exponential of rate μ_-" are not equivalent in general, the first one being weaker even for one vector (notice that C in the second characterization is independent of the vector). The proof of (4.4) uses essentially the spectral properties of $D\mathcal{T}_f(\mathbf{x})$ we assumed.

We want to prove that $\chi^{[<]}$ is of the form

$$\chi^{[<]} = \{\mathbf{v} \in \chi \mid v_i \in E_i^{[<]}\}$$

that is, whether $\mathbf{v}$ belongs or not to the space $\chi^{[<]}$ can be ascertained by testing successively the components. This will be true basically because to compute one coordinate of $D\mathcal{T}_f(\mathbf{x})\mathbf{v}$, we only need to know one component of $\mathbf{v}$.

Proposition 4.8. The vector $\mathbf{v} \in \chi$ belongs to $\chi^{[<]}$ if and only if, for every i, the vector $\mathbf{v}^i \equiv (..., 0,..., v_i, 0,..., 0,...) \in \chi^{[<]}$.

Proof. We will use (4.4) to prove both implications. We have

$$\|[D\mathcal{T}_f(\mathbf{x})]^n \mathbf{v}^i\| = |Df_{i+n}(x_{i+n}) \cdots Df_i(x_i)\, v_i|$$

$$|([D\mathcal{T}_f(\mathbf{x})]^n \mathbf{v})_{i+n}| \leqslant \|[D\mathcal{T}_f(\mathbf{x})]^n \mathbf{v}\| \leqslant C\mu_-^n \,\|\mathbf{v}\|$$

To prove the if part,

$$\|[D\mathcal{T}_f(\mathbf{x})]^n \mathbf{x}\| = \sup_i |([D\mathcal{T}_f(\mathbf{x})]^n \mathbf{v})_{i+n}|$$

$$= \sup_i \|[D\mathcal{T}_f(\mathbf{x})]^n \mathbf{v}^i\|$$

$$\leqslant \sup_i C\mu_-^n \,\|v_i\| = C\mu_-^n \,\|v\|$$

Hence, we can define

$$E_i^{[<]} = \{v \in \mathbf{R}^d \mid (..., 0,..., 0, v, 0,..., 0,...) \in \chi^{[<]}\}$$

Since

$$\mathcal{T}_f(\mathbf{x})(..., 0,..., 0, v, 0,..., 0,...) = (..., 0,..., 0, 0, Df_i(x_i)v,..., 0,...)$$

we see that $v \in E_i^{[<]} \Leftrightarrow Df_i(x_i) \in E_{i+1}^{[<]}$. So that (iic) of Lemma 4.7 is established.

To prove (iib), we observe that

$$\begin{aligned}
&\|[\mathcal{T}_f(\mathbf{x})]^n (..., 0,..., 0, v, 0,..., 0,..., 0)\| \\
&\quad = \|(..., 0,..., 0, 0, Df_{i+n}(x_{i+n})\, Df_{i+n-1}(x_{i+n-1}) \cdots Df_i(x_i)v, 0,...)\| \\
&\quad = |Df_{i+n}(x_{i+n} \cdots Df_i(x_i)v| \\
&\quad \leqslant C\mu_+^n \,\|(..., 0,..., 0, v, 0,..., 0)\| \\
&\quad = C\mu_+^n \,|v|
\end{aligned}$$

Analogous arguments and definitions work for $\chi^{[>]}$.

To prove (a), we observe that clearly $E_i^{[<]}$, $E_i^{[>]}$ are closed linear spaces. Moreover, their intersection is $\{0\}$, since the intersection of $\chi^{[>]}$, $\chi^{[<]}$ is the null vector. If there was a vector $v \in \mathbf{R}^d$, $v \notin E_i^{[<]} \oplus E_i^{[>]}$, we see that the vector $(..., 0,..., 0, v, 0,..., 0) \in \chi^{[<]} \oplus \chi^{[>]}$.

To prove (a), we observe that if we could find a vector $w \in \mathbf{R}^d$, then $w \notin E_i^{[<]} \oplus E_i^{[<]}$.

To show that the angle between spaces is bounded from below, we recall that as a consequence of the spectral theorem, $\Pi^{[<]}$ and $\Pi^{[>]}$, the spectral projections onto $\chi^{[<]}$ and $\chi^{[>]}$, are bounded.

By Proposition 4.8,

$$\Pi^{[<]}(..., v_i, v_{i+1},...) = (..., \pi_i^{[<]} v_i, \pi_{i+1}^{[>]} v_{i+1},...)$$

where $\pi_i^{[<]}$ and $\pi_i^{[>]}$ are the projections associated with the decomposition $\mathbf{R}^d = E_i^{[>]} \oplus E_i^{[>]}$. Since the spectral projections $\pi^{[<]}$ and $\pi^{[>]}$ are bounded, we have

$$\begin{aligned} |\pi_i^{[<]} v| &= \|\Pi^{[<]}(..., 0,..., 0, v, 0,..., 0)\| \\ &\leqslant \|\Pi^{[<]}\| \cdot \|(..., 0,..., 0, v, 0,..., 0,...)\| \\ &= \|\Pi^{[<]} 0\| \cdot |v| \end{aligned}$$

A similar argument shows $|\pi_i^{[>]} v| \leqslant \|\Pi^{[>]}\| \cdot |v|$. Hence, $\|\pi_i^{[<]}\|$ and $\|\pi_i^{[>]}\|$ are bounded independently of i. This is equivalent to saying that the angle between $E_i^{[<]}$ and $E_i^{[>]}$ is uniformly bounded from below.

This finishes the proof of (ii) of Proposition 4.8.

To prove (iii), it suffices to show that the equation

$$D\mathscr{T}_f(\mathbf{x})\mathbf{v} - z\mathbf{v} = \mathbf{w} \tag{4.5}$$

can be solved in $\mathbf{v}$ for any $\mathbf{w}$, $\mu_- < |z| < \mu_+$ and that $\|\mathbf{v}\| \leqslant C \|\mathbf{w}\|$.

Taking components, (4.5) is equivalent to

$$Df_i(x_i)\, v_i - z v_{i+1} = w_{i+1} \tag{4.6}$$

If $\pi_i^{[<]}$ and $\pi_i^{[>]}$ are the projections associated with the splitting $\mathbf{R}^d = E_i^{[<]} + E_i^{[>]}$, (iia) implies

$$\begin{aligned} Df_i(x_i)\, \pi_i^{[<]} &= \pi_{i+1}^{[<]} Df_i(x_i) \\ Df_i(x_i)\, \pi_i^{[>]} &= \pi_{i+1}^{[>]} Df_i(x_i) \end{aligned}$$

Hence, decomposing into the components along $E_{i+1}^{[<]}$ and $E_{i+1}^{[>]}$, we obtain that (4.6) is equivalent to

$$\begin{aligned} Df_i(x_i)\, \pi_i^{[<]} v_i - z\pi_{i+1}^{[<]} v_{i+1} &= \pi_{i+1}^{[<]} w_{i+1} \\ Df_i(x_i)\, \pi_i^{[>]} v_i - z\pi_{i+1}^{[>]} v_{i+1} &= \pi_{i+1}^{[>]} w_{i+1} \end{aligned} \tag{4.7}$$

We claim that these two equations can be solved by setting

$$\pi_{i+1}^{[<]} v_{i+1} = -\frac{1}{z} \pi_{i+1}^{[<]} w_{i+1} \sum_{j=0}^{\infty} \frac{-1}{z^{j+1}} [Df_i(x_i) \cdots Df_{i-j}(x_{i-j})] \pi_{i=j}^{[<]} w_{i-j}$$

$$\pi_{i+1}^{[>]} v_{i+1} = \sum_{j=0}^{\infty} z^j \{ [Df_i(x_i)]^{-1} \cdots Df_{i+j}(x_{i+j})^{-1} \} \pi_{i+1+j}^{[>]} w_{i+1+j} \tag{4.8}$$

In effect, we see that, using (iib),

$$\left| \frac{1}{z^{j+1}} [Df_i(x_i) \cdots Df_{i-j}] \pi_{i-j}^{[<]} w_{i-j} \right| \leqslant \frac{1}{|z|^{j+1}} C\mu_-^j \, |\pi_{i-j}^{[<]} w_{i-j}|$$
$$\leqslant \frac{C}{z} \left| \frac{\mu_-}{z} \right|^j \|\pi^{[<]}\| \, \|\mathbf{w}\|$$

and analogously for the other equation

$$|z^j \{ [Df_i(x_i)]^{-1} \cdots [Df_{i+j}(x_{i+j})]^{-1} \} \pi_{i+j+1}^{[>]} w_{i+j+1}|$$
$$\leqslant \left| \frac{z}{\mu_+} \right|^j C \, \|\pi^{[>]}\| \cdot \|\mathbf{w}\|$$

Hence, the two series in (4.8) converge uniformly and, by rearranging terms, it is easy to verify that they indeed are solutions.

Moreover, since the right-hand sides of (4.8) have norms that can be bounded by $C\|\mathbf{w}\|$ independently of i, we see that the sum will also be bounded in the same form, hence z is in the resolvent. This finishes the proof of Lemma 4.7. ▌

Lemma 4.9. Let $\mathscr{T}$ be a C^2 function on a Banach space X, $\mathbf{x} \in X$. Let M be a linear operator on X. Assume

$$\|\mathscr{T}(\mathbf{x}) - \mathbf{x}\| \leqslant \varepsilon$$
$$\|D\mathscr{T}(\mathbf{x}) - M\| \leqslant A$$
$$\|(M - I)^{-1}\| \leqslant B$$
$$\|D^2\mathscr{T}(\mathbf{y})\| \leqslant C \qquad \text{if} \quad \|\mathbf{x} - \mathbf{y}\| \leqslant \rho$$

Then, if

$$K \equiv AB + B \leqslant \rho < 1$$
$$\varepsilon + K\rho < \rho$$

there is one fixed point x^* of $\mathscr{T}$ in the set $\{y \mid \|y-x\| \leqslant \rho\}$. Moreover, $\|x-x^*\| \leqslant \varepsilon/1-K$.

Proof. Consider

$$\Phi(\mathbf{y}) = -(M-I)^{-1}[\mathscr{T}(\mathbf{y})-\mathbf{y}]+\mathbf{y}$$

A simple calculation shows that a fixed point of Φ is also a fixed point of $\mathscr{T}$.

Moreover, Φ is twice differentiable and we have

$$\begin{aligned} D\Phi(\mathbf{y}) &= -(M-I)^{-1}[D\mathscr{T}(\mathbf{y})-I]+I \\ &= -(M-I)^{-1}\{M-I+[D\mathscr{T}(\mathbf{x})-M]+[D\mathscr{T}(\mathbf{y})-D\mathscr{T}(\mathbf{x})]\}+I \\ &= -(M-I)^{-1}[D\mathscr{T}(\mathbf{x})-M]-(M-1)^{-1}[D\mathscr{T}(\mathbf{y})-D\mathscr{T}(\mathbf{x})] \end{aligned}$$

If we bound the first term using the second part of (i) and the second term by the mean value theorem, we obtain

$$\|D\Phi(\mathbf{y})\| \leqslant AB+B \leqslant \rho$$

Hence, Φ is a contraction in the ball around $\mathbf{x}$ of radius ρ. The second inequality implies that this ball is mapped into itself. From that, we can apply the elementary argument for the contraction mapping principle. ∎

We also recall some facts from the theory of nonuniformly hyperbolic systems. Our ultimate goal is to show that, for systems with positive Lyapunov exponents, we can construct pseudo-orbits with approximate inverses such that Lemma 4.9 applies.

Theorem 4.10. Let μ be Borel measure invariant under f. Then, for μ-almost all x, we have:

(i) $T_x M = \oplus_i E_i(x)$

with

(ii) if $v \in E_i$

$$\lim_{n\to\infty} \frac{1}{n} \log |Df^n(x)v| = \lim_{n\to\infty} \frac{1}{n} \log |Df^{-n}xv| = \lambda_i(x)$$

(iii) $\Delta_{ij}(x)$, the angle between $E_i(x)$ and $E_i(j)$, $i \neq j$, satisfies

$$\Delta_{ij}(f^n(x)) \geqslant e^{-|n|\varepsilon}(x)\, C_\varepsilon(x)$$

for some measurable function $C_\varepsilon(x) \neq 0$.

Proof. There are several proofs of this theorem in the literature. The original one is in ref. 29. A more modern one can be found in ref. 36. There, (iii) is proved explicitly as Corollary 3.3.

Notice that $\lambda_i(f(x)) = \lambda_i(x)$. Hence, if μ is ergodic under f, λ_i is a constant. Also, $Df(x)\, E_i(x) = E_i(f(x))$. Once we have these results, it is easy to prove the following:

Corollary 4.11. Let μ be a Borel probability measure, invariant under f and ergodic. Given $\varepsilon > 0$, we can find $l > 0$ and a set $\Lambda_{\varepsilon,l}$ such that:

(i) $\mu(\Lambda_{\varepsilon,l}) \geqslant 1 - \varepsilon$.

(ii) $\Lambda_{\varepsilon,l}$ is closed.

For all $x \in \Lambda_{\varepsilon,l}$ we have:

(i) If $n, m \in \mathbf{Z}$ and $v \in E_i(f^m x)$

$$l^{-1} e^{\lambda_i n} e^{\varepsilon\, |n| + |m|} \leqslant |Df^n(f^m x)\, v| \leqslant l e^{\lambda_i n} e^{E(|n| + |m|)}$$

(ii) The spaces $E_i(x)$ depend continuously on x when $x \in \Lambda_{\varepsilon,l}$.

(iii) $\Delta_{ij}(f^n(x)) \geqslant l^{-1} e^{\varepsilon\, |n|}$.

(iv) The sets $\Lambda_{\varepsilon,l}$ can be chosen in such a way that $\Lambda_{\varepsilon,l} \subset \Lambda_{\varepsilon,l'}$ if $l' > l$ and $\Lambda_\varepsilon = \bigcup_l \Lambda_{\varepsilon,l}$ has full measure.

Now we can go back to the proof of Theorem 2.4.

If $x \in \sup(\mu_\Gamma)$, we observe that we can find ε, l such that $x \in \operatorname{supp}(\mu \mid_{\Lambda_{\varepsilon,l}})$. Hence, for every $\delta > 0$

$$\mu(\Lambda_{\varepsilon,l} \cap B_\delta(x)) > 0$$

By the Poincaré recurrence theorem, we can find $x_0, \ldots, x_N$ such that $f(x_i) = x_{i+1}$, $x_0 \in \Lambda_\varepsilon \cap B_\delta(x)$, and $x_N \in \Lambda_\varepsilon \cap B_\delta(x)$.

If δ is small enough, we can take coordinate patches U_i around $x_0, \ldots, x_{N-1}$ as before in such a way that x_N will be on the coordinate patch of x_0.

Denote by $\tilde{x}_N$ the coordinate representation of x_N on the patch U_0.

Denote also $\tilde{E}^i_j(x) = D\Phi_i(\Phi_i^{-1} x)\, E_j(\Phi_i^{-1} x)$ [that is, $\tilde{E}^i_j(x)$ is the coordinate representation of the spaces corresponding to the jth Lyapunov exponent].

Notice that $Df_i(x)\, \tilde{E}^i_j(x) = \tilde{E}^{i+1}_j(f(x))$.

If the spaces $E_j(x)$ depend continuously on $x \in \Lambda_{\varepsilon,l}$, we see that $\tilde{E}^i_j$ will depend continuously also and the modulus of continuity can be estimated from the modulus of continuity of the coordinate changes.

In particular, we can find operators $\pi_i(x)$ such that

$$\pi_i(x)\,\tilde{E}^i_j(x) = \tilde{E}^i_j(0)$$
$$\|\pi_i(x) - Id\| \leqslant \omega(|x|)$$

with $\omega(t)$ decreasing $\omega(t) \to 0$ as $t \to 0$.

We claim that, for δ small enough, the pseudo-orbit given by

$$\mathbf{x} = \begin{cases} 0 & \text{if } \; i \neq kN \\ \tilde{x}_N & \text{if } \; i = kM \end{cases}$$

and the operator M defined by $(M\eta)_i = M_{i-1}\eta_{i-1}$ with $M_{i-1} = Df_{i-1}(0)$ if $i \neq kN$; $M_{i-1}\pi_i(\tilde{x}_N)\,Df_{i-1}(0)$ if $i = kN$.

Notice that we do not have upper bounds for M, but that Poincaré's recurrence theorem implies that there is a sequence of N's going to infinity. Hence, we will have to prove the estimates of Lemma 4.9 for N sufficiently large.

We will assume, without loss of generality, that N is large enough that

$$l\exp[(\mu^+ - \varepsilon)N] < 1$$
$$l\exp[(-\mu^- - 2\varepsilon)N] < 1$$

where μ^+ is the smallest positive Lyapunov exponent and μ^- is the negative Lyapunov exponent of smallest absolute value (heuristically, we are waiting long enough that the nonuniform hyperbolicity has had time to start acting).

Clearly, $\|\mathscr{T}_f(\mathbf{x}) - \mathbf{x}\| = |\tilde{x}_N|$, which can be estimated by $K\delta$, where, as before, K is a constant that only depends on the supremum of the derivatives of the coordinate mappings. We have

$$\begin{aligned}\|D\mathscr{T}_f - M\| &= \|Df_{N-1}(0) - \pi_N(\tilde{x}_N)\,Df_{N-1}(0)\| \\ &= \|Df_{N-1}(0)\|\;\omega(|\tilde{x}_N|)\end{aligned}$$

which also tends to zero with δ.

To estimate $(M - I)^{-1}$ we study the equation $(M - I)\eta = \mathbf{w}$ in a way quite similar to the proof of part (b) of Lemma 4.7. Notice that, by construction, M preserves the decomposition $\chi = \oplus_i \chi_i$ with $\chi_i = \{\mathbf{v} \mid v_k \in \tilde{E}^k_i\}$.

Proceeding as in the proof of Lemma 4.7(b), we find that the equation

$$M\eta - \eta = \mathbf{w}$$

admits the solutions [analogous to (4.8)] given by

$$\begin{aligned}\pi^{[<]}_{i+1}\eta_{i+1} &= -\pi^{[<]}_{i+1} w_{i+1} + \sum_{j=0} M_i \cdots M_{i-j} \pi^{[<]}_{i-j} w_{i-j} \\ \pi^{[>]}_{i+1}\eta_{i+1} &= \sum_{j=0}^{\infty} M_i^{-1} \cdots M_{i+j}^{-1} \pi^{[>]}_{i+1+j} w_{i+1+j}\end{aligned} \tag{4.9}$$

To bound the products in the sums in (4.8), we use the inequalities in the definition of the set $\Lambda_{\varepsilon,l}$. We see that

$$\begin{aligned}M_{kN} M_{kN+1} &\cdots M_{(k+1)N} \\ &= Df_{kN}(0) \cdots Df_{k+N}(0) \\ &\quad + Df_{kN}(0) \cdots Df_{(k+1)N-1}(0)\, \pi_{N-1}(\tilde{x}_N)\, Df_{(k+1)N}(0)\end{aligned}$$

Hence the norm is bounded by

$$l \exp[(\mu^- + \varepsilon)(N-1)]\, l \exp(\varepsilon N)\, \omega(\delta) + l \exp[(\mu^- + \varepsilon) N]$$

We see that, by making N sufficiently large and δ sufficiently small, we can ensure that this is bounded away from 1. Hence, the first series in (4.9) converges uniformly and we can bound the result by a constant times $\|\mathbf{w}\|$.

An analogous argument works for the second sum.

Applying Lemma 4.9, we conclude that there is a fixed point of $\mathscr{T}_f$ close to zero that is an orbit that shadows the ε pseudo-orbit $\mathbf{x}$.

We claim that this orbit has to be periodic of period N. To prove that, we recall that the fixed point of $\mathscr{T}_f$ was obtained by iterating the operator $\Phi = -(M-I)^{-1}(\mathscr{T}_f - I) + I$. Notice that $\mathscr{T}$ maps sequences of period N into sequences of period N and, using the formulas for the inverse of $M-I$, so does $(M-I)^{-1}$. Hence Φ maps periodic sequences into periodic sequences and, since the starting sequence is periodic with period N, so should be the fixed point.

To prove the claim about the Lyapunov exponents of the orbit, we use the characterization of Lyapunov exponents in terms of the spectrum of $D\mathscr{T}_f$.

Notice that, by making N sufficiently large and δ sufficiently small, we can guarantee that the fixed point of $\mathscr{T}_f$ would be as close as desired to zero.

Since

$$\|D\mathscr{T}_f(\mathbf{x}) - D\mathscr{T}_f(0)\| \leqslant \|\mathbf{x}\| \cdot \|D^2\mathscr{T}_f\| \leqslant \|\mathbf{x}\| \cdot \|D^2 f\|$$

(notice that this bound does not depend on N), we have that the spectrum of $D\mathscr{T}_f(\mathbf{x})$ would be arbitrarily close—in the sense of sets—to the spectrum of $D\mathscr{T}_f(0)$. By making δ sufficiently small, we can approximate $\operatorname{spec}(D\mathscr{T}_f(0))$ by $\operatorname{spec}(M)$.

The argument to show that $\operatorname{spec}(M)$ is close to being circles around the Lyapunov exponents is very similar to the arguments we have already used. It suffices to show that if z is away from the Lyapunov exponents, we can solve the equation

$$M\mathbf{v} - z\mathbf{v} = \mathbf{w} \tag{4.10}$$

If $\pi_i^{[<|z|]}$ and $\pi^{[>|z|]}$ denote the projection onto $\bigoplus_{e^{\lambda_i}<|z|} E_i$ and $\bigoplus_{e^{\lambda_i}>|z|} E_i$, respectively, by using the analogue of (4.8) and the estimates analogous to those used to bound (4.9), we can show that, provided that N is sufficiently large and δ sufficiently small, (4.10) has a solution if $|z|$ is not a Lyapunov exponent.

This finishes the proof of Theorem 2.4. ∎

5. DISCUSSION

The above results justify Greene's criterion for the families for which there is a sharp transition between the parameter values for which the Aubry–Mather set is an invariant circle and those parameter values for which it is a Cantor set with nonzero Lyapunov exponent.

Unfortunately, from our results it is impossible to distinguish a Cantor set with a zero Lyapunov exponent from a nonsmooth circle. Our results do not exclude either that, on an open interval of parameters, the neutral cases occur or the hyperbolic cases coexist with smooth tori in all the scales.

For the standard map and the golden mean rotation number there is a very convincing renormalization group picture[26,27] which suggests that at precisely one value of the parameter, the Aubry–Mather set of golden mean rotation number is a nonsmooth invariant circle and that, for smaller values of the parameter, the Aubry–Mather set is a smooth circle, whereas for values bigger that the critical value, it is a hyperbolic set.

This picture, due to McKay, is based on a detalied study of the dynamics of a renormalization operator acting on the space of maps.

The operator has an attractive fixed point—called trivial since it can be computed explicitly—and a nontrivial one which McKay computed quite convincingly. This nontrivial fixed point has a stable manifold of codimension 1 and a one-dimensional unstable manifold, one of whose sides ends on the trivial fixed point.

If a map under iteration of the renormalization group converges to the trivial fixed point, the Aubry–Mather Cantor set of golden mean rotation is a smooth circle. (This part of the picture and some generalizations have been justified rigorously in ref. 10.) If a map under repeated renormalization converges to the nontrivial fixed point, the Aubry–Mather set of golden mean rotation is a not very smooth invariant circle. If it approaches the unstable manifold on the side opposite to the trivial fixed point, the Aubry–Mather Cantor set of golden mean rotation is a hyperbolic Cantor set.

The remarkably simple behavior of the standard map family can be justified with the help of this picture by realizing that the curve described by the standard map family crosses transversally the stable manifold of the nontrivial fixed point and is very close to the unstable manifold of the renormalization map. Hence, by the λ-lemma, successive iterations of the renormalization group will make it converge to the unstable manifold.

Notice also that for families sufficiently close to the standard map the λ-lemma also applies and they will also converge to the unstable manifold. If we keep in mind that the effect of the renormalization group on a map is to make the space scale smaller and the time scale longer, the high iterations of the renormalization group capture phenomena that happen on small scales and for long times. The convergence onto the unstable manifold has the consequence that, for all families sufficiently close that the λ-lemma applies, the long-term behavior at small scales is the same.

Unfortunately, this very satisfactory picture has only a local nature. Refs. 14, 40, and 41 and our forthcoming paper[7] present evidence that the dynamics of the renormalization group operator has more complicated features than just a saddle point. These effects can be observed in standard-like families in which we substitute for the sine in the standard family a trigonometric polynomial of two coefficients.

In that case, it seems quite possible that there is not a sharp transition between the smooth behavior and the hyperbolic one. For those systems, Greene's criterion seems to work less efficiently than in the case of the standard map. Nevertheless, the discussion of our proofs suggests that Greene's criterion can be used very effectively and very safely as a negative criterion for the nonexistence of smooth invariant circles. When the residues are reasonably big for a periodic orbit of high period, we can be quite confident that, for this parameter value, there is no smooth invariant circle. Unfortunately, when the renormalization group picture does not hold, the values computed with different orbits do not stack up as predictably as in the case of the standard mapping and it does not seem possible to extrapolate.[7,14,40,41]

ACKNOWLEDGMENTS

The work of R.d.l.L. has been partially supported by National Science Foundation grants. He also would like to acknowledge stimulating discussions wit R. MacKay that took place while both were enjoying the hospitality of S. Wiggins.

REFERENCES

1. V. I. Arnold, *Geometric Methods in the Theory of Ordinary Differential Equations* (Springer-Verlag, New York, 1983).
2. A. Berretti and L. Chierchia, On the complex analytic structure of the golden invariant curve for the standard map, *Nonlinearity* **3**:39–44 (1990).
3. A. Berretti, A. Celletti, L. Chierchia, and C. Falcolini, Natural boundaries for area preserving twist maps, *J. Stat. Phys.*, to appear.
4. M. Brin and Y. Pesin, Partially hyperbolic dynamical systems, *Math. USSR Izv.* **8**:177–218 (1974).
5. A. Celletti and L. Chierchia, Invariant curves for area-preserving twist maps far from integrable, *J. Stat. Phys.*, to appear.
6. A. Celletti and L. Chierchia, Construction of analytic KAM surfaces and effective stability bounds, *Commun. Math. Phys.* **118**:119 (1988).
7. C. Falcolini and R. de la Llave, Numerical calculation of domains of analyticity for perturbation theories in the presence of small divisors, *J. Stat. Phys.* **67**:645–666 (1992).
8. A. Fathi, Une interprétation plus topologique de la démonstration du théorème du Birkhoff, Appendix to M. R. Herman, Sur les courbes invariantes par les difféomorphismes de l'anneau, *Astérisque* **103–104** (1983).
9. J. Greene, A method for determining a stochastic transition, *J. Math. Phys.* **20**:1183–1201 (1979).
10. N. Haydn, On invariant curves under renormalisation, *Nonlinearity* **3**:887–912 (1990).
11. R. Herman, Sur les courbes invariantes par les difféomorphismes de l'anneau, *Astérisque* **103–104** (1983).
12. R. Herman, Sur les courbes invariantes par les difféomorphismes de l'anneau, vol. 2, *Astérisque* **144** (1986).
13. I. Jungreis, A method for proving that monotone twist maps have no invariant circles, *Ergodic Theory Dynam. Syst.* **11**:79–84 (1991).
14. J. A. Ketoja and R. S. McKay, Fractal boundary for the existence of invariant circles for area preserving maps: Observations and a renormalisation explanation, *Physica* **35D**:318–334 (1989).
15. K. M. Khanin and Ya. G. Sinai, The renormalization group and Kolmogorov–Arnold–Moser theory, in *Nonlinear Phenomena in Plasma Physics and Hydrodynamics*, R. Z. Sagdeev, ed. (Mir, Moscow, 1986).
16. R. de la Llave and D. Rana, Accurate strategies for K.A.M. bounds and their implementation, in *Computer Aided Proofs in Analysis*, K. Meyer and D. Schmidt, eds. (Springer-Verlag, New York, 1991).
17. R. de la Llave and C. E. Wayne, Whiskered and low dimensional tori in nearly integrable Hamiltonian systems, preprint.
18. O. Lanford, Introduction to hyperbolic sets, in *Regular and Chaotic Motion in Dynamic Systems*, G. Velo and A. S. Wightman, eds. (Plenum, New York, 1985).
19. R. MacKay and I. C. Percival, Converse K.A.M.: Theory and practice, *Commun. Math. Phys.* **98**:469–512 (1985).

20. J. Mather, Non existence of invariant circles, *Ergodic Theory Dynam. Syst.* **4**:301–309 (1984).
21. J. Mather, A criterion for non existence of invariant circles, *Pub. Math. IHES* **63**:153–204 (1987).
22. J. Mather, Characterization of Anosov diffeomorphisms, *Indag. Mat.* **30**:479–483 (1968).
23. J. Mather, Destruction of invariant circles, *Ergodic Theory Dynam. Syst.* **8**:199–214 (1988).
24. J. Mather, More Denjoy minimal sets for area preserving diffeomorphisms, *Commun. Math. Helv.* **60**:508–557 (1985).
25. R. Mañé, An ergodic closing lemma, *Ann. Math.* **116**:503–541 (1982).
26. R. S. Mackay, Renormalisation in area preserving maps, Thesis, Princeton University, Princeton, New Jersey (1982).
27. R. S. Mackay, A renormalisation approach to invariant circles in area preserving maps, *Physica* **7D**:283–300 (1983).
28. R. S. MacKay, On Greene's residue criterion, preprint.
29. V. I. Osledec, A multiplicative ergodic theorem. Ljapunov characteristic numbers for dynamical systems, *Trans. Mosc. Math. Soc.* **19**:197–231 (1968).
30. A. Olvera and C. Simó, An obstruction method for the destruction of invariant curves, *Physica* **26D**:181–192 (1987).
31. A. Olvera and C. Simó, The obstruction method and some numerical experiments related to the standard map, in *Periodic Solutions of Hamiltonian Systems and Related Topics*, P. H. Rabinowitz *et al.*, eds. (Reidel, Amsterdam, 1987).
32. Y. Pesin, Characteristic Lyapunov exponents and smooth ergodic theory, *Russ. Math. Surv.* **32**:55–114 (1976).
33. Y. Pesin, Families of invariant manifolds corresponding to non-zero characteristic exponents, *Math. USSR Izv.* **10**:1261–1305 (1976).
34. H. Poincaré, *Les méthodes nouvelles de la mécanique céleste* (A. Blanchard, Paris, 1987).
35. D. Rana, Proof of accurate upper and lower bounds to stability domains in small denominator problems, Thesis, Princeton University, Princeton, New Jersey (1987).
36. D. Ruelle, Ergodic theory of differentiable dynamical systems, *Pub. Math. IHES* **50**:27–58 (1979).
37. G. Schmidt and J. Bialek, Fractal diagrams for Hamiltonian stochasticity, *Physica* **5D**:397–404 (1982).
38. J. Stoer and R. Burlisch, *Introduction to Numerical Analysis* (Springer-Verlag, New York, 1980).
39. C. L. Siegel and J. Moser, *Lectures on Celestial Mechanics* (Springer-Verlag, New York, 1971).
40. D. Salamon and E. Zehnder, KAM theory in configuration space, *Commun. Math. Helv.* **64**:84–132 (1989).
41. J. Wilbrink, Erratic behaviour of invariant circles in standard-like mappings, *Physica* **26D**:358–368 (1987).
42. J. Wilbrink, New fixed point of the renormalisation operator associated with the recurrence of invariant circles in generic Hamiltonian maps, *Nonlinearity* **3**:567–584 (1990).
43. A. Katok, Lyapunov exponents, entropy and periodic orbits for diffeomorphisms, *Pub. IHES* **51**:137–174 (1980).
44. V. I. Arnold and A. Avez, Ergodic problems of classical mechanics (Benjamin, New York, 1968).