Strange Attractors with One Direction of Instability

Qiudong Wang[1] and Lai-Sang Young[2]

TABLE OF CONTENT

[1]Department of Mathematics, University of Arizona, Tucson, AZ 85721, email dwang@math.arizona.edu. This research is partially supported by NSF grant #9970673 and an NSF Postdoctoral Research Fellowship

[2]Courant Institute of Mathematical Sciences, 251 Mercer St., New York, NY 10012, email lsy@cims.nyu.edu; and Department of Mathematics, UCLA, Los Angeles, CA 90095, email lsy@math.ucla.edu. This research is partially supported by a grant from NSF and a Guggenheim Fellowship

Strange attractors are of fundamental importance in dynamical systems; they have also been observed and recognized in many scientific disciplines. Up until now, most of the studies of strange attractors have relied on numerical simulations. Rigorous mathematical analysis has tended to be difficult, and progress has been slow. Among the not-so-many examples that have been studied are the Lorenz and Hénon attractors, both of which are closely related to certain one-dimensional maps.

The theory of one-dimensional maps, on the other hand, has experienced unprecedented growth in the last two decades. The purpose of this paper is to bring some of the techniques in one-dimension to bear on the analysis of attractors with a single direction of instability. More precisely, our aim is to develop a general theory of strange attractors with one unstable direction and $n-1$ directions of strong contraction, n being the dimension of the phase space. For simplicity, we will formulate our results in terms of perturbations of one-dimensional maps; what is important is that *locally* our dynamical systems have a one-dimensional character.

In this paper, we will treat only the case $n = 2$, where most of the interesting phenomena already occur, leaving the case of arbitrary phase dimension to be published elsewhere. In the rest of the introduction, we will focus on three aspects of this work that we regard as among the most important.

A. Conditions for the existence of strange attractors with known properties

One of the goals of this paper is to introduce an implementable scheme that would enable one to rigorously verify the existence of strange attractors with certain well defined "chaotic" properties. Leaving these properties to paragraph C below, we now give a rough description of the scheme we propose. Given a family of strongly dissipative maps with some expansion, i.e. a situation where a strange attractor potentially exists, we show that the following steps will ensure the desired conclusions for a positive measure set of parameters:

(a) First, pass to the singular limit by letting dissipation go to infinity. This gives a family of one-dimensional maps.

(b) Next, check that among these one-dimensional maps, there exist some with strong expanding properties (e.g. Misiurewicz maps).

(c) Then check that varying the parameter around the maps in (b) changes the dynamics effectively (this is a transversality condition).

(d) Finally, check that a nondegeneracy condition is satisfied in the unfolding, i.e. in the process that reverses (a) to recover the original dynamical systems.

These conditions are made precise in Sect. 1.1. We will show that *all* that one has to do to get the whole package of results described in paragraph C below (hyperbolicity, SRB measure, central limit theorem etc.) is to take the limit in (a), and then check (b)–(d); the latter two steps involve checking only that something is not equal to zero at a finite number of points.

The paper that paved the way for the use of one-dimensional techniques in two-dimensions is [BC2], in which Benedicks and Carleson showed, for certain parameters of the Hénon maps, that the attractor has positive Lyapunov exponents along its critical orbits (see paragraph B). Their estimates, however, use explicitly the formulas of the Hénon maps, making it difficult to apply directly the results in [BC2]. Extensions of [BC2] to small perturbations of Hénon maps have since been made, and some applications have been found; see e.g. [MV]. We do not claim by any means that this work is the first attempt to prove the existence of strange attractors, but we hope this is the most comprehensive attempt so far, both in terms of the clarity and generality of the conditions and in terms of the package of results that follow once these conditions are verified.

B. Geometry of critical regions

In this paragraph we discuss an object which dominates the landscape of the attractor, namely its *critical set*. In fact, it is important to understand not just the geometry of the critical set but the behavior of the map on its neighborhoods of various sizes; we call these *critical regions*.

Going back to one-dimension, there are basically two philosophically "different" ways to capture expanding properties for maps with critical points. There is the method of inducing used by Jakobson [J], which advocates, for an orbit passing near a critical point, to wait until it has regained a large derivative before looking at it again; and there is the idea first used by Collet and Eckmann [CE] and later in [BC1], which advocates imposing growth conditions directly on critical orbits.

It is the second approach described above that we will use in this paper to study the dynamics on attractors.

The idea of trying to identify a critical set for two-dimensional maps, that is to say, a set designated to play the role analogous to that played by critical points in one-dimension, goes back to [BC2]. The construction of the critical set in [BC2], however, is *ad hoc*, and the resulting object has no obvious intrinsic characterization. Moreover, while certain geometric relationships are satisfied, no coherent geometric picture of the critical regions follows from or is exploited in [BC2].

In order to develop a coherent geometric picture, we believe it is necessary to rework the entire inductive construction of [BC2] *with built-in geometric properties for the critical set as part of the induction.* This is what we have done in Part I of this paper. We have borrowed various pieces of local analysis from [BC2], but we have also added a geometric component to the story. This part is new, and as the reader will see, this departure from [BC2] will make a nontrivial difference when it comes to deriving dynamical consequences (see paragraph C). We remark that previous extensions of [BC2] have followed the inductive construction of [BC2] faithfully and are therefore also without these geometric considerations.

The critical set we introduce is an intrinsically defined object, characterized as those points on the attractor at which stable and unstable directions are interchanged. We prove that this set has a special geometric structure; it can be realized as the intersection of a nested sequence of rectangles with known geometric and dynamical properties. (For a quick desciption, see the statement of Theorem 1 in Sect. 1.2.) This geometric structure will be exploited heavily in the rest of the paper.

C. Dynamical consequences

The purpose of Part II of this paper is to develop, for the good parameters, properties that are consequences of the basic structures established in Part I. By "dynamical consequences", we refer to a comprehensive description of the attractor: its local and global structures, its dynamics as seen from statistical, geometric, combinatorial and symbolic points of view. We think of the first part of our paper as "sowing the seeds", and the second part as "reaping the harvest".

We state and prove in Part II of this paper more than a dozen results. Some of these results have been shown before for the Hénon maps; others are new even in that restricted context. Some require delicate proofs; others follow, with a little bit of work, from general theory. All are natural consequences of the picture established in Part I, namely hyperbolicity away from the critical regions, and the geometry of the critical regions. Taken together, they represent a fairly complete understanding of the class of attractors in question.

Statistical properties

From the statistical point of view, an inherent difficulty with dissipative dynamical systems is that *a priori* there is no natural invariant probability measure. By "natural", I refer to a measure that reflects the properties of Lebesgue-typical points. (The measure itself can be singular.) For systems with some hyperbolicity, there is the notion of a Sinai-Ruelle-Bowen or SRB measure introduced earlier in another context ([S2], [R1], [R2]). SRB measures are natural in the sense above. The problem is, not every attractor has an SRB measure.

Our first result is the existence of SRB measures for each of the attractors associated with a good parameter. These measures are not necessarily unique. In Section 8, we identify a finite number of ergodic SRB measures, called $\mu_1, \cdots, \mu_r$, and show that the asymptotic orbit distribution starting at Lebesgue-a.e. z in the basin is given by one of these μ_i. The domains of attraction of the μ_i can be quite complicated. In a way reminiscent of phase transitions, there are examples in which starting from certain arbitrarily small open sets, one has a positive probability of reaching several different μ_i. By appealing to some general results in [Y3] or [Y4], we prove exponential decay of correlations and a central limit theorem for the μ_i.

For the Hénon family near $a = 2,\ b = 0$ and their small perturbations, SRB measures and their statistical properties are studied in [BY1] and [BY2], and the basin property in [BV]. In this special case, the attractor admits only one SRB measure.

Geometric and other properties

It is useful conceptually to distinguish between the following two kinds of properties: properties that hold for Lebesgue-typical points, and properties carried by "small sets" or sets having Lebesgue measure zero. Because Lebesgue-typical points approach the attractor slowly, properties of the first kind require less precise information on the critical set. It is, in many ways, a greater challenge to understand the behavior of *every* orbit, for there is no control on how often or how close it comes to the critical set. Properties of the second kind, some of which we now describe, rely much more heavily on the detailed geometry of the critical set. All of our results in this category are new even for the Hénon maps.

A useful tool for keeping track of orbits in chaotic systems is to encode its orbits into symbolic sequences generated by a finite alphabet. Some encodings are more meaningful than others. Given that our attractors do not admit finite Markov partitions, we show in Section 10 that the situation is as good as can be: we have symbolic representations of orbits that reflect their true geometric locations, and a coding that is essentially one-to-one.

This coding allows us to identify our attractor with strings of symbols, bringing us closer to one-dimensional lattice models in statistical mechancis. A useful concept for dynamical systems borrowed from lattice models is that of an equilibrium state (see [S2], [R2]). We prove for our attractors the existence of equilibrium states, including measures of maximal entropy. We also prove various natural formulas for topological entropy, such as one given by counting the number of distinct "states" in n iterates.

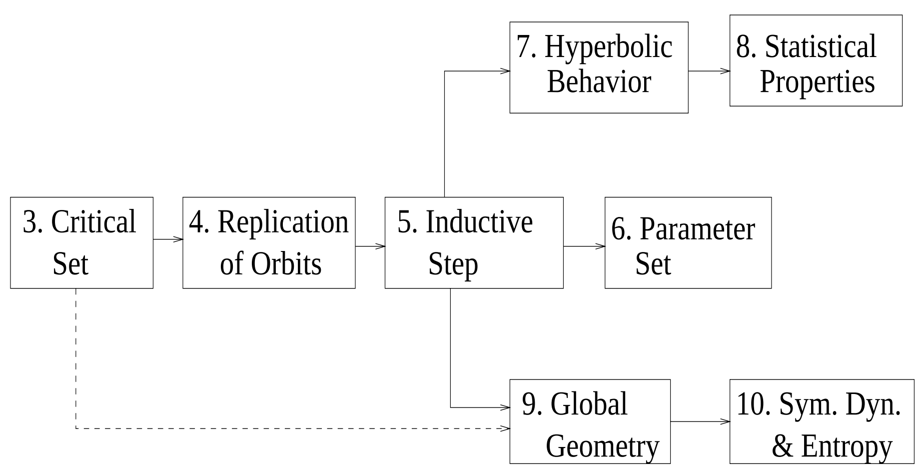
Uniformly hyperbolic or Axiom A attractors were among the first attractors to be understood (see [Sm], [Bo], [S2], [R2]). In Section 7, we show that our attractors, obviously nonuniformly hyperbolic as they are, can be seen as the limit of an increasing sequence of uniformly hyperbolic invariant sets.

Finally, in an entirely different direction, we give a finitary description of the approximate shape and complexity of these attractors, introducing a notion of "monotone branches". The way these branches fit together gives strong insight into the differences between one and two-dimensional maps.

This concludes our desciption of the content of this paper. Among the directions of research made more tractable by our results are questions on the zeta function and transfer operator (see e.g. [Bal], [PP], [R3]). With kneading sequences for critical orbits being well defined, it is reasonable to consider the possibility of a kneading theory (see [C], [MT]). Our topological discussion leads naturally to questions on prime ends (e.g. [Bar]).

There are many related works that have not yet been mentioned. First, there are various results in one-dimension (see [dMvS]) and on attractors with one direction of stability (not necessarily satisfying our conditions), including the solenoid and other Axiom A attractors [Sm], [W1], the Lorenz attractors (see [G], [Ro], [Ry], [W2]), dissipative twist maps [Bi], and certain periodically forced nonlinear oscillators ([Lev]; see also [GH]). Extending [BC2] and therefore closer to the setting of this paper are [DRV] and [V]. For results on piecewise uniformly hyperbolic attractors, see e.g. [CL], [I1], [I2], [M2] and [Y1]; and for statistical properties of hyperbolic billiards, see e.g. [S1], [BSC1], [BSC2] and [Y3].

This paper is by and large self-contained — with the exception of Section 6, where two results from 1-dimensional maps are quoted without proof, and Section 8, where previous work of the second-named author is used. Proofs that are computational in nature have been put in the Appendix so that they will not obstruct the main flow of ideas. In a paper as long as this one, it might be useful to indicate the logical connections among the various sections. After Section 1, we recommend at least looking through Section 2, in which we introduce much of the basic vocabulary for subsequent sections. The other sections are connected as indicated. (For example, the technical content of Section 6 is not needed for reading Sections 7-10.)



1 Statements of Results

1.1 Setting

For definiteness, Theorems 1–7 are stated in the context of attractors that arise from perturbations of *circle* maps. For the interval case, see Sect. 1.5.

Let $A = S^1 \times [-1,1]$. We consider 2-parameter families of maps $\{T_{a,b}\}$ where for each (a,b), $T_{a,b} : A \to A$ is a self-map of A and $(x,y,a,b) \mapsto T_{a,b}(x,y)$ is C^4. The class of 2-parameter families $\{T_{a,b}\}$ to which our results apply are constructed via the following four steps. The necessary smoothness is assumed in each step.

Step I. Let $f : S^1 \to S^1$ satisfy the following **Misiurewicz conditions**, i.e. letting $C = \{x : f'(x) = 0\}$, we assume:

1. $f''(x) \neq 0$ for all $x \in C$;
2. f has negative Schwarzian derivative on $S^1 \setminus C$; [3]
3. there is no $x \in S^1$ with $f^n(x) = x$ and $|(f^n)'(x)| \leq 1$;
4. for all $x \in C$, $\inf_{n>0} d(f^n x, C) > 0$.

Observe that for $p \in S^1$ with $\inf_{n\geq 0} d(f^n p, C) > 0$, if g is sufficiently near f in the C^2 sense, then there is a unique point $p(g)$ having the same symbolic dynamics with respect to g as p does with respect to f. If $\{f_a\}$ is a 1-parameter family through f, then for those a for which it makes sense, we will call $p(a) = p(f_a)$ the *continuation* of p. For $x \in C$, we let $x(a)$ denote the corresponding critical point of f_a.

Step II. Let f be as in Step I, and let $\{f_a\}$, $a \in [a_0, a_1]$, be a 1-parameter family of maps from S^1 to S^1 with $f = f_{a^*}$ for some $a^* \in [a_0, a_1]$. We require that $\{f_a\}$ satisfy the following **transversality condition**[4]: For every $x \in C$, if $p = f(x)$, then

$$\frac{d}{da} f_a(x(a)) \neq \frac{d}{da} p(a) \qquad \text{at} \quad a = a^*. \tag{1}$$

Step III. Let $\{f_a\}$ be as in Step II. Identifying S^1 with $S^1 \times \{0\} \subset A$, we extend $\{f_a\}$ to a 2-parameter family $\{f_{a,b}\}$, $a \in [a_0, a_1]$, $b \in [0, b_1]$, where $f_{a,b} : S^1 \to A$ is such that $f_{a,0} = f_a$ and $f_{a,b}$ is an *embedding* for $b > 0$.

Step IV. Finally, we extend $f_{a,b}$ to $T_{a,b} : A \to A$ in such a way that $T_{a,0}(A) \subset S^1 \times \{0\}$ and for $b > 0$, $T_{a,b}$ maps A diffeomorphically onto its image. We further impose the following **non-degeneracy condition**[5] on the map $T_{a^*,0}$:

$$\partial_y T_{a^*,0}(x,0) \neq 0 \quad \text{whenever} \quad f'_{a^*}(x) = 0. \tag{2}$$

[3] This condition can be dropped but the proofs would be more complicated.

[4] This transversality condition is used in [TTY].

[5] This condition is not assumed in [MV] or [BV]. Their regularity condition on $|\det(DT)|$ and bound on the perturbation term, however, imply a condition which is similar (though not equivalent) to (2) and which serves a similar purpose.

This completes our construction of admissible families $\{T_{a,b}\}$. We remark that the transversality and non-degeneracy conditions in Steps II and IV are generic. Thinking in terms of normal neighborhoods, one constructs easily for a given $f_{a,b}$ extensions of the type in Step IV; the signs of the ∂y-derivatives at the critical points of f_{a^*} are determined by the orientations of the turns of $f_{a,b}$ at the corresponding points. Step III is feasible if and only if the degree of f is $0, 1$ or -1. If $|\deg(f)| > 1$, an extra dimension is needed; this will be treated in a separate paper.

Finally, we observe that for $b > 0$, $T_{a,b}$ has the general form

$$T_{a,b} : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} F(x, y, a) + b\ u(x, y, a, b) \\ b\ v(x, y, a, b) \end{pmatrix}$$

where $F(x, y, a) = T_{a,0}(x, y)$ and the C^3 norms of $(x, y, a) \mapsto u(x, y, a, b)$ and $v(x, y, a, b)$ are uniformly bounded for all $b \in (0, b_1]$.[6] We may, in fact, replace the C^4 assumption at the beginning of Sect. 1.1 by the expression for $T_{a,b}$ above and the requirement of uniformly bounded C^3 norms.

Notation. Given $\{T_{a,b}\}$, constants that are determined entirely by the family $\{T_{a,b}\}$ will be referred to as **system constants**. Except where declared otherwise, the letter K is reserved throughout this article for use as a **generic system constant**, meaning a system constant that is allowed to change from statement to statement (the other system constants are fixed). We will use K_1, K_2 etc. where K appears in more than one role in the same statement.

Let K be such that $T_{a,b}(A) \subset R_0 := S^1 \times [-Kb, Kb]$ for all (a, b). It is convenient for us to work with R_0 instead of A. For $T = T_{a,b}$, let $R_n = T^n R_0$. Then $\{R_n\}$ is a decreasing sequence of neighborhoods of the **attractor** $\Omega := \cap_{n=0}^{\infty} R_n = \cap_{n=0}^{\infty} T^n A$.

1.2 Critical set and hyperbolic behavior

Our first theorem identifies, for each map T corresponding to a selected set of parameters, a fractal set $\mathcal{C}$ chosen to play the role of the critical set in 1-dimension. This set will be called the **critical set** of T. Our parameter selection imposes strong hyperbolic properties on orbits starting from $\mathcal{C}$ in the hope that these properties will be passed on to the rest of the system. The geometric structure near $\mathcal{C}$, which is described in some detail in Theorem 1, is crucial for many of our later results.

For $z_0 \in R_0$, let $z_i = T^i z_0$. If w_0 is a tangent vector at z_0, let $w_i = DT^i(z_0) w_0$. A curve in R_0 is called a **$C^2(b)$-curve** if the slopes of its tangent vectors are $\mathcal{O}(b)$ and its curvature is everywhere $\mathcal{O}(b)$.

[6] This is a calculus exercise: Observe that bu extends to a C^4 function g on $\{b \geq 0\}$ with $g|\{b = 0\} = 0$. Writing $\partial^3 = \frac{\partial^3}{\partial z_1 \partial z_2 \partial z_3}$ where $z_i = x, y$ or a, we then check that $\partial^3 u$ extends to a continuous function h on $\{b \geq 0\}$ with $h = \frac{\partial}{\partial b} \partial^3 g$ on $\{b = 0\}$.

Theorem 1 (Parameter selection and the critical set) *Given $\{T_{a,b}\}$ as in Sect. 1.1, there is a positive measure set $\Delta \subset [a_0, a_1] \times (0, b_1]$ such that (1) and (2) below hold for $T = T_{a,b}$ for all $(a, b) \in \Delta$. The set Δ is located near $a = a^*$ and $b = 0$; it has the property that for all sufficiently small b, $\Delta_b := \{a : (a, b) \in \Delta\}$ has positive 1-dimensional Lebesgue measure. The constants $\alpha, \delta, c > 0$ and $0 < \rho < 1$ below are system constants, and $b << \alpha, \delta, \rho, e^{-c}$ for all $(a, b) \in \Delta$.*

(1) **Geometry of critical regions and critical set.** *There is a Cantor set $\mathcal{C} \subset \Omega$ called the critical set given by $\mathcal{C} = \cap_{k=0}^{\infty} \mathcal{C}^{(k)}$ where the $\mathcal{C}^{(k)}$ are a decreasing sequence of neighborhoods of $\mathcal{C}$ called critical regions. More precisely,*

(i) $\mathcal{C}^{(0)} = \{(x, y) \in R_0 : d(x, C) < \delta\}$ where C is the set of critical points of f.

(ii) $\mathcal{C}^{(k)}$ has a finite number of components called $Q^{(k)}$ each one of which is diffeomorphic to a rectangle. The boundary of $Q^{(k)}$ is made up of two $C^2(b)$ segments of ∂R_k connected by two vertical lines: the horizontal boundaries are $\approx \min(2\delta, \rho^k)$ in length, and the Hausdorff distance between them is $\mathcal{O}(b^{\frac{k}{2}})$.

(iii) $\mathcal{C}^{(k)}$ is related to $\mathcal{C}^{(k-1)}$ as follows: $Q^{(k-1)} \cap R_k$ has at most finitely many components, each one of which lies between two $C^2(b)$ subsegments of ∂R_k that stretch across $Q^{(k-1)}$ as shown. Each component of $Q^{(k-1)} \cap R_k$ contains exactly one component of $\mathcal{C}^{(k)}$.

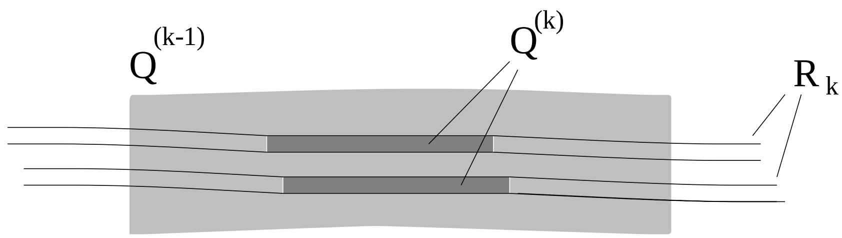


Figure 1 Critical regions

(2) **Properties of critical orbits.** *On each horizontal boundary γ of each component $Q^{(k)}$ of $\mathcal{C}^{(k)}$, $k = 0, 1, 2, \cdots$, there is a unique point z_0 characterized by the following two properties:*

(i) $\|DT^j(z_0)\binom{0}{1}\| \geq K^{-1}e^{cj}$ for all $j > 0$.

(ii) If τ is a unit tangent vector to γ at z_0, then $\|DT^n(z_0)\tau\| < (Kb)^n$ $\forall n > 0$.

The point z_0 is located within $\mathcal{O}(b^{\frac{k}{4}})$ of the midpoint of γ. Let Γ be the set of all of these points, and let $d_{\mathcal{C}}(\cdot)$ be the notion of "distance to the critical set" defined below. Then $z_0 \in \Gamma$ also satisfies

(iii) $d_{\mathcal{C}}(z_j) \geq K^{-1}e^{-\alpha j}$ for all $j > 0$.

Finally, since the critical set $\mathcal{C}$ is the accumulation set of Γ, properties (i) and (iii) of Γ are passed on to $\mathcal{C}$.

For $z \in R_0$, $d_{\mathcal{C}}(z)$ is defined as follows: For $z \in \mathcal{C}^{(0)} \setminus \mathcal{C}$, let k be the largest number with $z \in \mathcal{C}^{(k)}$, and let $d_{\mathcal{C}}(z)$ be the horizontal distance between z and the midpoint of the component of $\mathcal{C}^{(k)}$ containing z; for $z \notin \mathcal{C}^{(0)}$, use the component of $\mathcal{C}^{(0)}$ nearest to z.

Theorems 2-7 apply to $T = T_{a,b}$, $(a,b) \in \Delta$, where Δ is as in Theorem 1.

Our next theorem is about the abundance of hyperbolic behavior on the attractor and in the basin. A compact T-invariant set Λ is called **uniformly hyperbolic** if there is a splitting of the tangent bundle over Λ into invariant subbundles $E^u \oplus E^s$ such that for some $C, \lambda > 1$, we have, for all $n \geq 1$, $\|DT^n v\| \leq C\lambda^{-n}\|v\|$ for all $v \in E^s$ and $\|DT^{-n} v\| \leq C\lambda^{-n}\|v\|$ for all $v \in E^u$.

Theorem 2 (Hyperbolic behavior)

(1) Let

$$\Omega_\varepsilon := \{z_0 \in \Omega : d_{\mathcal{C}}(z_n) \geq \varepsilon \quad \forall n \in \mathbb{Z}\}.$$

(i) For every $\varepsilon > 0$, Ω_ε is uniformly hyperbolic. In fact, independent of ε, λ in the definition of hyperbolicity can be taken to be $\approx e^{\frac{c}{3}}$ where c is as in Theorem 1. In particular, for evevry periodic point $z \in \Omega$ with $T^q z = z$, $\|DT^q | E^u(z)\| \geq K^{-1} e^{\frac{c}{3}q}$.

(ii) As $\varepsilon \to 0$, the hyperbolicity on Ω_ε deteriorates in the sense that $C \to \infty$ and the minimum angle between E^u and E^s tends to zero.

(iii) $\Omega = \overline{\cup_{\varepsilon > 0} \Omega_\varepsilon}$ provided a surjective condition of the type (*) below is assumed.

(2) Under the regularity conditions (**) below, we have

$$\limsup_{n \to \infty} \frac{1}{n} \log \|DT^n(z_0)\| \geq \frac{c}{3}$$

for Lebesgue-almost every $z_0 \in R_0$.

The two technical conditions used in parts (1)(iii) and (2) of Theorem 2 are:

(*) Let $J_1, \cdots J_r$ be the intervals of monotonicity of f. Then for each i, there exists j such that $f(J_j) \supset J_i$.

(**) There exist $K_1, K_2 > 0$ such that for all $z \in R_0$,

$$K_1^{-1} b \leq |\det(DT_{a,b}(z))| \leq K_2 b$$

We remark that Theorem 2(1) confirms that $\mathcal{C}$ is the sole source of nonhyperbolicity in the system. Part (2) expresses the fact that many orbits experience at least some form of (nonuniform) hyperbolicity. A more detailed discussion is given in Section 7.

1.3 SRB Measures and their Statistical Properties

Definition 1.1 *Let $g : M \to M$ be a diffeomorphism of a manifold. A g-invariant Borel probability measure μ is called an* **SRB measure** *if g has a positive Lyapunov exponents $\mu -$ a.e. and the conditional measures of μ on unstable manifolds are absolutely continuous with respect to the Riemannian measure on these manifolds.*

In the absence of zero Lyapunov exponents, it follows from general hyperbolic theory that an SRB measure has at most a countable number of ergodic components, and that each ergodic component has a positive measure set of generic points. A point z is said to be **generic** with respect to μ if for every continuous function φ, $\frac{1}{n}\sum_{i=0}^{n} \varphi(g^i z) \to \int \varphi d\mu$ as $n \to \infty$. See [Led] and [PS].

Theorem 3 (Existence and ergodic properties of SRB measures)

(1) T admits an SRB measure.

Assuming condition (**) above, we have the following additional information:

(2) T admits at most r ergodic SRB measures μ_i where r is the cardinality of the critical set of the 1-dimensional map f.

(3) Lebesgue-a.e. $z_0 \in R_0$ is generic with respect to some μ_i; in fact, Lebesgue-a.e. $z_0 \in R_0$ lies in the stable manifold of a μ_i-typical point in Ω.

We know from general hyperbolic theory that without zero Lyapunov exponents, ergodic components of SRB measures are, up to finite factors, mixing [Led].

Theorem 4 (Decay of correlations and Central Limit Thoerem) *Let μ be an ergodic SRB measure, which, by taking a power of T if necessary, we assume to be mixing. Then*

(1) for each $\eta \in (0, 1]$, there exists $\lambda = \lambda(\eta) < 1$ such that if $\psi : A \to \mathbb{R}$ is Hölder continuous with exponent η and $\varphi \in L^\infty(\mu)$, then there exists $K(\varphi, \psi)$ such that

$$\left| \int (\varphi \circ T^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu \right| < K(\varphi, \psi)\lambda^n \quad \text{for all } n;$$

(2) the Central Limit Theorem holds for all Hölder φ with $\int \varphi d\mu = 0$, i.e.

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \varphi \circ T^i \quad \to \quad \mathcal{N}(0, \sigma)$$

where $\mathcal{N}(0,\sigma)$ is the normal distribution with variance σ^2; furthermore, $\sigma > 0$ if and only if $\varphi \circ T \neq \psi \circ T - \psi$ for any ψ.

We remark that the word "attractor" has different meanings in the literature (see [Mil] for a discussion). In this article, it is convenient for us to refer to Ω as "the attractor". Theorem 3 suggests, however, that from a measure-theoretic point of view, it may be more appropriate to regard the supports of the μ_i as attractors.

1.4 Global geometry, symbolic dynamics and topological entropy

A **monotone branch** of R_n is a region diffeomorphic to a rectangle and bordered by two subsegments of ∂R_n. Roughly speaking, it is the largest domain of this kind with the property that for $0 \leq i \leq n$, the x-coordinates of its T^{-i}-image stay inside some interval of monotonicity of f, where f is the initial 1-dimensional map from which $\{T_{a,b}\}$ is built. This notion is made precise in Section 9, where a combinatorial tree is introduced to describe the structure of a natural class of monotone branches.

Theorem 5 (Coarse geometry of attractor) *There is a sequence of neighborhoods $\tilde{R}_n$ of Ω with*

$$\tilde{R}_1 \supset \tilde{R}_2 \supset \tilde{R}_3 \supset \cdots \qquad \text{and} \qquad \cap_i \tilde{R}_i = \Omega$$

such that each $\tilde{R}_n$ is the union of a finite number of monotone branches of R_k, $n \leq k \leq n(1+K\theta)$, where $\theta \sim \frac{-1}{\log b}$.

Let $\{1, 2, \cdots, k\}$ be a finite alphabet and let Σ_k be the set of all bi-infinite sequences $\mathbf{s} = (\cdots, s_{-1}, s_0, s_1, \cdots)$ with $s_i \in \{1, 2, \cdots, k\}$. The shift operator $\sigma : \Sigma_k \to \Sigma_k$ is defined by $(\sigma \mathbf{s})_i = (\mathbf{s})_{i+1}$. For $\Sigma \subset \Sigma_k$, we call $\sigma|\Sigma : \Sigma \to \Sigma$ a **subshift** of the full shift on k symbols if Σ is a closed σ-invariant subset of Σ_k.

Let $x_1 < x_2 < \cdots < x_r < x_{r+1} = x_1$ be the critical points of f. Let $\mathcal{C}_i^{(0)}$ be the component of $\mathcal{C}^{(0)}$ containing x_i and let $\mathcal{C}_i = \mathcal{C} \cap \mathcal{C}_i^{(0)}$. We remark that each $\mathcal{C}_i$ is a fractal set – it is not contained in any smooth curve – and that *a priori* there is no well defined notion of whether a point lies to the left or to the right of $\mathcal{C}_i$.

Theorem 6 (Coding of orbits on attractor)

(1) The critical set $\mathcal{C}$ partitions $\Omega \backslash \mathcal{C}$ into disjoint sets $A_1, A_2, \cdots, A_r$ so that $z \in A_i$ has the interpretation of being "to the right" of $\mathcal{C}_i$ and "to the left" of $\mathcal{C}_{i+1}$.

(2) *There is a subshift* $\sigma : \Sigma \to \Sigma$ *of a full shift on finitely many symbols and a continuous surjection* $\pi : \Sigma \to \Omega$ *such that*

$$T \circ \pi \ = \ \pi \circ \sigma;$$

π *is 1-1 except on* $\cup_{i=-\infty}^{\infty} T^i \mathcal{C}$, *where it is 2-1.*

(3) *Under the additional assumption that* $f[x_j, x_{j+1}] \not\supset S^1$ *for any* j, *the coding in (2) is given by (1), i.e. for all* $z_0 \in \Omega \setminus \cup_{i=-\infty}^{\infty} T^i \mathcal{C}$, $\pi^{-1}(z_0)$ *is the unique sequence* $(s_i)_{i=-\infty}^{\infty}$ *with* $z_i \in A_{s_i}$.

Corollary 1 (Kneading sequences for critical points) *For every* $z_0 \in \mathcal{C}$, *the itinerary of* $\{z_1, z_2, \cdots\}$ *is uniquely represented by a sequence in* Σ.

Another consequence of Theorem 6 is the existence of equilibrium states. For a continuous map $g : X \to X$ of a compact metric space and a continuous function $\varphi : X \to \mathbb{R}$, a g-invariant Borel probability measure μ on X is called an **equilibrium state** for g with respect to the potential φ if μ maximizes the quantity

$$\sup \{ \ \mathrm{h}_\nu(\mathrm{g}) \ + \int \varphi \mathrm{d}\nu \ \}$$

where $h_\nu(g)$ denotes the metric entropy of g with respect to ν and the supremum is taken over all g-invariant Borel probability measures ν.

Corollary 2 (Existence of equilibrium states) *T has an equilibrium state for every continuous* $\varphi : \Omega \to R$. *In particular, T admits an invariant Borel probability measure maximizing entropy.*

The **topological entropy** of g, written $h_{top}(g)$, is usually defined in terms of open covers of arbitrarily small diameters or in terms of (n, ε)-spanning or separated sets. For precise definitions, see [Wa]. For the class of attractors studied in this paper, $h_{top}(g)$ can be computated in more concrete ways.

In Theorem 6 we saw that every $z_0 \in \Omega$ can be unambiguously associated with one (and occasionally two) symbol sequences in Σ determined by the locations of its iterates with respect to the components of the critical set. We will show in Section 10 that in like manner all the points in R_0 can be assigned symbol sequences – except that this assignment is not unique. Let us temporarily refer to this as the "fuzzy" coding on R_0. Let

N_n = number of distinct n-blocks in the coding of Ω;
$\tilde{N}_n$ = number of distinct n-blocks in the "fuzzy" coding of R_0;
P_n = number of fixed points of T^n;
$M_n^\pm$ = number of monotone segments in $\partial R_n^\pm$, the two boundary components of R_n (see Sect. 9.1 for the precise definition).

Theorem 7 (Formulas and inequalities for topological entropy)

(i)

$$h_{\text{top}}(T) \;=\; \lim_{n\to\infty} \frac{1}{n} \log N_n \;=\; \lim_{n\to\infty} \frac{1}{n} \log \tilde{N}_n \;=\; \lim_{n\to\infty} \frac{1}{n} \log P_n.$$

(ii)

$$\limsup_{n\to\infty} \frac{1}{n} \log M_n^{\pm} \;\le\; h_{\text{top}}(T) \;\le\; \left(\liminf_{n\to\infty} \frac{1}{n} \log M_n^{\pm}\right)(1 + \frac{K}{\log \frac{1}{b}}).$$

For a 1-dimensional piecewise monotonic map g, it is a well known fact that $h_{\text{top}}(g)$ is the growth rate of the number of intervals on which g^n is monotonic [MS]. The factor $(1 + \frac{K}{\log \frac{1}{b}})$ gives, in a sense, the potential *defect* in measuring the complexity of T via the 1-dimensional curves $\partial R_0^{\pm}$.

1.5 Hénon maps and homoclinic bifurcations

Theorems 1–7 are stated for attractors that arise from perturbations of circle maps. We state here, for the record, the corresponding results for *interval maps* and some of their applications. Reduction to the circle case is carried out in Appendix A.1.

Theorem 8 (Attractors arising from interval maps) *Let I be a closed interval of finite length, and let $f : I \to I$ be a Misiurewicz map with $f(I) \subset int(I)$. Let U be a neighborhood of $I \times \{0\}$ in $\mathbb{R}^2$, and let $\{T_{a,b}\}$ be a 2-parameter family of maps with $T_{a,b} : U \to \mathbb{R}^2$. We identify I with $I \times \{0\} \subset \mathbb{R}^2$, and assume that $\{T_{a,b}\}$ satisfies the conditions in Steps II, III and IV in Sect. 2.1 with $f_{a^*} = f$. Then*

(i) there exist $K > 0$ and a rectangle $\hat{\Delta} = [a_0, a_1] \times (0, b_1]$ arbitrarily near $(a^*, 0)$ such that for each $(a, b) \in \hat{\Delta}$, $T_{a,b}$ maps $R := I \times [-Kb, Kb]$ strictly into its interior, defining an attractor $\Omega := \bigcap_{n\ge 0} T_{a,b}^n R$;

(ii) there is a positive measure set $\Delta \subset \hat{\Delta}$ such that the conclusions of Theorems 1–7 hold for $T = T_{a,b} \mid R$ for all $(a, b) \in \Delta$.

Corollary 3 (The Hénon family) *Let*

$$T_{a,b} : (x, y) \mapsto (1 - ax^2 + y,\; bx), \qquad (x, y) \in \mathbb{R}^2.$$

Then for every $a^* \in [1.5, 2]$ for which $f_{a^*} : x \mapsto 1 - a^* x^2$ is a Misiurewicz map, the conclusions of Theorem 8 hold. In particular, there is a positive measure set Δ near $(a^*, 0)$ such that the conclusions of Theorems 1–7 hold for all $T = T_{a,b}$, $(a, b) \in \Delta$. These results are valid for both $b > 0$ and $b < 0$.

When specialized to $a^* = 2$ and $b > 0$, the part of Corollary 3 that corresponds to Theorem 1, part (2), in this paper is a version of the main result of [BC2]. The results in [BY1], [BV], and [BY2] are respectively the parts of Corollary 3 that the correspond to Theorem 3(1),(2), Theorem 3(3) and Theorem 4.

Our last result concerns the application of Theorems 1–7 to homoclinic bifurcations. Let g_μ, $\mu \in [0,1]$, be a C^∞ one-parameter family of surface diffeomorphisms unfolding at $\mu = 0$ a nondegenerate tangency of $W^u(p_0)$ and $W^s(p_0)$ where p_0 is a hyperbolic fixed point. We assume that the eigenvalues λ and σ of Dg_0 at p_0 satisfy $0 < \lambda < 1 < \sigma$ and $\lambda\sigma < 1$, and that they belong in the open and dense set of eigenvalue pairs that meet the hypotheses of Sternberg's linearization theorem. Under these conditions, it is well known (see [PT]) that for all sufficiently large k, there is a positive measure set of parameters $\hat{\Delta}_k$ such that for all $\mu \in \hat{\Delta}_k$, g_μ has a k-periodic attractor Ω_μ all but finitely many of whose periodic components are located near the fixed point p_μ.

Theorem 9 (Attractors arising from homoclinic bifurcations) *Let g_μ be as above. Then for all sufficiently large k, there is a positive measure set of parameters $\Delta_k \subset \hat{\Delta}_k$ for which the following hold: for all $\mu \in \Delta_k$, there is a component Ω^0_μ of Ω_μ with the property that if T_μ denotes the restriction of g^k_μ to a neighborhood of Ω^0_μ, then the conclusions of Theorems 1–7 hold for $T = T_\mu$.*

Our proof of Theorem 9, which is given in Appendix A.2, consists of observing that the maps g^k_μ meet the conditions of Theorem 8. The part of Theorem 9 that corresponds to Theorem 1, part (2), in this paper is the main result of [MV].

2 Preliminaries

We gather in this section a collection of technical facts used repeatedly in later sections. Most of the proofs are given in Appendix B. Sects. 2.1–2.4 contain material not specific to the family $\{T_{a,b}\}$, and K is not a "system constant" in these subsections.

2.1 Linear algebra

Let M be a 2×2 matrix. Assuming that M is not a scalar multiple of an orthogonal matrix, we say that a unit vector e defines **the most contracted direction** of M if $\|Mu\| \geq \|Me\|$ for all unit vectors u. For a sequence of matrices $M_1, M_2, \cdots$, we use $M^{(i)}$ to denote the matrix product $M_i \cdots M_2 M_1$ and e_i to denote the most contracted direction of $M^{(i)}$ when it makes sense.

Hypotheses for Sect. 2.1 The M_i are 2×2 matrices; they satisfy $|\det(M_i)| \leq b$ and $\|M_i\| \leq K_0$ where K_0 and b are fixed numbers with $K_0 > 1$ and $b << 1$.

Lemma 2.1 *There exists K depending only on K_0 such that if $\|M^{(i)}\| \geq \kappa^i$ and $\|M^{(i-1)}\| \geq \kappa^{i-1}$ for some $\kappa >> \sqrt{b}$, then e_i and e_{i-1} are well-defined, and*

$$\| e_i \times e_{i-1} \| \leq (\frac{Kb}{\kappa^2})^{i-1}.$$

Corollary 2.1 *If for $1 \leq i \leq n$, $\|M^{(i)}\| \geq \kappa^i$ for some $\kappa >> \sqrt{b}$, then:*

(a) $\| e_n - e_1 \| < \frac{Kb}{\kappa^2}$;

(b) $\|M^{(i)} e_n\| \leq (\frac{Kb}{\kappa^2})^i$ *for* $1 \leq i \leq n$.

Proof: (a) follows immediately from Lemma 2.1. For (b), since $\|e_n - e_i\| \leq (\frac{Kb}{\kappa^2})^i$, we have $\|M^{(i)} e_n\| \leq \|M^{(i)}(e_n - e_i)\| + \|M^{(i)} e_i\| < K_0^i \cdot (\frac{Kb}{\kappa^2})^i + (\frac{b}{\kappa})^i$. □

Next we consider for each i a 3-parameter family of matrices $M_i(s_1, s_2, s_3)$. For the purpose of the next corollary we make the additional assumptions that for $0 < j \leq 3$, $\|\partial^j M_i(s_1, s_2, s_3)\| \leq K_0^i$ and $|\partial^j \det(M_i(s_1, s_2, s_3))| < K_0^i b$ where ∂^j represents any one of the partial derivatives of order j with respect to s_1, s_2 or s_3. Let $\theta_i(s_1, s_2, s_3)$ denote the angle $e_i(s_1, s_2, s_3)$ makes with the positive x-axis, assuming it makes sense.

Corollary 2.2 *Suppose that for some $\kappa >> \sqrt{b}$, $\|M^{(i)}(s_1, s_2, s_3)\| \geq \kappa^i$ for every (s_1, s_2, s_3) and for every $1 \leq i \leq n$. Then for $j = 1, 2, 3$, $|\partial^j \theta_1| \leq K \kappa^{-(1+j)}$, and for $i \leq n$,*

$$|\partial^j (\theta_i - \theta_{i-1})| < (\frac{Kb}{\kappa^{(2+j)}})^{i-1}, \tag{3}$$

$$\|\partial^j M^{(i)} e_n\| < (\frac{Kb}{\kappa^{(2+j)}})^i. \tag{4}$$

Our next lemma is a perturbation result. Let M_i, M_i' be two sequences of matrices, let w be a vector, and let θ_i and θ_i' denote the angles $M^{(i)}w$ and $M'^{(i)}w$ make with the positive x-axis respectively.

Lemma 2.2 ([BC2], Lemma 5.5) *Let κ, λ be such that $\frac{Kb}{\kappa^2} < \lambda < K_0^{-12}\kappa^8$. If for $1 \leq i \leq n$, $\|M_i - M_i'\| \leq \lambda^i$ and $\|M^{(i)}w\| \geq \kappa^i$, then*
(a) $\|M'^{(n)}w\| \geq \frac{1}{2}\kappa^n$;
(b) $|\theta_n - \theta_n'| < \lambda^{\frac{n}{4}}$.

Proofs of Lemmas 2.1, 2.2 and Corollary 2.2 are given in Appendix B.1.

Hypothesis for Sects. 2.2 and 2.3: $T : A \to A$ is an embedding of the form

$$T(x, y) = (t_1(x, y), bt_2(x, y))$$

where the C^2-norms of t_1 and t_2 are $\leq K_0$, and $K_0 > 1$ and $b << 1$ are fixed numbers.

2.2 Stable curves

Lemma 2.3 *Let κ, λ be as in Lemma 2.2 and $z_0 \in A$ be such that for $i = 1, \cdots, n$, $\|DT^i(z_0)\| \geq \kappa^i$. Then there is a C^1 curve γ_n passing through z_0 such that*
(a) for all $z \in \gamma_n$, $d(T^i z_0, T^i z) \leq (\frac{Kb}{\kappa^2})^i$ for all $i \leq n$;
(b) γ_n can be extended to a curve of length $\sim \lambda$ or until it meets ∂A.

A proof of this lemma is given in Appendix B.2.

We call γ_n a **stable curve of order** n. It will follow from this lemma that if $\|DT^i(z_0)\| \geq \kappa^i$ for all $i > 0$, then there is a **stable curve** γ_∞ passing through z_0 obtained as a limit of the γ_n's.

2.3 Curvature estimates

Let $\gamma_0 : [0, 1] \to A$ be a C^2 curve, and let $\gamma_i(s) = T^i(\gamma_0(s))$. We denote the curvature of γ_i at $\gamma_i(s)$ by $k_i(s)$.

Lemma 2.4 *Let $\kappa > b^{\frac{1}{3}}$. We assume that for every s, $k_0(s) \leq 1$ and*

$$\|DT^j(\gamma_{n-j}(s))\gamma_{n-j}'(s)\| \geq \kappa^j \|\gamma_{n-j}'(s)\|$$

for every $j < n$. Then

$$k_n(s) \leq \frac{Kb}{\kappa^3}.$$

A proof is given in Appendix B.3.

2.4 One-dimensional dynamics

We begin with some properties of maps satisfying the Misiurewicz condition. Let f be as in Sect. 1.1, and let $C_\delta := \{x \in S^1 : d(x, C) < \delta\}$.

Lemma 2.5 *There exist $\hat{c}_0, \hat{c}_1 > 0$ such that the following hold for all sufficiently small $\delta > 0$: Let $x \in S^1$ be such that $x, fx, \cdots, f^{n-1}x \notin C_\delta$, any n. Then*

(i) $|(f^n)'x| \geq \hat{c}_0 \delta e^{\hat{c}_1 n}$;

(ii) if, in addition, $f^n x \in C_\delta$, then $|(f^n)'x| \geq \hat{c}_0 e^{\hat{c}_1 n}$.

A proof is given in Appendix B.4.

Corollary 2.3 *Let $c_0 < \hat{c}_0$ and $c_1 < \hat{c}_1$. Then for all sufficiently small δ, there exists $\varepsilon = \varepsilon(\delta)$ such that for all g with $\|g - f\|_{C^2} < \varepsilon$, (i) and (ii) above hold for g with c_0 and c_1 in the places of $\hat{c}_0$ and $\hat{c}_1$.*

Proof: Let N be such that $\delta e^{\hat{c}_1 N} > e^{c_1 N}$, and choose ε small enough so that for all $i \leq N$, if $x, gx, \cdots, g^{i-1}x \notin C_\delta(g)$, then $(g^i)'x \approx (f^i)'x$. □

The results in the rest of this subsection are not needed in this article. We include them only as motivation for the corresponding results in 2-dimensions.

Temporarily write $C = C(g)$. To control $(g^n)'x$ when $g^i x \in C_\delta$ for some $i < n$, we need to impose further conditions on g. Following [BC1] and [BC2], we assume there exist $\lambda > 1$ and $0 < \alpha << 1$ such that for all $\hat{x} \in C$ and $n \geq 0$:

(a) $d(g^n \hat{x}, C) \geq c_0 e^{-\alpha n}$ and

(b) $|(g^n)'(g\hat{x})| \geq c_0 \lambda^n$.

We define for each $x \in C_\delta$ a **bound period** $p(x)$ as follows. Fix $\beta > \alpha$. Let $\hat{x} \in C$ be such that $|x - \hat{x}| < \delta$. Then $p(x)$ is the smallest p such that

$$|g^p x - g^p \hat{x}| > c_0 e^{-\beta p}.$$

Lemma 2.6 (Derivative recovery) *There exists K such that for g satisfying the conditions above, if $|x - \hat{x}| = e^{-\mu} < \delta$ for some $\hat{x} \in C$, then*

(i) $K^{-1}\mu \leq p(x) \leq K\mu$;

(ii) $K^{-1}(x - \hat{x})^2 |(g^{i-1})'(g\hat{x})| < |g^i x - g^i \hat{x}| < K(x - \hat{x})^2 |(g^{i-1})'(g\hat{x})|$;

(iii) $|(g^p)'x| \geq K^{-1}\lambda^{\frac{p}{2}}$ where $p = p(x)$.

Proof: For this result there is no substantive difference between the situation here and that of the quadratic family $x \mapsto 1 - ax^2$. See [BC1] and [BC2], Section 2. □

Standing hypotheses for the rest of the paper: $\{T_{a,b}\}$ is as in Sect. 1.1. In particular, it has the form

$$T_{a,b}(x, y) \;=\; (F_a(x, y) + b u_{a,b}(x, y),\; b v_{a,b}(x, y)).$$

Where no ambiguity arises, we will write $T = T_{a,b}$. The phrase "for (a, b) sufficiently near $(a^*, 0)$" will appear (finitely) many times in the next few sections. Each time it appears, the rectangle in parameter space for which our results apply may have to be reduced. From here on K is the generic system constant as declared in Section 1.

2.5 Dynamics outside of $\mathcal{C}^{(0)}$

The first system constant to be chosen is δ. A number of upper bounds for δ will be specified as we go along. For now we think of it as a very small positive number with $d(f^n\hat{x}, C) >> \delta$ for all $\hat{x} \in C$ and $n > 0$. We assume also that a is sufficiently near a^* that the Hausdorff distances between the critical sets of f_{a^*} and f_a are $<< \delta$.

Recall that we will be working in $R_0 = \{(x, y) \in A : |y| \leq Kb\}$. Our zeroth critical region $\mathcal{C}^{(0)}$ is defined to be

$$\mathcal{C}^{(0)} = \{(x, y) \in R_0 : |x - \hat{x}| < \delta \quad \text{for some } \hat{x} \in C\}.$$

Let $s(u)$ denote the slope of a vector u. Assuming that $b^{\frac{1}{4}} << \delta$, an easy calculation shows that for $z \notin \mathcal{C}^{(0)}$, if $|s(u)| < \frac{b}{\delta^4}$, then $|s(DT(z)u)| = \mathcal{O}(\frac{b}{\delta})$. Also, if $\kappa_0 := \min \|DT(z)u\|$ where the minimum is taken over all $z \notin \mathcal{C}^{(0)}$ and unit vectors u with $|s(u)| < \frac{b}{\delta^4}$, then $\kappa_0 > K^{-1}\delta$. Let $K(\delta) := \frac{K}{\kappa_0^3}$, so that $K(\delta)b$ is the upper bound for k_n in Lemma 2.4. We call a vector u a **b-horizontal vector** if $|s(u)| < K(\delta)b$. A curve γ is called a **$C^2(b)$-curve** if its tangent vectors are b-horizontal and its curvature is $\leq K(\delta)b$ at every point.

Lemma 2.7 *(a) For $z \notin \mathcal{C}^{(0)}$, if u is b-horizontal, then so is $DT(z)u$.*

(b) If γ is a $C^2(b)$-curve outside of $\mathcal{C}^{(0)}$, then $T(\gamma)$ is again a $C^2(b)$-curve.

Proof: (a) has already been explained; (b) is an immediate consequence of (a) and Lemma 2.4. □

Our next lemma describes the dynamics of b-horizontal vectors outside of $\mathcal{C}^{(0)}$.

Lemma 2.8 *There exist constants $c_0, c_1 > 0$ independent of δ such that the following holds for $T = T_{a,b}$ for all (a, b) sufficiently near $(a^*, 0)$. Let $z \in R_0$ be such that $z, Tz, \cdots, T^{n-1}z \notin \mathcal{C}^{(0)}$, and let u be a b-horizontal vector. Then*

(i) $\|DT^n(z)u\| \geq c_0\delta e^{c_1 n}$;

(ii) if, in addition, $T^n z \in \mathcal{C}^{(0)}$, then $\|DT^n(z)u\| \geq c_0 e^{c_1 n}$.

Proof: As with Corollary 2.3, this follows from Lemma 2.5 by perturbation. □

2.6 Critical points inside $\mathcal{C}^{(0)}$

Wherever it makes sense, let e_m denote the field of most contracted directions of DT^m and let q_m be the slope of e_m. When working with a curve γ parameterized by arc length, we write $q_m(s) = q_m(\gamma(s))$. We begin with some easy observations about e_1.

Lemma 2.9 *For all (a, b) sufficiently near $(a^*, 0)$, e_1 is defined everywhere on R_0, and there exists $K > 0$ such that*

(a) $|q_1| > K^{-1}\delta$ outside of $\mathcal{C}^{(0)}$, and q_1 has opposite signs on adjacent components of $R_0 \setminus \mathcal{C}^{(0)}$;

(b)

$$|\frac{dq_1}{ds}| > K^{-1}$$

on every $C^2(b)$-curve γ in $\mathcal{C}^{(0)}$.

Proof: The existence of e_1 follows from the fact that everywhere on R_0, $\|DT\| > K^{-1}$ (this uses the non-degeneracy condition in Step IV, Sect. 1.1) while $|\det(DT)| = \mathcal{O}(b)$. For $a = a^*$, $b = 0$ and $\{y = 0\}$, the assertion in (a) is obvious, and part (a) of Lemma 2.9 follows by a perturbative argument. The estimate for $|\frac{dq_1}{ds}|$ uses the non-degeneracy condition above and the fact that $f''_{a^*} \neq 0$ on C. See Appendix B.5 for details. □

Definition 2.1 *Let γ be a $C^2(b)$-curve in $\mathcal{C}^{(0)}$. We say that z_0 is a* **critical point of order** m **on** γ *if*

(a) $\|DT^i(z_0)\| \geq K^{-1}$ for $i = 1, 2, \cdots, m$;

(b) at z_0, e_m coincides with the tangent vector to γ.

It follows from Lemma 2.9 that on every $C^2(b)$-curve that stretches across a component of $\mathcal{C}^{(0)}$, there is a unique critical point of order 1. The next two lemmas are used in the "updating" of existing critical points and the creation of new ones. Their proofs are given in Appendix B.5

Lemma 2.10 ([BC2], p. 113) *Let γ be a $C^2(b)$-curve in $\mathcal{C}^{(0)}$ where $\gamma(0) = z$ is a critical point of order m. We assume that*

(a) $\|DT^i(z)\| \geq 1$ for $i = 1, 2, \cdots, 3m$;

(b) $\gamma(s)$ is defined for $s \in [-(Kb)^{\frac{m}{2}}, (Kb)^{\frac{m}{2}}]$.

Then there exists a unique critical point $\hat{z}$ of order $3m$ on γ, and $|\hat{z} - z| < (Kb)^m$.

Lemma 2.11 ([BC2], Lemma 6.1) *For $\varepsilon > 0$, let γ and $\hat{\gamma}$ be two disjoint $C^2(b)$-curves in $\mathcal{C}^{(0)}$ defined for $s \in [-4K_1\sqrt{\varepsilon}, 4K_1\sqrt{\varepsilon}]$ where K_1 is the constant K in Lemma 2.9(b). We assume*

(a) $\gamma(0)$ is a critical point of order m;

(b) the x-coordinates of $\gamma(0)$ and $\hat{\gamma}(0)$ coincide, and $|\gamma(0) - \hat{\gamma}(0)| < \varepsilon$.

Then there exists a critical point of order $\hat{m}$ at $\hat{\gamma}(\hat{s})$ with $|\hat{s}| < 4K_1\sqrt{\varepsilon}$ and $\hat{m} = min\{m, K\log\frac{1}{\varepsilon}\}$.

2.7 Tracking DT^n: a splitting algorithm

The purpose of this section is to recall an algorithm introduced in [BC2] that gives, under suitable circumstances, a direct relation between DT^n and 1-dimensional derivatives.

Let $z_0 \in R_0$, and let w_0 be a unit vector at z_0 that is b-horizontal. We write $z_n = T^n z_0$ and $w_n = DT^n(z_0)w_0$. In the case where $z_i \notin \mathcal{C}^{(0)}$ for all i, the resemblance to 1-d is made clear in Lemmas 2.5 and 2.8. Consider next an orbit $z_0, z_1, \cdots$ that visits $\mathcal{C}^{(0)}$ exactly once, say at time $t > 0$. Assume:

(a) there exists $\ell > 0$ such that $\|DT^i(z_t)\binom{0}{1}\| \geq 1$ for all $i < \ell$, so that in particular e_ℓ, the most contracted direction of DT^ℓ, is defined at z_t, and
(b) $\theta(w_t, e_\ell)$, the angle between w_t and e_ℓ, is $\geq b^{\frac{\ell}{2}}$.

Then $DT^i(z_0)$ can be analyzed as follows. (Note that our notation is different from that in [BC2].) We split w_t into $w_t = \hat{w}_t + \hat{E}$ where $\hat{w}_t$ is parallel to the vector $\binom{0}{1}$ and $\hat{E}$ is parallel to e_ℓ. For $i \leq t$ and $i \geq t + \ell$, let $w_i^* = w_i$. For i with $t < i < t+\ell$, let $w_i^* = DT^{i-t}(z_t)\hat{w}_t$. We claim that all the w_i^* are b-horizontal vectors, so that $\{\|w_{i+1}^*\|/\|w_i^*\|\}_{i=0,1,2,\ldots}$ resemble a sequence of 1-d derivatives. In particular, $\|w_{t+1}^*\|/\|w_t^*\| \sim \theta(w_t, e_\ell)$ simulates a drop in the derivative when an orbit comes near a critical point in 1-dimension.

To justify the statement about the slope of the w_i^*, we note that $DT(z_t)\binom{0}{1}$ is b-horizontal, so that in view of lemma 2.7 we need only to consider $w_{t+\ell}^*$. We have

$$\|DT^\ell(\hat{E})\| \leq b^\ell \frac{\|\hat{w}_t\|}{\theta(w_t, e_\ell)} \leq b^{\frac{\ell}{2}} \|\hat{w}_t\| \leq b^{\frac{\ell}{2}} \|DT^\ell(z_t)\hat{w}_t\|,$$

the first and third inequalities following from (a) and the second from (b). Since the slope of $DT^\ell(z_t)\hat{w}_t$ is smaller than $\frac{Kb}{2\delta}$, it follows that $w_{t+\ell}^* = DT^\ell(z_t)\hat{w}_t + DT^\ell(z_t)\hat{E}$ remains b-horizontal.

The discussion above motivates the following splitting algorithm introduced in [BC2]. Consider $\{z_i\}_{i=0}^\infty$, and let $t_1 < \cdots < t_j < \cdots$ be the times when $z_i \in \mathcal{C}^{(0)}$. We let w_0 be a b-horizontal unit vector, and assume as before that e_{ℓ_i} makes sense at z_i for $i = t_j$. Define w_i^* as follows:

1. For $0 \leq i \leq t_1$, let $w_i^* = DT^i(z_0)w_0$.
2. At $i = t_j$, we split w_i^* into

$$w_i^* = \hat{w}_i + \hat{E}_i$$

where $\hat{w}_i$ is parallel to $\binom{0}{1}$ and $\hat{E}_i$ is parallel to e_{ℓ_i}.

3. For $i > t_1$, let

$$w_i^* = DT(z_{i-1})\hat{w}_{i-1} \;+ \sum_{j:\ t_j+\ell_{t_j}=i} DT^{\ell_{t_j}}(z_{t_j})\hat{E}_{t_j} \tag{5}$$

and let $\hat{w}_i = w_i^*$ if $i \neq t_j$ for any j.

This algorithm does not give anything meaningful in general. It does, however, in the scenario of the next lemma.

Lemma 2.12 *Let z_i, w_i and w_i^* be as above. Assume*
(a) for each $i = t_j$, $\theta(w_i^*, e_{\ell_i}) \geq b^{\frac{\ell_i}{2}}$;

(b) the time intervals $I_j := [t_j, t_j + \ell_{t_j}]$ *are strictly nested, i.e. for* $j \neq j'$, *either* $I_j \cap I_{j'} = \emptyset$, $I_j \subset I_{j'}$, *or* $I_{j'} \subset I_j$, *and* $t_j + \ell_{t_j} \neq t_{j'} + \ell_{t_{j'}}$.

Then $w_i = w_i^*$ *for* $i \notin \cup_j I_j$, *and the* w_i^**'s are all b-horizontal vectors. The sequence* $\{|w_i^*\|\}$ *has the property that* $\|w_{i+1}^*\|/\|w_i^*\| \sim \theta(w_i^*, e_{\ell_i})$ *for* $i = t_j$, *and* $\|w_{i+1}^*\| \approx \|DT(z_i)w_i^*\|$ *for* $i \neq t_j$.

Proof: The nested condition in (b) allows us to consider the I_j's one at a time beginning with the innermost time intervals. This reduces to the case of a single visit to $\mathcal{C}^{(0)}$ treated earlier on. □

3 The Critical Set

Many authors, including [BC1], [CE], [J], [M1], and [NS], have studied 1-dimensional maps by controlling their critical orbits. These ideas were mimicked in [BC2] where the authors developed techniques for identifying, for certain Hénon maps, a set they called the "critical set". This is done via an inductive procedure involving parameter selection. The first step in our analysis of the family $\{T_{a,b}\}$ is to carry out a similar parameter selection, and the aim of this section is to formulate suitable inductive hypotheses.

3.1 What is the critical set?

In 1-dimension, the critical set is where all previous expansion is destroyed. Tangencies of stable and unstable manifolds play a similar role in higher dimensions. Here is how we propose to capture the set $\mathcal{C}$ that we will prove in Section 7 to be the *origin* of all *nonhyperbolic behavior*.

Let $\mathcal{F}_0$ be the foliation on R_0 with leaves $\{y = \text{constant}\}$, and let $\mathcal{F}_k$ be its image under T^k. In Sect. 2.5 we defined the 0th critical region $\mathcal{C}^{(0)}$. Suppose that $T^i\mathcal{C}^{(0)} \cap \mathcal{C}^{(0)} = \emptyset$ for all $i \leq i_0$. Then for $i \leq i_0$, $\mathcal{F}_i$ restricted to $\mathcal{C}^{(0)} \cap R_i$ consists of finitely many bands of roughly horizontal leaves whose tangent vectors have been expanded the previous i iterates (Lemma 2.8). From Corollaries 2.1, 2.2 and Lemma 2.9, we see also that in $\mathcal{C}^{(0)}$, DT^i has a well-defined field of most contracted directions, namely e_i, whose integral curves are roughly parabolas. It is natural to take the set of tangencies in $\mathcal{C}^{(0)} \cap R_i$ between the leaves of $\mathcal{F}_i$ and the integral curves of e_i to be our ith approximation of $\mathcal{C}$. Since these approximations stabilize quickly with i, they would converge to $\mathcal{C}$ if this picture could be maintained indefinitely, i.e. if the "turns" of $\mathcal{F}_i$ could be prevented from entering $\mathcal{C}^{(0)}$ for all i.

For $i \leq i_0$, we think of $\mathcal{C}^{(i)} := \mathcal{C}^{(0)} \cap R_i$ as our ith *critical region*. The strategy as explained above, then, is essentially to solve for tangencies of *temporary stable* *and unstable manifolds* in $\mathcal{C}^{(i)}$ and call the resulting set our ith approximation of $\mathcal{C}$. Observe that $\mathcal{C}^{(i)}$ is the union of at most K^i rectangles with a transparent geometry. This geometry will be passed on to the critical set.

Now experience from 1-dimension tells us that in order to retain a positive measure set of parameters, we must allow our "turns" to approach the critical set as i increases. We will allow them to return *slowly*, and to maintain a picture similar to that for $i \leq i_0$, we will shrink the critical regions $\mathcal{C}^{(i)}$ *sideways* at a rate faster than this rate of approach. Justification is needed to show that this process can be continued indefinitely and to prove the stabilization of the approximate critical sets. In the end, an alternate characterization of $\mathcal{C}$ will be $\mathcal{C} = \bigcap_{i\geq 0} \mathcal{C}^{(i)}$.

In order for the contractive fields above to be defined, it is necessary that the derivative along orbits starting from $\mathcal{C}$ experience some exponential growth. This growth, which is also useful for controlling the movements of the "turns", is brought about in two ways: (i) by arranging for critical orbits to stay away from the critical set for a very long time, hyperbolicity is guaranteed for a long initial period; (ii) when an orbit of $\mathcal{C}$ gets near a point $z \in \mathcal{C}$, it copies the initial segment of the orbit of z, thereby *replicating* the growth properties created in (i).

A version of these ideas will be made precise in the inductive assumptions.

3.2 Getting started

We first introduce our main system constants. They are α, β, ρ, c, n_0, θ and δ (which we have already met):

- $e^{-\alpha n}$ and $e^{-\beta n}$, with $\alpha << \beta << 1$, represent two small length scales.

- $c > 0$ is our target Lyapunov exponent; it is $< c_1$ where c_1 is as in Lemma 2.8.

- $0 < \rho < K^{-1}$ is an arbitrary number of order 1. It determines the rate at which our critical regions decrease in size (see Sect. 3.1).

- n_0 is the number of iterates the critical orbits are required to stay a preassigned distance away from $\mathcal{C}$; see below for more precise specifications.

- θ is chosen so that b^θ is a number of order 1 and $< K^{-1}$; one use of θ is the following: critical orbits originating from the same component of $\mathcal{C}^{([\theta N])}$ are indistinguishable in their first N iterates. For this reason, critical points of generation $> \theta N$ are not constructed in the first N steps of the induction (see (IA1) below).

These constants are chosen in the following order: c and ρ are determined by the derivative of T; α and β are then chosen. This is followed by δ, which is $<< \delta_0$ to start with and shrunk a number of times as needed in the course of our argument.

The value of n_0 is not determined until very late in the proof: it is used to ensure sufficient hyperbolicity at the start (and to overcome various "irregularities" that occur at initial stages); it depends on all the other system constants except for b. Observe that increasing n_0 is at the expense of shrinking the size of the parameter set at the start. The magnitude of b, which is used to beat everything, is the last to be chosen; θ as we have defined it is, of course, determined by b.

At the start of our induction, we assume we have a parameter set Δ_0 with the following properties: Let f be the Misiurewicz map from which we are perturbing, and let $\delta_0 = \frac{1}{4} \inf \{d(f^n x, C), x \in C, n > 0\}$. First, by considering a sufficiently near a^*, we may assume that for all a and for every critical point x of f_a, $d(f_a^n x, C) \geq 2\delta_0$ for all $0 < n \leq n_0$. Next, by choosing b sufficiently small, we may assume, through Corollaries 2.1, 2.2 and Lemmas 2.9, 2.10, that $T_{a,b}$ has on each connected segment of $\partial R_0 \cap \mathcal{C}^{(0)}$ a unique critical point z_0 of order n_0, and that z_0 is close enough to the corresponding critical point of f_a that $d_{\mathcal{C}}(z_n) \geq \delta_0$ for all $n \leq n_0$. These are our **critical points of generation** 0. They comprise the set we call Γ_0.

Parameters are deleted at each stage of our induction. **Sections 3 –5 are concerned with the dynamics of the maps corresponding to the parameters retained. Issues pertaining to the measure of the set of retained parameters** (including whether or not it is nonempty) **are postponed to Section 6.**

3.3 Inductive assumptions

Let $N \geq n_0$ be a large number, and let Δ_N be the set of parameters retained after N iterates. We now formulate a set of inductive assumptions that describes the desired dynamical picture for $T = T_{a,b}$, $(a,b) \in \Delta_N$. While we will continue to provide motivations and explanations, (IA1)–(IA6) below are to be viewed as formal inductive hypotheses. As before, let $z_i = T^i z_0$.

3.3.1 Critical points and critical regions

(IA1) *For all $k \leq \theta N$, the critical regions $\mathcal{C}^{(k)}$ are defined and have the geometric properties stated in* (1)(i), (ii) *and* (iii) *of Theorem 1. Moreover, on each horizontal boundary of each component of $\mathcal{C}^{(k)}$, there is a unique critical point of order N located within $\mathcal{O}(b^{\frac{k}{3}})$ of the midpoint of the segment.*

Critical points of order N on $\partial\mathcal{C}^{(k)}$ are called **critical points of generation** k **and order** N. The set of critical points of generation $\leq k$ is denoted by Γ_k. As the induction progresses, the orders of the critical points are updated, and the precise locations of Γ_k are modified accordingly. At the end of the induction process, $\Gamma := \cup_k \Gamma_k$, where Γ_k now refers to the set of critical points of generation k and order ∞, is the set in the statement of part (2) of Theorem 1.

3.3.2 Distance to critical set and loss of hyperbolicity

If the critical set is where would-be stable and unstable directions are interchanged, then distance to the critical set might provide a measure of loss of hyperbolicity. This is indeed the case under suitable circumstances and for a suitable notion of "distance".

If Q is a component of $\mathcal{C}^{(k)}$, we let L_Q denote the vertical line midway between the two vertical boundaries of Q.

Definition 3.1 *We say $z \in \mathcal{C}^{(0)}$ is* **horizontally related** *or simply* **h-related** *to $\Gamma_{\theta N}$ if there exists a component Q of $\mathcal{C}^{(k)}$, $k \leq \theta N$, such that $z \in Q$ and $dist(z, L_Q) \geq b^{\frac{k}{20}}$. When this holds, we say z is h-related to z_0 for all $z_0 \in \Gamma_{\theta N} \cap Q$.* [7]

This is an attempt to describe the location of a point relative to $\Gamma_{\theta N}$, which, as $N \to \infty$, converges to a fractal set. From Lemma 4.1, we see that $\Gamma_{\theta N} \cap Q$ is contained in a region of width $\mathcal{O}(b^{\frac{k}{4}})$ in the middle of Q, so that z and $\Gamma_{\theta N} \cap Q$ have a very obviously horizontal relationship. We caution, however, that there may be points in $\Gamma_{\theta N}$ that are directly above or below z, and quite possibly both to its left and to its right. Observe also that if Q' is a component of $\mathcal{C}^{(k')}$ such that $z \in Q' \subset Q$, then $dist(z, L_{Q'}) \geq b^{\frac{k'}{20}}$.

Definition 3.2 *For $z \in R_0$, we define its* **distance to the critical set***, denoted $d_{\mathcal{C}}(z)$, as follows: for $z \in \mathcal{C}^{(0)}$, we let $d_{\mathcal{C}}(z) = dist(z, L_Q)$ where Q is the component of $\mathcal{C}^{(k)}$ containing z and k is the largest number $\leq \theta N$ with $z \in \mathcal{C}^{(k)}$; for $z \notin \mathcal{C}^{(0)}$, let Q be the component of $\mathcal{C}^{(0)}$ nearest to z. We further let $\phi(z)$ be one of the two points in $\partial Q \cap \Gamma_{\theta N}$ if z is h-related to $\Gamma_{\theta N}$.*

For $z \in \mathcal{C}^{([\theta N])}$, the definitions of $d_{\mathcal{C}}(z)$ and $\phi(z)$ are temporary and will be modified as the induction progresses. We remark that for z in an h-related position, the distance from z to $\phi(z)$ is a very good approximation of $d_{\mathcal{C}}(z)$.

To secure growth properties for the orbits of $\Gamma_{\theta N}$, we forbid them to approach the critical set too closely too soon. (IA2) is a result of parameter selection.

(IA2) *For all $z_0 \in \Gamma_{\theta N}$ and all $i \leq N$, $d_{\mathcal{C}}(z_i) \geq min(\delta, e^{-\alpha i})$.*

We will assume, for convenience, that $e^{-\alpha n_0} < \delta$. Under this assumption, (IA2) reads $d_{\mathcal{C}}(z_i) \geq e^{-\alpha i}$ for $i > n_0$.

(IA2) implies that for all $z_0 \in \Gamma_{\theta N}$ and $i \leq N$, z_i is h-related to $\Gamma_{\theta N}$ whenever it is in $\mathcal{C}^{(0)}$. Intuitively, this is because z_i is in a very "deep" layer relative to its distance to $\Gamma_{\theta N}$. Formally, let $z_i \in Q \subset \mathcal{C}^{(k)}$ where Q and k are as in Definition 3.2. Then $k << i$ since $\rho^k \geq e^{-\alpha i}$. Now $z_i \in R_i$. If $k < [\theta N]$, then $z_i \in Q \cap R_{k+1}$, proving $d_{\mathcal{C}}(z_i) \geq \rho^{k+1} >> b^{\frac{k}{20}}$. If $k = [\theta N]$, then $d_{\mathcal{C}}(z_i) \geq e^{-\alpha i} \geq e^{-\alpha N} >> b^{\frac{1}{20}\theta N}$ provided that b^θ is chosen to be $< e^{-20\alpha}$.

[7]When studying the dynamics of T on ∂R_k, it will be convenient to include the following in the definition of h-relatedness: Let γ be a horizontal boundary of a component of $\mathcal{C}^{(k)}$, $k \leq \theta N$, and let $\hat{z} \in \gamma \cap \Gamma_{\theta N}$. Then $z \in \gamma$ is also said to be h-related to $\hat{z}$.

Definition 3.3 *(a) For arbitrary $z \in \mathcal{C}^{(0)}$, we define its* **fold period** *$\ell(z)$ to be the nonnegative integer $\ell \geq 1$ such that $b^{\frac{\ell}{2}}$ is closest to $d_{\mathcal{C}}(z)$.*

(b) Given $z_0 \in R_0$ and unit vector w_0, we let w_i^*, $i = 0, 1, 2, \cdots$, be given by the splitting algorithm in Sect. 2.7 with $\ell_i = \ell(z_i)$ assuming $e_{\ell(z_i)}$ is defined at z_i.

For $\ell \leq N$, Lemma 2.2 gives an estimate on the size of the neighborhood of $\Gamma_{\theta N}$ on which e_ℓ is well defined. In particular, if z is h-related to $\Gamma_{\theta N}$, then $e_{\ell(z)}$ is defined at z.

Recall that q_1 is the slope of e_1. We fix $\varepsilon_0 > 0$ such that $\varepsilon_0 << |\frac{\partial q_1}{\partial x}|$ in $\mathcal{C}^{(0)}$. For $z \in \partial R_k$, let $\tau(z)$ denote a unit tangent vector to ∂R_k at z. In the angle estimates below, τ and e_ℓ are assumed to point in roughly the same direction as w.

Definition 3.4 *Let $z \in \mathcal{C}^{(0)}$ be h-related to $\Gamma_{\theta N}$, and let w be a vector at z. We say w* **splits correctly** *if $|\frac{w}{\|w\|} - \tau(\phi(z))| < \varepsilon_0 d_{\mathcal{C}}(z)$.*

(IA3) *For $z_0 \in \Gamma_{\theta N}, w_0 = \binom{0}{1}$ and $i \leq N$, w_i^* splits correctly whenever $z_i \in \mathcal{C}^{(0)}$.*

The sense in which this splitting is "correct" is as follows. We wish to use Lemma 2.12 to understand the evolution of w_i, and (IA3) implies condition (a) of the lemma. This is because $|e_{\ell_i}(z_i) - \frac{w_i^*}{\|w_i^*\|}| \geq |e_{\ell_i}(z_i) - e_{\ell_i}(\phi(z_i))| - |e_{\ell_i}(\phi(z_i)) - \tau(\phi(z_i))| - |\tau(\phi(z_i)) - \frac{w_i^*}{\|w_i^*\|}| \geq |\frac{\partial q_{\ell_i}}{\partial x}| d_{\mathcal{C}}(z_i) - \mathcal{O}(b^{\ell_i}) - \varepsilon_0 d_{\mathcal{C}}(z_i) \geq \frac{1}{2}|\frac{\partial q_1}{\partial x}| d_{\mathcal{C}}(z_i) \sim b^{\frac{\ell_i}{2}}$. Condition (b) of Lemma 2.12 is discussed in Sect. 4.1.

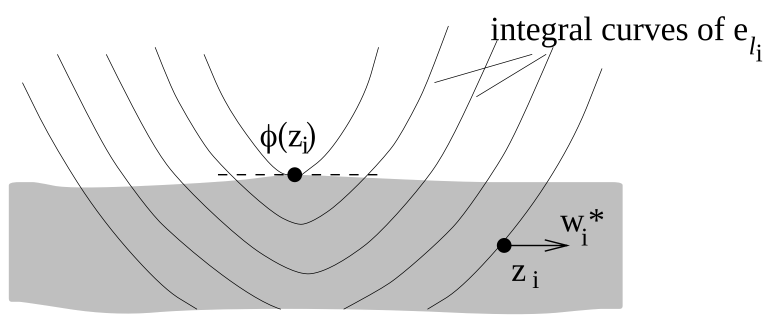


Figure 2 Correct splitting of w_i^*

3.3.3 Derivative along critical orbits

We saw in the last paragraph that for $z_0 \in \Gamma_{\theta N}$, as z_i enters $\mathcal{C}^{(0)}$, w_i^* suffers a loss of hyperbolicity proportional to $d_{\mathcal{C}}(z_i)$. Combining this with (IA5)(c) below applied to an earlier step, we see that this loss will be partially – but not fully – compensated for at the end of a certain period. To prevent a downward spiral in Lyapunov exponent, further parameter exclusion is needed.

(IA4) *For all $z_0 \in \Gamma_{\theta N}$ and $0 \leq i \leq N$, $\|w_i^*(z_0)\| > c_0 e^{ci}$.*

In future steps of the induction, orbits of length N starting from $\Gamma_{\theta N}$ will be replicated; in other words, they will serve as guides for other points that enter $\mathcal{C}^{(0)}$.

Definition 3.5 *For arbitrary ξ_0 and $\xi_0' \in \mathcal{C}^{(0)}$, we define their* **bound period** *to be the largest integer p such that for all $0 < j \leq p$,*

$$|\xi_j - \xi_j'| \;\leq\; e^{-\beta j}.$$

Consider the situation where $\xi_0' = z_0 \in \Gamma_{\theta N}$. An important observation is that for $j \leq p$, $|\xi_j - z_j| \;<< d_{\mathcal{C}}(z_j)$. Observe also that by taking δ small enough, we have $d_{\mathcal{C}}(\xi_j) > \frac{1}{2}\delta_0$ for all $j \leq$ min (p, n_0) independent of n_0. (To achieve this, choose n_1 with $e^{-\beta n_1} < \frac{1}{2}\delta_0$, and require $K\delta^2 \|DT\|^{n_1} < e^{-\beta n_1}$). Taking n_0 large also ensures that $d_{\mathcal{C}}(\xi_j) > \frac{\delta}{2}$ whenever z_j is outside of $\mathcal{C}^{(0)}$.

Our last two inductive assumptions deal with the properties z_0 passes along to ξ_0.

(IA5) *Let $z_0 \in \Gamma_{\theta N} \cap \partial \mathcal{C}^{(k)}$, and let $\gamma : [0, \varepsilon] \to \mathcal{C}^{(0)}$ be a $C^2(b)$-curve with $\gamma(0) = z_0$ and $\gamma'(0)$ tangent to $\partial \mathcal{C}^{(k)}$. We regard all $\xi_0 \in \gamma$ as bound to z_0, and let $p(\xi_0)$ denote their bound periods. Then:*

(a) There exists K such that for $\xi_0 \in \gamma$ with $|\xi_0 - z_0| = e^{-h}$,

$$\frac{1}{K}h \leq p(\xi_0) \leq Kh \quad \textit{provided} \quad Kh < N;$$

moreover, $p(\xi_0)$ increases monotonically with the distance between ξ_0 and z_0;

(b) for $\ell \leq j \leq min(p, N)$, $|\xi_j - z_j| \approx |\xi_0 - z_0|^2 \|w_j(z_0)\|$ where "$\approx$" means up to a factor of $(1 \pm \varepsilon_1)$ for some $\varepsilon_1 > 0$;

(c) $\|w_p(\xi_0)\| \cdot |\xi_0 - z_0| \geq e^{\frac{cp}{3}}$ provided $p < N$.

(IA5) describes the quadratic nature of the "turn" as γ is mapped forward. For comparison with 1-dimensional behavior, see Lemma 2.6.

The following distortion estimates are used in the proof of (IA5). Let $w_0(\xi_0) = w_0(z_0) = \binom{0}{1}$, and let $\hat{w}_i^*(\xi_0)$ be given by Definition 3.3(b) except that $e_{\ell(z_i)}$ (and not $e_{\ell(\xi_i)}$) is used for splitting at time i. (IA6) compares $w_i^*(z_0)$ and $\hat{w}_i^*(\xi_0)$. Let $M_i(\cdot)$ and $\theta_i(\cdot)$ denote the magnitude and argument of the vectors in question. Define

$$\Delta_i(\xi_0, z_0) = \sum_{s=0}^{i} (Kb)^{\frac{s}{4}} \mid \xi_{i-s} - z_{i-s} \mid . \tag{6}$$

(IA6) *Given $z_0 \in \Gamma_{\theta N}$ and any $\xi_0 \in \mathcal{C}^{(0)}$, we regard ξ_0 as bound to z_0 and let p be the bound period. Then for $i \leq min\{p, N\}$,*

$$\frac{M_i(z_0)}{M_i(\xi_0)},\;\; \frac{M_i(\xi_0)}{M_i(z_0)} \;\leq\; \exp\{K \sum_{j=1}^{i-1} \frac{\Delta_j}{d_{\mathcal{C}}(z_j)}\} \tag{7}$$

and

$$|\ \theta_i(\xi_0) - \theta_i(z_0)\ | \leq (Kb)^{\frac{1}{2}} \Delta_{i-1}. \tag{8}$$

The estimates above also hold with $w_i^*(z_0)$ replaced by $\hat{w}_i^*(\xi_0')$ where ξ_0' is another point in $\mathcal{C}^{(0)}$ also thought of as bound to z_0, and p is the minimum of the two bound periods.

We remark that the right side of (7) is finite and can be made arbitrarily close to 1 by choosing δ small (see Appendix B.7).

Let us return for a moment to Definition 3.1. From the geometry of $\mathcal{C}^{(k)}$ (see (IA1) and Lemma 4.1) it is an exercise in calculus to show that if ξ_0 is h-related to $z_0 \in \Gamma_{\theta N}$, then it lies on a $C^2(b)$-curve through z_0 tangent to $\tau(z_0)$. In particular, (IA5) applies.

Our rules of parameter exclusion, namely (IA2) and (IA4), are similar to those used in [BC2], but they are applied to different orbits and with a different definition of "$d_{\mathcal{C}}(\cdot)$". The notions of bound and fold periods are borrowed from [BC2], as are (IA5) and (IA6). Our construction of $\mathcal{C}$, however, has a distinctly different flavor.

4 Replication of Orbit Segments

In Sect. 3.1 we outlined a scheme for obtaining derivative growth along critical orbits, namely to choose a start-up geometry that guarantees some initial growth, and then to try to replicate this behavior. Section 4 contains a detailed analysis of the replication process. The main results are stated in Sect. 4.3, after some technical preparations in Sects. 4.1 and 4.2, including amending slightly the definitions of bound and fold periods. Throughout Section 4, (IA1)–(IA6) are assumed up to time N.

4.1 Nested properties of bound and fold periods

Consider $z_0 \in \Gamma_{\theta N}$. When z_i enters $\mathcal{C}^{(0)}$, it is natural to assign to it a **bound period** $p(z_i)$ defined using $\phi(z_i)$. An unsatisfactory aspect of this definition is that two bound periods so defined may overlap without one being completely contained in the other. The purpose of this subsection is to adjust slightly the definition of $p(z_i)$ to create a simpler binding structure. A similar adjustment is made in [BC2].

First we fix some notation. Let $Q^{(j)}$ denote the components of $\mathcal{C}^{(j)}$, and let $\hat{Q}^{(j)}$ be the component of $R_j \cap \mathcal{C}^{(j-1)}$ containing $Q^{(j)}$. For $z \in \partial R_j$, let $\tau(z)$ be a unit vector at z tangent to ∂R_j.

Lemma 4.1 *For $z, z' \in \Gamma_{\theta N} \cap Q^{(k)}$, we have*

$$|z - z'| = \mathcal{O}(b^{\frac{k}{4}}) \quad \text{and} \quad \|\tau(z) \times \tau(z')\| = \mathcal{O}(b^{\frac{k}{4}}).$$

Proof: Let $z^{(k)}$ be a critical point in $\partial Q^{(k)}$. For $k \le i < [\theta N]$, let $z^{(i+1)}$ be a critical point of generation $i+1$ in $Q^{(i)}(z^{(i)})$, the component of $Q^{(i)}$ containing $z^{(i)}$. From (IA1) we know that the Hausdorff distance between the two horizontal boundaries of $Q^{(i)}(z^{(i)})$ is $\mathcal{O}(b^{\frac{i}{2}})$. Lemma 2.11 then tells us that $|z^{(i)} - z^{(i+1)}| = \mathcal{O}(b^{\frac{i}{4}})$. The angle estimate also follows from the proof of Lemma 2.11 □

Lemma 4.2 *Let ξ_0 be h-related to $z_0 \in \Gamma_{\theta N}$. If during their bound period z_i returns to $\mathcal{C}^{(k)}$, then $\xi_i \in \hat{Q}^{(k)}(z_i)$.*

Proof: Let γ be a $C^2(b)$-curve joining z_0 and ξ_0. Then $T^i\gamma \subset R_i$. Since $e^{-\alpha i} \le d_{\mathcal{C}}(z_i) \le \rho^k$, we have $k < i$ and therefore $T^i\gamma \subset R_k$. By the monotonicity of bound periods, every point in $T^i\gamma$ is within a distance of $< e^{-\beta i}$ from z_i. This puts $\xi_i \in R_k \cap Q^{(k-1)}(z_i)$. □

Lemma 4.3 *Let $z_0 \in \Gamma_{\theta N}$ be such that $z_i \in \mathcal{C}^{(0)}$ at times $t_1 < t_2 < \cdots < t_r$, and that for each $j < r$ the bound period p_j initiated at time t_j extends beyond time t_{j+1}. Then $p_j < (K\alpha)^{j-1}p_1$.*

Proof: Let $\tilde{z}_0 = \phi(z_{t_1})$. We claim that $|z_{t_2} - \phi(z_{t_2})| \approx |\tilde{z}_{t_2-t_1} - \phi(\tilde{z}_{t_2-t_1})|$, which is $> e^{-\alpha(t_2-t_1)}$. If true, this will imply, by (IA5)(a), that $p_2 < K\alpha(t_2-t_1) < K\alpha p_1$, and the assertion in the lemma will follow inductively. Since $|z_{t_2} - \tilde{z}_{t_2-t_1}| < e^{-\beta(t_2-t_1)} << e^{-\alpha(t_2-t_1)}$, it suffices to show that $|\phi(\tilde{z}_{t_2-t_1}) - \phi(z_{t_2})| << |\tilde{z}_{t_2-t_1} - \phi(\tilde{z}_{t_2-t_1})|$. Let k be the largest number such that $\tilde{z}_{t_2-t_1} \in \mathcal{C}^{(k)}$. By Lemma 4.2, $z_{t_2} \in Q^{(k-1)}(\tilde{z}_{t_2-t_1})$, so $\phi(\tilde{z}_{t_2-t_1})$ and $\phi(z_{t_2})$ must both be in $Q^{(k-1)}(\tilde{z}_{t_2-t_1})$. By Lemma 4.1 they are $\le b^{\frac{k-1}{4}}$ apart, and this is $<< |\tilde{z}_{t_2-t_1} - \phi(\tilde{z}_{t_2-t_1})|$. □

Definition 4.1 *For $z_0 \in \Gamma_{\theta N}$ with $z_i \in \mathcal{C}^{(0)}$, the* **adjusted bound period** *$p^*(z_i)$ is defined to be the smallest number p^* with the property that for all j with $i \le j < i+p^*$, if $z_j \in \mathcal{C}^{(0)}$, then $j + p(z_j) \le i + p^*$.*

Adjusted bound periods, therefore, have a nested structure by definition.

Corollary 4.1 *(a) $p^* \le p + K\alpha p$.*

(b) For $z_i \in \mathcal{C}^{(0)}$ with $\phi(z_i) = \hat{z}_0$, we have for all $j \le p^*$,

$$|z_{j+i} - \hat{z}_j| < e^{-\beta^* j}$$

for some β^* smaller than β and $>> \alpha$.

The proof is left as an exercise. We assume from here on that all bound periods for all critical orbits are adjusted, and write p and β instead of p^* and β^*.

This amended definition gives critical orbits the following simple structure of **bound** and **free states**. We call z_i a **return** if $z_i \in \mathcal{C}^{(0)}$. Then z_i is free for $i \le n_1$ where $n_1 > 0$ is the time of the first return, and it is in bound state for $n_1 < i \le n_1+p_1$ where p_1 is the bound period initiated at time n_1. After time $n_1 + p_1$, z_i remains free until its next return at time n_2, is bound for the next p_2 iterates, and so on. The times n_j are called **free return** times. A **primary bound period** begins at each n_j. Inside the time interval $[n_j, n_j + p_j]$, there may be **secondary bound periods** which comprise disjoint time intervals, and so on.

Next we consider fold periods, which are denoted by ℓ and defined in Sect. 3.3.2. As with bound periods, if z_i enters $\mathcal{C}^{(0)}$ at times t_1 and t_2 with $t_1 < t_2 \le N$, and if the fold period begun at t_1 remains in effect at t_2, then using Lemma 4.2 we see that $\ell_{t_2} < \frac{\alpha}{\log \frac{1}{b}} \ell_{t_1}$, so that **adjusted fold periods** can be defined similarly to give a nested structure. This is condition (b) of Lemma 2.12 . A further simplifying arrangement, which we will also adopt, is that no fold periods expire at returns to $\mathcal{C}^{(0)}$ or at the step immediately after. The proof of the following lemma is straightforward and will be omitted.

Lemma 4.4 (cf. [BC2], Lemma 6.5) *Let $z_0 \in \Gamma_{\theta N}$. Then for every $i < N$, there exist $i_1 \le i \le i_2$ with*

$$i_2 - i_1 < K\theta\alpha i$$

such that i_1 and i_2 are out of all fold periods.

4.2 Orbits controlled by $\Gamma_{\theta N}$

In this subsection we consider (z_0, w_0) where z_0 is an arbitrary point in R_0 and w_0 is a unit vector. We write $z_i = T^i z_0$ and $w_i = DT^i(z_0)w_0$.

Definition 4.2 *We say (z_0, w_0) is* **controlled** *by $\Gamma_{\theta N}$ up to time m (with m possibly $> N$) if the following hold.*

- *Initial conditions: if $z_0 \notin \mathcal{C}^{(0)}$, then w_0 is a b-horizontal vector; if $z_0 \in \mathcal{C}^{(0)}$, then either $w_0 = \binom{0}{1}$, or z_0 is h-related to $\Gamma_{\theta N}$ and w_0 splits correctly.*
- *For $0 < i \le m$, if $z_i \in \mathcal{C}^{(0)}$, then z_i is h-related to $\Gamma_{\theta N}$ and w_i^* splits correctly in the sense of Definition 3.4 with ε_0 replaced by $2\varepsilon_0$.*

No h-relatedness property is required for $z_0 \in \mathcal{C}^{(0)}$ when $w_0 = \binom{0}{1}$ because for practical purposes, one may think of the sequence as starting with (z_1, w_1).

Let (z_0, w_0) be as above. Then the orbit of z_0 has a natural bound/free structure defined as follows: If $z_0 \in \Gamma_{\theta N}$, then it is natural to regard $z_0, z_1, \cdots, z_i$ as free until z_i returns to $\mathcal{C}^{(0)}$. For $z_0 \in \mathcal{C}^{(0)} \setminus \Gamma_{\theta N}$, we may regard z_0 as bound to any $\hat{z} \in \Gamma_{\theta N}$ for a period p provided that $(\max \|DT\|)^p \; |z_0 - \hat{z}| < e^{-\beta p}$. (This trivial bound period is used to ensure that Lemma 4.2 continues to work.) When z_i is h-related to $\Gamma_{\theta N}$, we take the bound period to be that between z_i and $\phi(z_i)$ (which is longer than the trivial one). Observe that Lemma 4.3 is equally valid for controlled orbits as for orbits starting from $\Gamma_{\theta N}$, so that a nested structure can also be assumed for the bound and fold periods of controlled orbits.

In the language of Definition 4.2, the situation can be summed up as follows. First, it follows from (IA2) and (IA3) that for all $\hat{z}_0 \in \Gamma_{\theta N}$, $(\hat{z}_0, \binom{0}{1})$ is controlled by $\Gamma_{\theta N}$ up to time N. (In fact, the angle of splitting is better than that in the definition of "control".) Second, for (z_0, w_0) controlled by $\Gamma_{\theta N}$, (IA5) and (IA6) apply to give information during its bound periods. In particular, the orbit of (z_0, w_0) has similar bound/free structures and "derivative recovery" estimates as those of $(\hat{z}_0, \binom{0}{1})$, $\hat{z}_0 \in \Gamma_{\theta N}$, except that (IA2) and (IA4) need not hold.

In the remainder of this subsection we record some basic facts on the growth of $\|w_i\|$ and $\|w_i^*\|$. Their proofs are given in Appendix B.6. In Lemmas 4.6–4.8, it is assumed that (z_0, w_0) is controlled by $\Gamma_{\theta N}$ up to time m, and all time indices are $\leq m$.

Lemma 4.5 *Suppose (z_0, w_0) satisfies the initial conditions in Definition 4.2, and for $0 < i \leq m$, z_i is h-related to $\Gamma_{\theta N}$ at all returns. Then (z_0, w_0) is controlled up to time m if the angle condition on w_i^* is satisfied at all free returns.*

Lemma 4.6 *Under the additional assumption that $d_{\mathcal{C}}(z_i) > e^{-\alpha i}$ for all $i \leq m$, we have*

$$K^{-\varepsilon i}\|w_i^*\| \ \leq \ \|w_i\| \ \leq \ K^{\varepsilon i} e^{\alpha i}\|w_i^*\|, \qquad \varepsilon = K\alpha\theta.$$

Lemma 4.7 *There exists $c' > 0$ such that for every $0 \leq k < n$,*

$$\|w_n^*\| \geq K^{-1} d_{\mathcal{C}}(z_j) e^{c'(n-k)} \|w_k^*\|$$

where j is the first time $\geq k$ when a bound period extending beyond time n is initiated. If no such j exists, the factor $d_{\mathcal{C}}(z_j)$ in the inequality above is replaced by δ in general, by 1 if z_n is a free return.

Lemma 4.8 *Let $k < n$ and assume z_n is free. Then*

$$\|w_n\| > K^{-1}\delta \ e^{c'(n-k)}\|w_k\|,$$

with δ omitted if $z_n \in \mathcal{C}^{(0)}$.

4.3 Controlled orbits as "guides" for other orbits

(IA2)–(IA6) are about orbits starting from $\Gamma_{\theta N}$. In Sect. 4.2 we introduced a class of orbits that successfully use orbits from $\Gamma_{\theta N}$ as their "guides". We now let these orbits serve as guides for other orbits and study the properties they pass along. This is the essence of the replication process.

Throughout Sect. 4.3 we assume that

(1) $z_0 \in \mathcal{C}^{(0)}$, $w_0 = \binom{0}{1}$, and (z_0, w_0) is controlled by $\Gamma_{\theta N}$ up to time m;
(2) $d_{\mathcal{C}}(z_i) > \delta_0$ for $i \leq n_0$ and $> e^{-\alpha i}$ for all $n_0 < i \leq m$.

Observe that these conditions are satisfied by all $z_0 \in \Gamma_{\theta N}$. Our first order of business is to establish that for all ξ_0 bound to z_0, $\hat{w}_i^*(\xi_0)$ copies $w_i^*(z_0)$ faithfully. A detailed proof of the following lemma is given in Appendix B.7.

Lemma 4.9 (cf. [BC2], Lemma 7.8) *Let (z_0, w_0) be as above, and let $\xi_0 \in \mathcal{C}^{(0)}$ be an arbitrary point which we think of as bound to z_0. Let $M_\mu(\cdot)$ and $\theta_\mu(\cdot)$ have the same meaning in (IA6). Then the estimates for*

$$\frac{M_\mu(\xi_0)}{M_\mu(z_0)}, \quad \frac{M_\mu(z_0)}{M_\mu(\xi_0)} \quad \textit{and} \quad |\theta_\mu(\xi_0) - \theta_\mu(z_0)|$$

as stated in (IA6) hold for all $\mu \leq min(p, m)$. The corresponding distortion estimates for two points ξ_0 and ξ_0' bound to z_0 apply as well.

In the rest of this subsection we consider the situation where z_0 is a critical point on a $C^2(b)$-curve in the sense of Sect. 2.6 and study the quadratic behavior as this curve is iterated. More precisely, let e_m be the contractive field of order m, which we know from Lemmas 4.6 and 4.7 is defined at z_0. We assume

(3) z_0 lies on a $C^2(b)$-curve $\gamma \subset \mathcal{C}^{(0)}$, and $e_m(z_0)$ is tangent to γ.

For $\xi_0 \in \gamma$, let $p = p(\xi_0)$ denote the bound period between z_0 and ξ_0. We assume that during its bound period, the orbit of ξ_0 inherits the secondary and higher order bound structures of the orbit of z_0.

Lemma 4.10 *In the part of γ where $p < m$, p increases monotinically with distance from z_0.*

Proof: Proceeding inductively, we assume that on a connected subsegment γ_k of γ one of whose end points is z_0, the minimum bound period is k. It suffices to show that at time $k + 1$, the part of γ_k that remains bound to z_0 is connected. We may assume $T^k(\gamma_k)$ is not in a secondary fold period (otherwise all of $T^{k+1}(\gamma_k)$ will be in a bound period), and that $d_{\mathcal{C}}(\xi_0) > \frac{1}{2}\delta$ for all $\xi_0 \in T^k(\gamma_k)$.

Let $T^k(\gamma_k) = \gamma^{(1)} \cup \gamma^{(2)}$ where $\gamma^{(1)}$ consists of points for which the primary fold period remains in effect and $\gamma^{(2)}$ its complement. Then $\gamma^{(1)}$ is contained in a disk B of radius $K^k b^{\frac{k}{2}}$ centered at z_k, and the bound period on no part of B can expire at time $k+1$. If the bound period of any part of $\gamma^{(2)}$ is to expire at time $k+1$, then the far end of $\gamma^{(2)}$ must be $> K^{-1}e^{-\beta(k+1)}$ from z_k. Also, its tangent vectors are b-horizontal. One concludes that $T^k(\gamma) \setminus B$ is a b-horizontal connected segment which will remain horizontal in the next iterate, forcing the desired picture. □

Let $s \to \xi_0(s)$ be the parametrization of γ by arc length with $\xi_0(0) = z_0$. The following lemma, whose proof is given in Appendix B.8, contains a **distance formula** for $|\xi_\mu(s) - z_\mu|$. See Sect. 2.5 for comparison with 1-d.

Lemma 4.11 *Let $\varepsilon_1 > 0$ be given. Then for all $\mu \in \mathbb{Z}^+$ and $s > 0$ satisfying $\mu \leq m$, $(Kb)^{\frac{\mu}{2}} < s$ and $p(\xi_0(s)) \geq \mu$, we have*

$$(1-\varepsilon_1)\ \|w_\mu(0)\|\ K_1 s^2 \ < \ |\xi_\mu(s) - z_\mu| \ < \ (1+\varepsilon_1)\ \|w_\mu(0)\|\ K_1 s^2 \qquad (9)$$

where $K_1 = \frac{1}{2}|\frac{dq_1}{dx}(z_0)|$.

Corollary 4.2 *Assume in addition to (1)–(3) above that* $\|w_j^*(z_0)\| > e^{cj}$ *for all* $j \le m$*. Let* $\xi_0 \in \gamma$*. Suppose that* $|\xi_0 - z_0| = e^{-h}$ *and* $p(\xi_0) \le m$*. Then*

(a) $\frac{h}{3K_2} \le p \le \frac{3h}{c}$ *where* $K_2 = \log \|DT\|$*;*

(b) $\|w_p(\xi_0)\| \cdot |\xi_0 - z_0| \ge e^{\frac{cp}{3}}$*.*

Proof: (a) The lower bound for p follows from the fact that for all $j \le \frac{h}{3K_2}$, $|\xi_j - z_j| < \|DT\|^j |\xi_0 - z_0| < e^{-\frac{2h}{3}} << e^{-\beta \frac{h}{3K_2}}$. By Lemma 4.11, p is the smallest μ such that $\|w_\mu(0)\| \cdot |z_0 - \xi_0|^2 > K_1^{-1} e^{-\beta\mu}$. This must happen for some $\mu \le \frac{3h}{c}$ because $\|w_{\frac{3h}{c}}(z_0)\| \cdot |z_0 - \xi_0|^2 > K^{-\varepsilon \frac{3h}{c}} \|w^*_{\frac{3h}{c}}(z_0)\| \cdot |z_0 - \xi_0|^2 > K^{-\varepsilon \frac{3h}{c}} e^{c \cdot \frac{3h}{c}} e^{-2h} > 1$.

(b) This follows from the fact that $\|w_p(\xi_0)\| \approx \|w_p(z_0)\|$ (Lemma 4.9) and $|z_0 - \xi_0| \cdot \| w_p(\xi_0) \| > e^{-\frac{\beta}{2}p} \| w_p(\xi_0) \|^{\frac{1}{2}} > e^{-\frac{\beta}{2}p} e^{\frac{cp}{2}} > e^{\frac{cp}{3}}$. □

In analogy with Definition 3.3, we define for $\xi_0(s) \in \gamma$ the notion of a *fold period with respect to* z_0. This is the number ℓ such that $(Kb)^{\frac{\ell}{2}} \approx s$. If $\tau_0(\xi_0)$, the unit tangent vector to γ at ξ_0, is split according to this definition, then the rejoining of the E_i-vector for $\ell < i < p$ has negligible effect. We may assume also that as we iterate, the sub-segment of γ bound to z_0 acquires the same fold periods as z_i, and think of these as *secondary fold periods* for ξ_i .

Corollary 4.3 *Let the assumptions and notation be as in Corollary 4.2. We let* $p = p(\xi_0)$ *where* $|\xi_0 - z_0| = e^{-h}$ *and assume that* z_p *is not in a fold period. Then*

(a) the subsegment of $T^p\gamma$ *between* ξ_p *and* z_p *contains a curve* $\ge e^{-K\beta h}$ *in length and with b-horizontal tangent vectors;*

(b)

$$\|\tau_p(\xi_0)\| \ge K^{-1} e^{h(1-\beta K)}.$$

Proof: (a) By definition, $|\xi_p - z_p| > e^{-\beta p}$. The part of $T^p\gamma$ in a fold period with respect to z_0 has length $\le (Kb)^{\frac{p}{2}} \|DT\|^p$, and the rest have b-horizontal tangent vectors. To convert these estimates in p into bounds involving h, use Corollary 4.2(a).

(b) Splitting τ_0 using e_p, we see that $\|w_p\| \sim e^h \|\tau_p\|$. Combining this with Lemmas 4.11 and 4.9, we have $e^h \|\tau_p(\xi_0)\| \sim \|w_p(\xi_0)\| \approx \|w_p(z_0)\| > K^{-1} |\xi_p - z_p| e^{2h} \ge K^{-1} e^{-K\beta h} e^{2h}$. □

5 Pushing the Induction Forward

The goal of this section is to define Δ_{3N} and to prove that (IA1)–(IA6) hold up to time $3N$ for parameters in Δ_{3N}. The key to this inductive step is the correct splitting of the w_i^*-vectors at free returns (Proposition 5.2). This is proved with the aid of another important fact, namely the control of points in ∂R_k (Proposition 5.1).

5.1 Control of ∂R_k, $k \le \theta N$

For $z \in \partial R_k$, let $\tau(z)$ denote a unit tangent vector to ∂R_k at z.

Proposition 5.1 *For every $\xi_0 \in \partial R_0$ and every $k \le \theta N$, (ξ_0, τ_0) with $\tau_0 = \tau(\xi_0)$ is controlled up to time k by Γ_k.*

Proof: The proof proceeds by induction. The correctness of splitting of τ_0 is evident. We assume all (ξ_0, τ_0) have been controlled up to time $k-1$, so that it makes sense to speak of ξ_k as being in a bound or free state. Suppose ξ_k is bound to z_i for some $z_0 \in \Gamma_{k-1}$. Since $d_{\mathcal{C}}(z_i) > e^{-\alpha i}$, we have $z_i \in \mathcal{C}^{(j)} \setminus \mathcal{C}^{(j+1)}$ for some $j << i \le k$. By Lemma 4.2, ξ_k is h-related to Γ_k, and by Lemma 4.5, τ_k^* splits correctly, proving control at step k. Before proceeding to the free case, we state a lemma of independent interest:

Lemma 5.1 *Let γ be a subsegment of ∂R_k. If all the points on γ are free, then γ is a $C^2(b)$-curve.*

Proof: That τ_k is a b-horizontal vector is an immediate consequence of the splitting algorithm. As for curvature, we appeal to Lemma 2.4 after using Lemma 4.8 to establish that $\|\tau_k\| > K^{-1}\delta e^{c'(k-i)}\|\tau_i\|$ for all $i < k$. $\diamondsuit$

Returning to the proof of Proposition 5.1, let ξ_k be a free return, and let γ be the maximal free subsegment of ∂R_k containing ξ_k. Since the end points of γ are in bound state, they cannot be in $\mathcal{C}^{(k-1)}$ as explained earlier. This leaves two possibilities for the relation between γ and $\mathcal{C}^{(k-1)}$.

Case 1. γ passes through the entire length of a component of $\mathcal{C}^{(k-1)}$. In this case we know from (IA1) that there is a critical point $z_0 \in \gamma$. To see that every $\xi' \in \gamma \cap \mathcal{C}^{(0)}$ is h-related to Γ_k, start from z_0 and move away from it along γ. Using the $C^2(b)$ property of γ, the structure of critical regions (see (IA1)) and the fact that $\gamma \cap \partial R_i = \emptyset\ \forall i < k$, we observe that after leaving $\partial Q^{(k)}(z_0)$ one gets into $Q^{(k-1)}(z_0)$, then $Q^{(k-2)}(z_0)$, and so on, with $d_{\mathcal{C}}(\xi') \ge \rho^i$ for $\xi' \in Q^{(i-1)}(z_0) \setminus Q^{(i)}(z_0)$. For the splitting of $\tau(\xi')$, it follows from Lemma 4.1 and the $C^2(b)$ property of γ that for $\xi' \in \gamma \cap Q^{(i-1)}(z_0) \setminus Q^{(i)}(z_0)$, $\angle(\tau(\xi'), \tau(\phi(\xi')) \le \angle(\tau(\xi'), \tau(z_0)) + \angle(\tau(z_0), \tau(\phi(\xi')) < (Kb)|\xi' - z_0| + (Kb)^{\frac{i-1}{4}} < \varepsilon_0 d_{\mathcal{C}}(\xi')$.

Case 2. γ does not intersect $\mathcal{C}^{(k-1)}$. Let $j < k$ be the largest integer such that $\gamma \cap \mathcal{C}^{(j-1)} \neq \emptyset$. Then there exists $z \in \gamma \cap (\hat{Q}^{(j)} \setminus Q^{(j)})$ for some $Q^{(j)}$. Suppose for definiteness that z lies in the right component of $\hat{Q}^{(j)} \setminus Q^{(j)}$. Moving left along γ from z, we note that since $\gamma \cap Q^{(j)} = \emptyset$, the left end point $\hat{z}$ of γ must also be in the same component of $\hat{Q}^{(j)} \setminus Q^{(j)}$. H-relatedness and correct splitting are now proved as in Case 1 with $\hat{z}$ playing the role of z_0. We know $\tau(\hat{z})$ splits correctly because $\hat{z}$ is, by definition, in a bound state. $\square$

5.2 Extending control of $\Gamma_{\theta N}$-orbits to time $3N$

We continue to assume (IA1)–(IA6). Let $z_0 \in \Gamma_{\theta N}$ and $w_0 = \binom{0}{1}$. The next proposition plays a key role in the inductive process.

Proposition 5.2 *If $z_0 \in \Gamma_{\theta N}$ satisfies $d_{\mathcal{C}}(z_i) > e^{-\alpha i}$ for all $i \leq 3N$, then (z_0, w_0) is automatically controlled by $\Gamma_{\theta N}$ up to time $3N$. In fact, we have the following stronger results on the angle of splitting:*

(i) If z_i is a free return, then

$$\angle(w_i, \tau(\phi(z))) << \varepsilon_0 d_{\mathcal{C}}(z).$$

(ii) If z_i is a bound return, then

$$\angle(w_i^*, \tau(\phi(z))) < \varepsilon_0 d_{\mathcal{C}}(z).$$

Proof: From the condition that $d_{\mathcal{C}}(z_i) > e^{-\alpha i}$, we know that z_0 is h-related to $\Gamma_{\theta N}$ up to time $3N$ (see the remark following (IA2) in Sect. 3.3.2), and that $p < K\alpha 3N << N$. To prove the assertions on the angle of splitting, we proceed inductively, assuming they are valid up to time $k-1$ for some $k \leq 3N$.

First, we consider the case where z_k is a free return. Then either $z_k \in \hat{Q}^{(j)} \setminus Q^{(j)}$ for some $j \leq \theta N$, or $z_k \in \mathcal{C}^{([\theta N])}$. In the latter case we let $j = [\theta N]$ for purposes of the following arguments.

Claim 5.1 *There exists j', $\frac{1}{2}j \leq j' < j$, such that if*

$$\xi_0 = z_{k-j'} \qquad \textit{and} \qquad u_0 = \frac{w_{k-j'}(z_0)}{\|w_{k-j'}(z_0)\|}\ ,$$

then for $0 \leq s < j'$,

$$\|DT^s(\xi_0)u_0\| \geq \|DT\|^{-s}.$$

Proof of Claim 5.1: We consider the graph $\mathcal{G}$ of $i \mapsto \log \|w_i(z_0)\|$ for $k-j < i \leq k$. Let L be the (infinite) line through $(k,\ \log\|w_k\|)$ with slope $\log\|DT\|$. Then clearly, all the points in $\mathcal{G}$ lie above L. Let P be the intersection of L with the line $x = k - \frac{1}{2}j$. We let L be pivoted at P and rotate it clockwise until it hits some point in $\mathcal{G}$. (Draw a picture!) Let $k - j'$ be the first coordinate of the first point hit. Then $\frac{1}{2}j \leq j' < j$, and Claim 5.1 is proved if we can show that in its final position, the slope of L is $\geq -\log\|DT\|$. This is true because z_k being a free return, $\|w_{k-j}\| < \|w_k\|$ by Lemma 4.8, so the straight line joining the two points $(k-j,\ \log\|w_{k-j}\|)$ and $(k,\ \log\|w_k\|)$ has slope $\geq -\log\|DT\|$.[8] $\diamondsuit$

Now by Lemma 2.3, there exists an integral curve γ of the most contracted field of order j' through ξ_0 having length $\mathcal{O}(1)$. Since γ follows roughly the direction of e_1, it has slope $> K^{-1}\delta$ outside of $\mathcal{C}^{(0)}$ and is roughly a parabola inside $\mathcal{C}^{(0)}$ (Lemma 2.9). In both cases, γ meets ∂R_0. Let $\xi_0' \in \gamma \cap \partial R_0$. Then

$$|\xi_s - \xi_s'| < (K^2 b)^s$$

for all $0 \leq s \leq j'$. Our next claim is made possible by Proposition 5.1.

Claim 5.2 $\xi'_{j'}$ *is a free return.*

Proof of Claim 5.2: If not, then $\xi'_{j'}$ would be bound to $\hat{z}$, a point on a critical orbit, and we would have $\xi_{j'}, \xi'_{j'} \in \hat{Q}^{(i)}(\hat{z})$ for some $i << j' < j$ with $d_{\mathcal{C}}(\xi_{j'}) \approx d_{\mathcal{C}}(\xi'_{j'}) \approx d_{\mathcal{C}}(\hat{z}) > e^{-\alpha j'}$. This contradicts our assumption that $\xi_{j'} = z_k$ is in $\hat{Q}^{(j)}$ or in $\mathcal{C}^{([\theta N])}$, for in either case, $d_{\mathcal{C}}(z_k) < \rho^{j-1}$. $\diamondsuit$

Claim 5.3 *With u_0 as in Claim 5.1, let*

$$\tau_i = DT^i(\xi_0')\tau_0, \qquad u_i = DT^i(\xi_0)u_0,$$

and let θ_i be the angle between u_i and τ_i. Then $\theta_{j'} \leq b^{\frac{j'}{2}}$.

Proof of Claim 5.3: Write $A = DT(\xi_{i-1})$ and $A' = DT(\xi'_{i-1})$. Then

$$\begin{aligned}\theta_i = \frac{\|\tau_i \times u_i\|}{\|\tau_i\| \cdot \|u_i\|} &= \frac{1}{\|\tau_i\| \cdot \|u_i\|}\|A'\tau_{i-1} \times A'u_{i-1} + A'\tau_{i-1} \times (A - A')u_{i-1}\| \\ &\leq \frac{\|\tau_{i-1}\|}{\|\tau_i\|} \cdot \frac{\|u_{i-1}\|}{\|u_i\|} \cdot (|\det(A')|\theta_{i-1} + K|\xi_i - \xi_i'|) \\ &\leq \frac{\|\tau_{i-1}\|}{\|\tau_i\|} \cdot \frac{\|u_{i-1}\|}{\|u_i\|} \cdot (b\theta_{i-1} + K(K^2b)^{i-1}).\end{aligned}$$

Applying this relation for θ_i recursively, we obtain

$$\theta_{j'} < \left(\sum_{i=0}^{j'} \frac{\|\tau_i\|}{\|\tau_{j'}\|} \cdot \frac{\|u_i\|}{\|u_{j'}\|}\right)(K^2b)^{j'-1}.$$

[8] This result also follows from a lemma of Pliss; see e.g. [Ma].

Since both z_k and $\xi_{j'}$ are free returns, we may use Lemma 4.8 to bound the sum in brackets, completing the proof of Claim 5.3. $\diamondsuit$

We are finally ready to prove our assertion on the angle of splitting for the free return z_k. Recall that $\xi_{j'} = z_k \in \hat{Q}^{(j)}$ or $Q^{([\theta N])}$. Since $|\xi_{j'} - \xi'_{j'}| < (K^2 b)^{j'}$, $\xi'_{j'} \in \partial R_{j'}$ and $j' < j$, we have $\xi'_{j'} \in \partial Q^{(j')}(z_k)$. By our inductive hypothesis, $\tau_{j'}(\xi'_0)$ splits correctly. Since $\angle(w_k(z_0), \tau(\xi'_{j'})) \leq b^{\frac{j'}{2}}$ (Claim 5.3), $\angle(\tau(\phi(\xi'_{j'})), \tau(\phi(z_k))) = \mathcal{O}(b^{\frac{j'}{4}})$ and $|d_{\mathcal{C}}(\xi'_{j'}) - d_{\mathcal{C}}(z_k)| = \mathcal{O}(b^{\frac{j'}{4}})$ (Lemma 4.1), it suffices to prove that $b^{\frac{j'}{4}} << d_{\mathcal{C}}(z_k)^2$. In the case where $z_k \in \hat{Q}^{(j)} \setminus Q^{(j)}$, this is trivial as $d_{\mathcal{C}}(z_k) \sim \rho^j$. In the case where $z_k \in Q^{([\theta N])}$, since $d_{\mathcal{C}}(z_k) > e^{-\alpha k}$, we have $d_{\mathcal{C}}(z_k)^2 > e^{-6\alpha N}$, which we may assume is $>> b^{\frac{1}{12}\theta N} \geq b^{\frac{1}{4}j'}$.

To complete our proof of Proposition 5.2, we now consider the case when z_k is a bound return. Our argument is along the lines of Lemma 4.5, with the following modifications to get the sharper result claimed in assertion (ii).

The argument in the proof of Lemma 4.5 transfers the problem of estimating the angle of splitting in question to that of estimating $\angle(DT^{k-j}(\hat{z}_0)u,\ \tau(\phi(\hat{z}_{k-j}))$ where $\hat{z}_0 \in \Gamma_{\theta N}$ is a binding point for z_j for some $j < k$ (and u is as in Lemma 4.5). If $\hat{z}_{k-j}$ is a free return, then this angle is $<< \varepsilon_0 d_{\mathcal{C}}(z_k)$ according to assertion (i) in this proposition, and we are done.

Suppose $\hat{z}_{k-j}$ is not free. For definiteness, let us first consider the case where $\hat{z}_{k-j}$ is bound to the orbit of $\hat{z}'_0$ with a binding initiated at time j', $j < j' < k$, and that $\hat{z}'_{k-j'}$ is a free return. We first apply assertion (i) of this proposition to $\angle(DT^{k-j'}(\hat{z}'_0)u,\ \tau(\phi(\hat{z}'_{k-j'})))$. Based on this information, we use our modified version of Lemma 4.5 to first estimate $\angle(DT^{k-j}(\hat{z}_0)u,\ \tau(\phi(\hat{z}_{k-j}))$, and then repeat the argument to estimate the angle of splitting of w^*_k. Observe that here $d_{\mathcal{C}}(z_k) \approx d_{\mathcal{C}}(\hat{z}'_{k-j'})$.

In general, $\hat{z}'_{k-j'}$ may not be free, in which case we consider the critical orbit it is following, and so on. It may take several steps before we arrive at the situation of a guiding critical orbit making a free return. We need to argue that the errors in these successive approximations do not add up. They do not, because for the same reason that $d_{\mathcal{C}}(z_k) \approx d_{\mathcal{C}}(\hat{z}'_{k-j'})$ above, each approximation guarantees that $d_{\mathcal{C}}(z_k)$ is bigger, so that the errors in the constant in front of $d_{\mathcal{C}}(z_k)$ form a geometric series, the sum of which we may assume is $< \varepsilon_0$. This completes the proof of Proposition 5.2. $\square$

5.3 Verification of (IA1)–(IA6) up to time $3N$

Step 1 *Deletion of parameters.* We delete from Δ_N all (a, b) for which there exists $z_0 \in \Gamma_{\theta N}$ and i, $N < i \leq 3N$, such that

$$d_{\mathcal{C}}(z_i) < e^{-\alpha i} \qquad \text{or} \qquad \|w^*_i(z_0)\| < e^{ci}.$$

The set of remaining parameters is called Δ_{3N}. We do not claim in (IA1)–(IA6) that Δ_{3N} has positive measure or even that it is nonempty; this is discussed in Section 6.

Steps 2–5 below apply to $T = T_{a,b}$ for $(a, b) \in \Delta_{3N}$.

Step 2 *Updating of* $\Gamma_{\theta N}$. For each $z_0 \in \Gamma_{\theta N}$, since $\|w_i\|$ grows exponentially (Step 1 and Lemma 4.6), there exists a unique z'_0 on the component of $\partial\mathcal{C}^{(k)}$ containing z_0 that is a critical point of order $3N$ (Lemma 2.10). Let $\Gamma'_{\theta N}$ be the set of these z'_0, i.e. $\Gamma'_{\theta N}$ is a copy of $\Gamma_{\theta N}$ updated to order $3N$.

Step 3 *Construction of* $\Gamma_{3\theta N}$ *and* $\mathcal{C}^{(k)}$, $\theta N < k \leq 3\theta N$. We proceed inductively, assuming all has been accomplished for $k-1$.

First we establish control of ∂R_k as in Sect. 5.1, with one minor (technical) difference to be explained below. It follows that R_k meets $\mathcal{C}^{(k-1)}$ in at most a finite number of components bounded by free, and hence $C^2(b)$, curves.

Next, we construct critical points on ∂R_k. Let Q be one of the components of $R_k \cap \mathcal{C}^{(k-1)}$, and let γ be one of its horizontal boundaries. By Lemma 2.11, there exists a critical point $\hat{z}_0 \in \gamma$ of order $\hat{m} = min\{3N, -\log d(z_0, \gamma)^{\frac{1}{2}}\}$ where $z_0 \in \Gamma'_{\theta N}$ lies on the boundary of the component $Q^{([\theta N])}$ containing γ. Since $d(z_0, \gamma) = \mathcal{O}(b^{\frac{\theta N}{2}})$, we have, assuming θ is chosen with $e^{-3N} > K^{-N} > b^{\frac{\theta N}{4}}$, that $\hat{z}_0$ is of order $3N$.

The critical regions $\mathcal{C}^{(k)}$ are then constructed as follows: For each Q as above, choose one of the critical points on ∂Q, and define $Q^{(k)} := \{\xi \in Q :$ the horizontal distance between ξ and $\hat{z}_0$ is $\leq \rho^k\}$. This is the component of $\mathcal{C}^{(k)}$ in Q.

To continue, we need to set bindings for points in ∂R_k. Technically, only $z_0 \in \Gamma_{\theta N}$ (and not the critical points on ∂R_i, $\theta N < i \leq k$) can be used. This is of no concern to us for the following reason: for k' with $k < k' \leq 3\theta N$, only those parts of $\partial R_{k'}$ that are free are involved in the construction of $\mathcal{C}^{(k')}$; and for $\xi_0 \in \partial R_k \cap \mathcal{C}^{([\theta n])}$, independent of which $z_0 \in Q^{([\theta n])}(\xi_0)$ we think of it as bound to, ξ_i will remain in bound state through time $3\theta N$ because $|\xi_i - z_i| \leq K^{3\theta N}\rho^{\theta N} << e^{-3\beta\theta N}$.

This completes the constructive procedure. The critical points in ∂R_k, $N < k \leq 3N$, together with $\Gamma'_{\theta N}$ form $\Gamma_{3\theta N}$. To complete the verification of (IA1) up to time $3N$, we need to explain the uniqueness of $\hat{z}_0$ as a critical point of order $3N$ on γ. Since e_1 is defined everywhere in $\mathcal{C}^{(0)}$ and has derivative $> K^{-1}$, while γ is a $C^2(b)$, curve, Corollary 2.1 limits the possibility of any critical points to an interval of length $\mathcal{O}(b)$. On this interval, e_2 is defined, further limiting the candidates for critical points to an interval of length $\mathcal{O}(b^2)$ etc. Finally, The bound on the Hausdorff distance between the two horizontal boundaries of $Q^{(k)}$ as stated in Theorem 1 is a triviality and not an inductive fact: it is true because $\text{area}(R_k) \leq |\det(DT^k)| < (Kb)^k$ and the two horizontal boundaries of $Q^{(k)}$ are roughly parallel.

Step 4 *Updating the definitions of* $d_{\mathcal{C}}(\cdot)$ *and* $\phi(\cdot)$. Using $\Gamma_{3\theta N}$ and $\mathcal{C}^{(k)}$, $k \leq [3\theta N]$, we reset these definitions for $z \in \mathcal{C}^{([\theta N])}$ in accordance with Definition 3.2. Since $|\text{old}\phi(z) - \text{new}\phi(z)| = \mathcal{O}(b^{\frac{\theta N}{4}})$ and $|\tau(\text{old}\phi(z)) - \tau(\text{new}\phi(z))| = \mathcal{O}(b^{\frac{\theta N}{4}})$ (Lemma 4.1), these changes have essentially no effect on the correctness of splitting for points with $d_{\mathcal{C}}(\cdot) > b^{\frac{3\theta N}{20}}$. The relations in (IA5) are also not affected.

Step 5 *Verification of (IA2)–(IA6) for* $i \leq 3N$. This is carried out in 3 stages.

(1) First we argue that for $z_0 \in \Gamma_{\theta N}$ (we really mean $\Gamma_{\theta N}$, not $\Gamma'_{\theta N}$), (IA2)–(IA6) hold for $i \leq 3N$: (IA2) and (IA4) hold by design; (IA3) is given by Proposition 5.2, and (IA5) and (IA6) are proved in Sect. 4.3 with $m = 3N$.

(2) With the properties of $\Gamma_{\theta N}$ in (1) having been established, we observe that continuing to use $\Gamma_{\theta N}$ as the source of control, the material in Sects. 4.2 and 4.3 are now valid for times up to $min(m, 3N)$.

(3) We are now ready to argue that (IA2)–(IA6) hold for all $z'_0 \in \Gamma_{3\theta N}$. For each $z'_0 \in \Gamma_{3\theta N}$, whether it is in $\Gamma'_{\theta N}$ or of generation $> \theta N$, there exists $z_0 \in \Gamma_{\theta N}$ such that $|z'_0 - z_0| = \mathcal{O}(b^{\frac{\theta N}{4}})$. This implies, for $i \leq 3N$, that $|z'_i - z_i| < b^{\frac{\theta N}{4}} \|DT\|^{3N} << e^{-\beta 3N}$ provided θ is chosen so that $b^{\frac{\theta}{4}} \|DT\|^3 < \frac{1}{2} e^{-\beta}$. (IA2) follows immediately from the corresponding condition for z_0. Regarding z'_0 as bound to z_0 for at least $3N$ iterates, (IA3) and (IA4) follow from property (IA6) of z_0. Finally, regarding $(z'_0, \binom{0}{1})$ as controlled by $\Gamma_{\theta N}$ up to time $3N$, we obtain (IA5) and (IA6) from Lemmas 4.9-4.11 and Corollary 4.2.

Conclusions from Sections 3–5: After letting N go to infinity, we have defined for each $T = T_{a,b}$ with $(a, b) \in \Delta := \cap_N \Delta_N$ a set $\mathcal{C}$ given by $\mathcal{C} = \cap_{i \leq 0} \mathcal{C}^{(i)}$. This is the *critical set* in Theorem 1. Let Γ be the set to which $\Gamma_{\theta N}$ converges as $N \to \infty$. An equivalent characterization of $\mathcal{C}$ is that it is the set of accumulation points of Γ. Clearly, the properties that $d_{\mathcal{C}}(z_i) \geq e^{-\alpha i}$ and $\|w_i\|$ grows exponentially are passed on to points in $\mathcal{C}$. We have thus completed the proof of Theorem 1 modulo the positivity of the measure of Δ.

6 Measure of Selected Parameters

In this section we fix $b > 0$ and consider the 1-parameter family $a \mapsto T_{a,b}$. Let $\Delta_b = \{a : (a,b) \in \Delta\}$. The Lebesgue measure of a set $A \subset \mathbb{R}$ is denoted by $|A|$. More generally, we use $|\cdot|$ to denote the measure on curves induced by arc length. The purpose of this section is to prove that $|\Delta_b| > 0$ for all sufficiently small $b > 0$.

6.1 Phase-space dynamics and curves of critical orbits

Assuming $\delta = e^{-\mu^*}$ for some $\mu^* \in \mathbb{Z}^+$, let $\mathcal{P} = \{I_{\mu j}\}$ be the partition of the interval $(-\delta, \delta)$ defined as follows: for $\mu \geq \mu^*$, let $I_\mu = (e^{-(\mu+1)}, e^{-\mu})$, and let each I_μ be further subdivided into μ^2 subintervals of equal length called $I_{\mu j},\ j = 1, 2, \cdots \mu^2$; for $\mu \leq -\mu^*$, let $I_{\mu j} = -I_{(-\mu)j}$.

Next let γ be a curve with nearly horizontal tangent vectors. We assume for simplicity that γ meets only one component $Q^{(0)}$ of $\mathcal{C}^{(0)}$, and let $\hat{z} = (\hat{x}, \hat{y})$ be a point near the center of $Q^{(0)}$. The partition $\mathcal{P}_{\gamma,\hat{z}}$ on γ is defined to be $(\psi^{-1}\mathcal{P})|\gamma \cup \{I^\pm\}$ where $\psi(x,y) = x - \hat{x}$ and $I^\pm$ are the two components of $\gamma \setminus \psi^{-1}(-\delta, \delta)$. An element of $\mathcal{P}_{\gamma,\hat{z}}$ is said to have "full length" if its image under ψ is either equal to some $I_{\mu j}$ or longer than all the $I_{\mu j}$'s. When γ and $\hat{z}$ are understood, we often refer to $\mathcal{P}_{\gamma,\hat{z}}$ simply as $\mathcal{P}$ and $(\psi^{-1} I_{\mu j}) \cap \gamma$ as $I_{\mu j}$.

Before proceeding to the estimation of $|\Delta_b|$, we consider first the following problem in phase-space dynamics. The estimation of $|\Delta_b|$ includes an argument parallel to and more complicated than this.

A model problem in phase-space dynamics

Let $T = T_{a,b}$ with $(a,b) \in \Delta$. Recall from the proof of Proposition 5.1 that if $\gamma \subset \partial R_k$ is a maximal free segment meeting some $Q^{(0)}$, then either $\gamma \cap Q^{(0)}$ contains a critical point $\hat{z} \in \Gamma$ or the entire segment $\gamma \cap Q^{(0)}$ is h-related to some $\hat{z} \in \Gamma$. In both cases, $\mathcal{P}_{\gamma,\hat{z}}$ is the partition of choice on γ. Note that for $z \in I_{\mu j},\ d_{\mathcal{C}}(z) \approx e^{-|\mu|}$.

Let ω_0 be a subsegment of ∂R_0, and write $\omega_i := T^i \omega_0$. We assume that (i) for all $z_0 \in \omega_0,\ d_{\mathcal{C}}(z_i) > e^{-\alpha i}$ for all $i \leq N$, and (ii) ω_N is free and is approximately equal to some $I_{\mu_0 j_0}$. The problem is to find a lower estimate for the measure of $\{z_0 \in \omega_0 : d_{\mathcal{C}}(z_i) > e^{-\alpha i} \text{ for all } i\}$.

We may assume that all the points in ω_N have the same bound period, and let $i_1 > N$ be the first moment in time after the expiration of this bound period when $\omega_{i_1} \cap \mathcal{C}^{(0)}$ contains an $I_{\mu j}$ of full length. This must happen at some point, for the length of ω_i grows by a factor $> K$ between successive free returns (Corollary 4.3). It is easy to check that $d_{\mathcal{C}} > e^{-\alpha i}$ is not violated between times N and i_1. Let $\{\omega\}$ be the partition $\mathcal{P}$ on ω_{i_1} with end segments attached to their neighbors if they are not of full length. We delete those ω's that contain some z with $d_{\mathcal{C}}(z) < e^{-\alpha i_1}$. For each ω that is kept, we repeat the procedure above with ω in the place of ω_N, that is, we iterate until ω makes a free return at time $i_2 = i_2(\omega)$ with $T^{i_2 - i_1}\omega$ containing an $I_{\mu j}$ of full length. We then partition $T^{i_2-i_1}\omega$, discard subsegments that violate $d_{\mathcal{C}} > e^{-\alpha i_2}$, and continue to iterate the rest.

We estimate the fraction of ω_{i_1} deleted at time i_1 as follows. Since $\omega_N \approx I_{\mu_0 j_0}$, the bound period p is $\leq K|\mu_0|$. From Corollary 4.3, we see that ω_{i_1} has length $> \frac{K^{-1}}{\mu_0^2}e^{-\beta K|\mu_0|} > e^{-2\beta|\mu_0|K}$. Now $|\mu_0| \leq \alpha N$ and $i_1 > N + p_0$ where $p_0 > 0$ is a lower bound for all bound periods. Then

$$\frac{|\{z \in \omega_{i_1} : d_{\mathcal{C}}(z) < e^{-\alpha i_1}\}|}{|\omega_{i_1}|} \;<\; \frac{2e^{-\alpha(N+p_0)}}{e^{-2K\alpha\beta N}} \;<\; e^{-\frac{1}{2}\alpha N}$$

assuming N is sufficiently large. Similarly, for each subsegment $\omega \approx I_{\mu j}$ of ω_{i_1} that is kept, the fraction of $T^{i_2-i_1}\omega$ deleted at time i_2 is $< e^{-\frac{1}{2}\alpha i_1} < e^{-\frac{1}{2}\alpha(N+p_0)}$, and so on. To estimate the total measure of ω_0 deleted, these fractions have to be pulled back to ω_0. This involves a *distortion estimate* for DT^i along certain subsegments of ∂R_k. Using the fact that this distortion is uniformly bounded (Lemma 8.2), we see that the fraction of ω_0 deleted in this procedure is $< K\sum_i e^{-\frac{1}{2}\alpha(N+ip_0)} < Ke^{-\frac{1}{2}\alpha N}$.

We remark that the scheme in this paragraph relies on the fact that ω_N has a certain minimum length depending on N, otherwise the entire segment may be obliterated before time i_1 is reached.

Strategy for estimating $|\Delta_b|$

Since b is fixed throughout this discussion, let us for notational simplicity omit mention of it and write Δ, Δ_N and T_a instead of $\Delta_b, \Delta_b \cap \Delta_N$ and $T_{a,b}$. Let N be fixed. The problem is to estimate the measure of parameters deleted between times N and $3N$. Our strategy is as follows: For $\hat{a} \in \Delta_N$ and $z_0 \in \Gamma_{\theta N}(\hat{a})$, let $a \mapsto z_0(a)$ be defined on an interval containing $\hat{a}$. We consider

$$\gamma_0 \;\to\; \gamma_1 \;\to \gamma_2 \;\to\; \cdots \qquad \text{where} \qquad \gamma_i(a) := z_i(a) = T_a^i(z_0(a)),$$

and estimate the measure of the set of a for which $z_i(a)$ violates (IA2) or (IA4).

The idea behind this line of proof is that *qualitatively, the evolution of γ_0 is similar to that of* ω_0 in the model phase-space problem. If this is true, then the measure deleted on account of (IA2) can be estimated analogously. To understand why the γ_i's behave like phase curves, i.e. curves that are obtained through the iteration of T_a, observe the way in which $\frac{d}{da}\gamma_i$, the tangent vector to the curve $a \mapsto \gamma_i(a)$, is transformed: if $\|\frac{d}{da}\gamma_i(a)\| >> 1$, then $\frac{d}{da}\gamma_{i+1}(a) \approx DT_a(\gamma_i(a))\frac{d}{da}\gamma_i(a)$; that is to say, $\gamma_{i+1} \approx T_a \circ \gamma_i$ near $\gamma_i(a)$.

Issues to be addressed

1. *Similarity of space- and a-derivatives.* This is the first and most important step in justifying the thinking in the last paragraph. Let γ_0 be as above. In Sect. 6.2, we show that $\frac{d}{da}\gamma_i \sim DT^i(\frac{d}{da}\gamma_0)$ or $DT^i\binom{0}{1}$. As we will see, this is made possible by

our transversality condition on $\{f_a\}$ in Sect. 1.1. The only other prerequisite for this comparison is that the slopes of γ_0 be suitably bounded. This is verified in Sect. 6.3 for curves corresponding to critical points of all generations and all orders.

2. *Dynamics of the curves* $a \mapsto \gamma_i(a)$. Our next step is to show that *as curves parametrized by* a, the γ_i have properties similar to those of ω_i. For example, with $\Gamma_{\theta N}$ moving with a, how is $d_{\mathcal{C}}(z_i(a))$ affected? Other properties include the geometry of free segments, quadratic behavior of the type in Sect. 4.3, distortion estimates along γ_i etc. These questions are discussed in Sect. 6.4.

3. *Deletions of parameters in violation of (IA2) or (IA4).* We consider $z_0 \in \Gamma_{\theta N}$ one at a time, and let γ_0 be the corresponding curve of critical points. Assuming the success of the last step, deletions on γ_0 on account of (IA2) are estimated following the scheme outlined in the model problem. Estimates for the measure of parameters deleted on account of (IA4) are discussed in Sect. 6.5.

4. *Combined effect of deletions corresponding to all* $z_0 \in \Gamma_{\theta N}$. Obviously, we need to multiply the measure of the parameters deleted on each γ_0 by the cardinality of $\Gamma_{\theta N}$, but there are technical considerations: As in our phase-space model, to get started we need γ_N to have a certain minumum length. This raises the question of the length of the parameter interval on which each $a \mapsto z_0(a)$ can be continued (this problem appears already in Sect. 6.3). Also relevant is the combined effect of deletions on all critical curves prior to time N. The final estimate is made in Sect. 6.6.

The idea to relate parameter-space dynamics to phase-space dynamics is, of course, not new. Two results on 1-dimensional maps are cited without proof and used in this section: a transversality condition from [TTY] is used in Sect. 6.2 and a large deviation estimate from [BC2] is used in Sect. 6.5.

6.2 Equivalence of space- and a-derivatives

The setting of this subsection is as follows: For fixed $b > 0$, let $\hat{a}$ be such that $z_0 = z_0(\hat{a}) \in \Gamma_{\theta N}(\hat{a})$ obeys the conditions in (IA2) and (IA4) and the conclusions of Lemmas 4.6–4.8 up to time n. This assumes implicitly that all the binding structures needed for the last sentence to make sense are in place. (See the first part of Sect. 6.5 for a more detailed discussion.) We assume also that $z_0(\hat{a})$ has a smooth continuation $a \mapsto z_0(a)$ to an a-interval containing $\hat{a}$. Let $w_i = DT^i_{\hat{a}}(z_0(\hat{a}))\binom{0}{1}$ and $\tau_i = \frac{dz_i}{da}(\hat{a})$. The goal of this subsection is to compare w_i and τ_i. Let $\tau_0 = (\tau_{0,1}, \tau_{0,2})$.

Proposition 6.1 *Given $\bar{\tau} > 0$, there exist constants $\lambda_2 > \lambda_1 > 0$ and a small $\varepsilon > 0$ such that the following holds: If $(\hat{a}, b)$ is sufficiently near $(a^*, 0)$, $z_0(\hat{a})$ is as above, $\|\tau_0\| < \bar{\tau}$ and $|\tau_{0,2}| < \varepsilon$, then for all $i \le n$,*

$$\lambda_1 \le \frac{\|\tau_i\|}{\|w_i\|} \le \lambda_2.$$

We will show below that once we have $\|\tau_i\| \sim \|w_i\|$ for some i with $\|w_i\|$ sufficiently large, then this relationship will hold from there on. The estimate for the initial stretch is guaranteed by our transversality condition on the 1-dimensional family $\{f_a\}$. We recall a relevant result from 1-dimension:

Let f and $\{f_a\}$ be as in Sect. 1.1. Let x_0 be a critical point of f, and let $p = f(x_0)$. Since $f = f_{a^*}$, we write $x_0(a^*) = x_0,\ p(a^*) = p$, and let $a \mapsto x_0(a)$ and $a \mapsto p(a)$ be the continuation of x_0 and p as defined in Sect. 1.1. Let $x_k(a) = f_a^k(x_0(a))$. We will use $(\cdot)'$ to denote differentiation with respect to x.

Lemma 6.1 ([TTY], Proposition VII.7) *As $k \to \infty$,*

$$Q_k(a^*) \ := \ \frac{\frac{dx_k}{da}(a^*)}{(f^{k-1})'(x_1(a^*))} \quad \to \quad \lambda_0 := \frac{dx_1}{da}(a^*) - \frac{dp}{da}(a^*).$$

The transversality condition in Sect. 1.1, Step II, states that $\lambda_0 \neq 0$. We will also need the following technical lemma the proof of which is given in Appendix B.9.

Lemma 6.2 *There exist constants K and $c' > 0$ such that for every $0 \leq s < i$, we have*

$$\|DT^{i-s}(z_s)\| \leq Ke^{-c's}\|w_i\|.$$

Proof of Proposition 6.1: Since

$$\tau_i = DT(z_{i-1})\tau_{i-1} + \psi(z_{i-1})$$

where $\psi(z) = \frac{\partial(T_a z)}{\partial a}(\hat{a})$, we have inductively

$$\tau_i = DT^i(z_0)\tau_0 + \sum_{s=1}^{i} DT^{i-s}(z_s)\psi(z_{s-1}).$$

The upper estimate for $\frac{\|\tau_i\|}{\|w_i\|}$ follows from Lemma 6.2 and the uniform boundedness of $\|\psi(\cdot)\|$:

$$\begin{aligned}\frac{\|\tau_i\|}{\|w_i\|} &\leq \frac{\|DT^i(z_0)\tau_0\|}{\|w_i\|} + \sum_{s=1}^{i} \frac{\|DT^{i-s}(z_s)\psi(z_{s-1})\|}{\|w_i\|} \\ &< K\|\tau_0\| + K\sum_{s=1}^{\infty} e^{-c's} := \lambda_2.\end{aligned}$$

To obtain a lower bound for $\frac{\|\tau_i\|}{\|w_i\|}$, we pick k_0 large enough that $|Q_{k_0}(a^*)| > \frac{1}{2}|\lambda_0|$ where Q_{k_0} and λ_0 are as in Lemma 6.1, and decompose τ_i into $\tau_i = I + II$ where

$$I = DT^i(z_0)\tau_0 + \sum_{s=1}^{k_0} DT^{i-s}(z_s)\psi(z_{s-1}),$$

$$II = \sum_{s=k_0+1}^{i} DT^{i-s}(z_s)\psi(z_{s-1}).$$

Again by Lemma 6.2, we have

$$\frac{\|II\|}{\|w_i\|} < \sum_{s=k_0+1}^{\infty} Ke^{-c's}.$$

We will show $\frac{\|I\|}{\|w_i\|} > K_0^{-1}|\lambda_0|$ for some K_0, and assume k_0 is chosen so that $\sum_{s>k_0} Ke^{-c's}$ $<< K_0^{-1}|\lambda_0|$. Write

$$I = DT^{i-k_0}(z_{k_0})V$$

where

$$V = DT^{k_0}(z_0)\tau_0 + \sum_{s=1}^{k_0} DT^{k_0-s}(z_s)\psi(z_{s-1}).$$

Claim 6.1

$$\|V\| \ > \ \frac{1}{3}\,\frac{\|w_{k_0}\|}{\|w_1\|}\,|\lambda_0|,$$

and the second component of V tends to 0 as $(\hat{a}, b) \to (a^*, 0)$.

Proof of Claim 6.1: Let $z_0 \to (x_0, 0)$ as $(\hat{a}, b) \to (a^*, 0)$. The two terms of V are estimated as follows:

(i) $\|DT^{k_0}(z_0)\tau_0\| < K|\tau_{0,2}|$ for $(\hat{a}, b)$ sufficiently near $(a^*, 0)$. This is because k_0 is a system constant, and writing $T^{k_0}_{a^*,0} = (T^1, T^2)$, we have

$$DT^{k_0}_{\hat{a},b}(z_0)\tau_0 \ \to \ \left(\frac{\partial T^1}{\partial x}(x_0,0)\tau_{0,1} + \frac{\partial T^1}{\partial y}(x_0,0)\tau_{0,2},\ 0\right) = \left(\frac{\partial T^1}{\partial y}(x_0,0)\tau_{0,2},\ 0\right).$$

(ii) For $(\hat{a}, b)$ sufficiently near $(a^*, 0)$, z_s stays out of $\mathcal{C}^{(0)}$ for $> k_0$ iterates, and

$$\frac{\sum_{s=1}^{k_0} DT^{k_0-s}(z_s)\psi(z_{s-1})}{\|w_{k_0}\|/\|w_1\|} \ \to \ \left(\frac{\sum_{s=1}^{k_0}(f^{k_0-s})'(x_s(a^*))\frac{d}{da}(f_a(x_{s-1}))(a^*)}{\pm(f^{k_0-1})'(x_1(a^*))},\ 0\right)$$

$$= \ \left(\pm\sum_{s=1}^{k_0}\frac{\frac{d}{da}(f_a(x_{s-1}))(a^*)}{(f^{s-1})'(x_1(a^*))},\ 0\right),$$

which by a simple computation is equal to $(\pm Q_{k_0}(a^*),\ 0)$. ◇

Assuming that $n_0 > k_0$, so that $d_{\mathcal{C}}(z_{k_0}) > \frac{1}{2}\delta_0$, we have that the slope of $e_{i-s}(z_s)$ is bounded below by some K^{-1}. This together with Claim 6.1 gives

$$\|DT^{i-k_0}(z_{k_0})V\| > K^{-1}\|DT^{i-k_0}(z_{k_0})w_{k_0}\|\ \frac{\|V\|}{\|w_{k_0}\|} > K_0^{-1}\|w_i\|\ |\lambda_0|.$$

□

We will also need an estimate on the angle between τ_i and w_i, which we denote by θ_i. The assumptions are as in Proposition 6.1 .

Lemma 6.3 *If z_i is free, then $\theta_i < \frac{K}{\|\tau_i\|}$.*

Proof:

$$| \sin\theta_i | \le \frac{1}{\|\tau_i\|}\left(\sum_{s=1}^{i}\frac{1}{\|w_i\|}\|w_i \times DT^{i-s}(z_s)\psi_(z_{s-1})\| + \frac{\|w_i \times DT^i(z_0)\tau_0\|}{\|w_i\|}\right)$$

$$\le \frac{1}{\|\tau_i\|}\left(\sum_{s=1}^{i}\frac{\|w_s\|}{\|w_i\|}\left\|\frac{w_s}{\|w_s\|}\times\psi(z_{s-1})\right\| b^{i-s} + \frac{\|\tau_0\|}{\|w_i\|}\, b^i\right) \le \frac{K}{\|\tau_i\|}\sum_{s=0}^{\infty} b^s.$$

The last inequality is valid if, for example, $\|w_s\| \le K\frac{1}{\delta}\|w_i\|$ for all $s \le i$, which is the case when z_i is free. □

6.3 Initial data for critical curves

The goal of this subsection is to verify the conditions on τ_0 in Proposition 6.1 for critical curves of all generations and all orders. Our plan of proof is as follows:

1. We obtain information on the slopes of critical curves of generation i by comparing them to critical curves of generation $i-1$. Following [BC2], this is done using a lemma of Hadamard, which requires that the intervals of definition of the critical curves be sufficiently long. We are thus led to the following question: on how long of a parameter interval can one continue a critical curve with reasonable properties?

2. As the order of a critical point tends to infinity, the length of the parameter interval on which it is defined goes to zero. This makes it necessary for us to prove our results in two steps, to first work with critical points having orders commensurate with their generations, and then to pass the bounds on to curves corresponding to higher orders.

6.3.1 Stability of critical regions

In Sections 3–5, we construct for $N = N_0, 3N_0, 3^2N_0, \cdots$ a parameter set Δ_N such that for $a \in \Delta_N$, $\Gamma_{\theta N}$ is well defined and consists of critical points of generation θN and order N. Let us denote this set by $\Gamma_{\theta N,N}$. In the discussion to follow, it will be convenient to consider $\Gamma_{i,n}$ for arbitrary $i \le n$. We define these sets formally as follows:

First we fix $a \in \Delta_N$, and define $\Gamma_{i,N}$, $\theta N < i \le N$, inductively by carrying out the steps in Section 5 in a slightly different order. Assuming that $\Gamma_{i-1,N}$ is defined and all the points in ∂R_0 are controlled for $i-1$ iterates, we define $\mathcal{C}^{(i)}$ and $\Gamma_{i,N}$.

Immediately, we observe that the newly constructed critical points are controlled by $\Gamma_{\theta N,N}$. In particular, they satisfy (IA2) and (IA4) (with possibly slightly weaker constants) and can be used for binding. For free segments of ∂R_i that lie in $\mathcal{C}^{(0)}$, we may then set binding as in the proof of Proposition 5.1, and proceed to step $i+1$.

For n with $N < n \le 3N$, let $\Delta_n := \{a \in \Delta_N$: (IA2) and (IA4) are satisfied up to time n for orbits from $\Gamma_{\theta N}\}$. A slight extension of the argument above defines $\Gamma_{i,n}$ for all $a \in \Delta_n$ and $i \le n$.

Finally, we introduce for each n the parameter set $\tilde{\Delta}_n$, which has the same definition as Δ_n except that in the definition of $\tilde{\Delta}_N$, $N = N_0, 3N_0, \cdots$, (IA2) and (IA4) are replaced by $d_{\mathcal{C}}(z_j) > \frac{1}{2}e^{-\alpha j}$ and $\|w_j^*\| > \frac{1}{2}e^{cj}$. One checks easily that all the results in Sections 3–5 are valid under these slightly relaxed rules, as is the discussion in the last two paragraphs, so that $\Gamma_{i,n}$ is defined for all $a \in \tilde{\Delta}_n$ and $i \le n$.

We remark before proceeding further that built into our definition of $\Gamma_{i,n}$ for $\frac{N}{3} < n < N$ is the property that $z_0 \in \Gamma_{i,n}$ has all the properties of $\tilde{z}_0 \in \Gamma_{\theta N,N}$ (except for the factor $\frac{1}{2}$) up to time n. In particular, Proposition 6.1 applies to $\hat{a} \in \tilde{\Delta}_n$ and $z_0 = z_0(\hat{a}) \in \Gamma_{i,n}$.

Definition 6.1 *For $i \le n$, an interval $J \subset \tilde{\Delta}_n$ and $\hat{a} \in J$, we say $\Gamma_{i,n}(\hat{a})$ has a* **smooth continuation** *to J if there is a map $g : \Gamma_{i,n}(\hat{a}) \times J \to R_0$ such that*

- $g(\cdot, a) = \Gamma_{i,n}(a)$ for all a and

- for each $z \in \Gamma_{i,n}(\hat{a})$, $a \mapsto g(z,a)$ is smooth.

Likewise one has the notion of the **critical regions $\mathcal{C}^{(i)}$ deforming continuously** *as a ranges over J.*

Lemma 6.4 *Let $\hat{a} \in \Delta_n$ and $J = [\hat{a} - \rho^{2n}, \hat{a} + \rho^{2n}]$. Then $J \subset \tilde{\Delta}_n$; moreover, $\Gamma_{n,n}(\hat{a})$ has a smooth continuation to J, and $\mathcal{C}^{(i)}$, $i \le n$, deform continuously on J.*

The structural stability of the critical regions comes from the fact that the components of $\mathcal{C}^{(i)}$ are stacked together in a very rigid way, and their relations to the components of $\mathcal{C}^{(i-1)}$ are equally rigid. As a varies over J, the entire structure may move up or down by amounts $>> b^{\frac{i}{2}}$, the maximum height of the components of $\mathcal{C}^{(i)}$, but it takes a relatively large horizontal displacement to slide these components past each other. A proof of Lemma 6.4 is given in Appendix B.10.

6.3.2 Comparing τ_0-vectors for different critical curves

Lemma 6.5 *There exists K such that the following holds for all n: Consider $\hat{a} \in \Delta_n$ and $J = [\hat{a} - \rho^{2n}, \hat{a} + \rho^{2n}]$. Let $z^{(n)} \in \Gamma_{n,n}(\hat{a})$, $z^{(n-1)} \in \Gamma_{n-1,n-1}(\hat{a}) \cap Q^{(n-1)}(z^{(n)})$, and let $z^{(n)}(a)$ and $z^{(n-1)}(a)$ be the continuations of $z^{(n)}$ and $z^{(n-1)}$ on J. Then*

$$\|\frac{dz^{(n)}}{da}(a) - \frac{dz^{(n-1)}}{da}(a)\| \ < \ (Kb)^{\frac{n}{9}}.$$

From this lemma it follows inductively that $\|\frac{dz^{(n)}}{da} - \frac{dz^{(0)}}{da}\| < Kb^{\frac{1}{9}}$ where $z^{(0)}$ is a critical point of generation 0 and order 1 lying in $Q^{(0)}(z^{(n)})$. Since there is only a finite number of critical curves of generation 0 and order 1, and for them $\tau_{0,2} = 0$, Lemma 6.5 proves that the hypotheses on τ_0 in Proposition 6.1 are met for curves corresponding to all $z^{(n)} \in \Gamma_{n,n}$. It remains to pass these properties to critical curves of higher order.

Lemma 6.6 *Let $m > n$, $\hat{a} \in \Delta_m$, and let $z^m \in \Gamma_{n,m}(\hat{a})$ be the updating of $z^n \in \Gamma_{n,n}(\hat{a})$ to order m. Then for all $a \in [\hat{a} - \rho^{2m}, \hat{a} + \rho^{2m}]$,*

$$\|\frac{dz^m}{da}(a) - \frac{dz^n}{da}(a)\| < (Kb)^{\frac{n}{4}}.$$

Lemmas 6.5 and 6.6 are proved in Appendix B.10.

6.4 Dynamics of critical curves

We fix a parameter interval J and a critical point z_0 which we assume can be smoothly continued to all of J. As usual, let $\gamma_i(a) = z_i(a)$. The purpose of this subsection is to make precise the parallel between the dynamics of $\gamma_0 \to \gamma_1 \to \gamma_2 \to \cdots$ and the action of T_a^i on ∂R_0. Let $\tau_i(a) = \frac{d\gamma_i}{da}(a)$.

Lemma 6.7 *Suppose Proposition 6.1 holds for all the parameters a and time indices i in question. Then there is a small number $k(\delta) > 0$ and an integer $i_0 > 0$ such that for all $i > i_0$, if $\gamma_i(a)$ is free and $\notin \mathcal{C}^{(0)}$, then $|\text{slope}(\tau_i(a))| < k(\delta)$ and $\tau_{i+1}(a) \approx DT_a(\gamma_i(a))\tau_i(a)$. In particular, if γ_i is as above, then it grows exponentially in length as long as it stays outside of $\mathcal{C}^{(0)}$.*

Proof: By Proposition 6.1, there is i_0 such that for all $i > i_0$, $\tau_i(a)$ is very close to $w_i(a)$ both in length and in angle. Assertion (i) is immediate; (ii) follows once $\|\tau_i\|$ is sufficiently large, and the last assertion is a consequence of (ii) and Lemma 2.8.

□

We assume (a, b) is sufficiently near $(a^*, 0)$ that $n_0 > i_0$. The reason we assert only that $|\text{slope}(\tau_i(a))| < k(\delta)$ (where $k(\delta) >> b$) is that for a very long period at the beginning – the length of this period depending on b – one cannot expect τ_i to be b-horizontal.

Next we allow γ_i to intersect $\mathcal{C}^{(0)}$. For each fixed a, we have introduced in Sections 3–5 definitions of distance to the critical set, binding point, bound period, etc. To emphasize their dependence on a, we write $d_{\mathcal{C}(a)}(\cdot)$, $\phi_a(\cdot)$ and $p_a(\cdot)$ when referring to definitions that belong to the map T_a. Even for a fixed map, these quantities depend sensitively on the location of the point in question; vertical displacements of z, for example, may dramatically change $\phi_a(z)$. In the "dynamics" of critical curves, the problem is all the more delicate, for not only does $z_i(a)$ move with a, the entire critical set moves as well. The goal of the next few lemmas is to establish some viable notions of $d_{\mathcal{C}}(\cdot)$ and bound/free states that work in a coherent fashion for all points in γ_i.

We assume for the rest of this subsection that

(i) $J \subset \tilde{\Delta}_{K\alpha n}$, so that for each a the binding structure is in place for points with $d_{\mathcal{C}(a)}(\cdot) > e^{-\alpha n}$;

(ii) z_0 obeys (IA2) and (IA4) up to time n, and

(iii) all time indices are $\leq n$.

In the next lemma, we let $|\cdot - \cdot|_h$ denote the horizontal distance between two points, and assume for simplicity that γ_i is contained in one component of $\mathcal{C}^{(0)}$.

Lemma 6.8 *Suppose* $|\text{slope}(\tau_i)| < k(\delta)$*. Then there exists* $\bar{z} \in \mathcal{C}^{(0)}$ *such that whenever* $d_{\mathcal{C}(a)}(\gamma_i(a)) > \frac{1}{2}e^{-\alpha i}$*,*

$$\big||\gamma_i(a) - \bar{z}|_h - d_{\mathcal{C}(a)}(\gamma_i(a))\big| < Ke^{-ci} d_{\mathcal{C}(a)}(\gamma_i(a)).$$

Thus we may put the partition $\mathcal{P}_{\gamma_i,\bar{z}}$ on γ_i and define $d_{\mathcal{C}}(\cdot) = |\cdot - \bar{z}|_h$ (the precise definition of $d_{\mathcal{C}}(\gamma_i(a))$ is irrelevant for a with $d_{\mathcal{C}(a)}(\gamma_i(a)) < \frac{1}{2}e^{-\alpha i}$).

Lemma 6.9 *Let* γ_i *be as above. We assume further that* $z_i(a)$ *is a free return for every* a*. Then for each* $\omega_0 = I_{\mu j} \in \mathcal{P}_{\gamma_i,\bar{z}}$ *with* $|\mu| < \alpha i$*, there exists* $\tilde{p} = \tilde{p}(\omega_0) < K|\mu|$ *such that for all* a*,* a' *with* $z_i(a), z_i(a') \in \omega_0$*,*

(a) $|z_{i+j}(a) - z_{i+j}(a')| < e^{-\beta j}$ *for* $j \leq \tilde{p}$*;*

(b) $z_{i+\tilde{p}}$ *is out of all fold periods,* $|\text{slope}(\tau_{i+\tilde{p}})| < k(\delta)$ *and* $|\omega_{\tilde{p}}| \geq \frac{1}{\mu^2}e^{-\beta K|\mu|}$*;*

(c) $\|w_{i+\tilde{p}}\| > K^{-1}e^{\frac{\tilde{p}}{3}}\|w_i\|$*, and* $\|\tau_{i+\tilde{p}}\| > K^{-1}e^{\frac{\tilde{p}}{3}}\|\tau_i\|$*.*

Lemma 6.9 allows us to define a natural notion of bound/free states for the curves γ_i that agrees essentially with the dynamical notion previously defined for each $z_i(a)$.

Proposition 6.2 *We assume the following hold for all* $a \in J$ *and* $i \leq n$*:*

(i) for each i*, the entire segment* γ_i *is bound or free simultaneously, and* γ_i *is contained in three contiguous* $I_{\mu j}$*'s at all free returns;*

(ii) γ_n *is a free return.*

Then there exists K *(independent of* γ_0 *or* n*) such that for all* $a, a' \in J$*,*

$$\frac{1}{K} \leq \frac{\|\tau_n(a)\|}{\|\tau_n(a')\|} \leq K.$$

Lemmas 6.8 and 6.9 are proved in Appendix B.11. Proposition 6.2 is proved in Appendix B.12.

6.5 Deletions on account of a single critical point

Let J be a parameter interval on which $\Gamma_{\theta N,\, 3K\alpha N}$ is well defined, that is to say, for each (fixed) $a \in J$, all critical points of generation $\leq \theta N$ have been introduced, and they obey (IA2) and (IA4) up to time $3K\alpha N$. Moreover, we assume that all of the critical points have smooth continuations on J. In this subsection we follow the evolution of one fixed $z_0 \in \Gamma_{\theta N,\, 3K\alpha N}$ and consider the set of $a \in J$ that will be excluded on account of its behavior between times N and $3N$.

Let a and z_0 be fixed. Before embarking on the main discussion, let us first review what can be said about z_n and $w_n = DT_a^n(z_0)\binom{0}{1}$ for $n \leq 3N$ based on the information available. (The precise location of z_0 will have to be "updated" as we go along; these issues have been dealt with in previous subsections and will not be discussed here.)

(i) There is no ambiguity as to whether $d_{\mathcal{C}}(z_n) > e^{-\alpha n}$.

(ii) If $z_n \in \mathcal{C}^{(0)}$ and $d_{\mathcal{C}}(z_n) > e^{-\alpha n}$, then it has a binding point and the ensuing bound period has the properties in (IA5). (This uses the fact that $p < 3K\alpha N$.)

(iii) If $d_{\mathcal{C}}(z_i) > e^{-\alpha i}$ for all $i < n$, then (z_0, w_0) is controlled in the sense of Definition 4.2 up to time n. (This follows from the proof of Proposition 5.2. Notice that the argument uses only the critical structures guaranteed above.)

(iv) If for some subinterval $J' \subset J$, every $a \in J'$ has the property that z_0 obeys the estimate in (IA2) up to time n, then the discussion in Sect. 6.4 holds for the critical curve γ_i defined on J' for $i < n$. (In Sect. 6.4 we have assumed for simplicity that z_0 obeys the estimate in (IA4); we can do without that because the estimate $\|w_i^*\| \geq e^{\frac{c}{3}i}$, which we have from (IA2) alone, will also suffice.)

Observe (1) $\Gamma_{\theta N,\, 3K\alpha N}$ is essentially the minimum structure needed for the discussion above, and (2) this discussion is entirely independent of the behavior after time $3K\alpha N$ of critical points other than z_0. It is this *independence* that allows us to consider one critical point at a time up to time $3N$ and to make deletions on the basis of its behavior alone. On the other hand, the *dependence* on early behavior of other critical points is a strong reminder that distinct critical orbits cannot be treated completely separately through their entire lifetimes.

We now proceed to the main topic of discussion. We assume for the rest of this subsection that J is as above, that for all $a \in J$, the estimates in (IA2) and (IA4) hold for z_0 up to time N, and that γ_N, which is a free return, is $\approx I_{\mu j}$ for some μ with $|\mu| < \alpha N$.

We begin with deletions on account of (IA2). At this point, we ask the reader to go to Sect. 6.1, and to consider γ_N in the place of ω_N in the model phase-space problem. The construction we make in the parameter case is identical to that in Sect. 6.1. For the analogy between the dynamics of critical curves and true dynamical curves, see Lemma 6.7 (outside $\mathcal{C}^{(0)}$), Lemmas 6.8 and 6.9 (bound estimates) and Proposition

6.2 (distortion). We summarize the result in the following proposition, the proof of which we omit.

Proposition 6.3 *There is a set* $D_{N,z_0} \subset J$ *with*

$$|D_{N,z_0}| \ < \ Ke^{-\frac{1}{2}\alpha N}|J|$$

such that for all $a \in J \setminus D_{N,z_0}$*, the estimate in (IA2) holds for* z_0 *up to time* $3N$*.*

The structure of $J \setminus D_{N,z_0}$, which will be relevant in Sect. 6.6, can be described as follows: In the procedure outlined in Sect. 6.1, pulling back the subdivisions at each stage to the parameter interval J results in a partition defined on the subset of J that has not yet been deleted at that time. Let us call these partitions $\mathcal{Q}_{n,z_0}$, $n = N, N+1, \cdots, 3N$. Each element J' of $\mathcal{Q}_{n,z_0}$ is an interval. One step later, J' may again be an element of $\mathcal{Q}_{n+1,z_0}$, or it may be subdivided into shorter subintervals some of which may be discarded. For all $a \in J'$, $z_i(a)$ can be thought of as having "indistinguishable" itineraries up to time n, that is to say, for each $i \leq n$, the critical curve γ_i defined on J' is either entirely outside of $\mathcal{C}^{(0)}$ or entirely contained in some $I_{\mu j}$, and all points are either in a bound state or in a free state simultaneously. Moreover, at its last free return before time n, γ_i occupies the full length of some $I_{\mu j}$.

Finally we move on to deletions on account of (IA4). We use the construction above, deleting those elements of $\mathcal{Q}_{3N,z_0}$ that correspond to z_i having an abnormally high frequency of close returns between times N and $3N$.

Proposition 6.4 *Given* $\varepsilon > 0$, $\exists \delta_0 = \delta_0(\varepsilon)$ *such that if* $\delta < \delta_0$ *and the parameters in question are sufficiently near* $(a^*, 0)$ *(depending on* δ*), then the following holds: If* J*,* z_0 *and* γ_N *are as above, then there is a set* $E_{N,z_0} \subset J \setminus D_{N,z_0}$ *with*

$$|E_{N,z_0}| < e^{-\varepsilon n}|J \setminus D_{N,z_0}|$$

such that for all $a \in J \setminus (D_{N,z_0} \cup E_{N,z_0})$*, the estimates in (IA2) and (IA4) hold for* z_0 *up to time* $3N$*.*

As is evident from its formulation, this estimate is a little more delicate than the previous one. A one-dimensional version of this result is proved in [BC2], pages 81-86. (For an alternate proof, see [TTY].) After the discussion at the beginning of this subsection and the groundwork in Sect. 6.4, the adaptation of this result to our setting is straightforward.

6.6 Estimating $|\Delta|$

The initial parameter set Δ_0 is chosen as follows. Let $C = \{x_i\}$ be the critical set of f, and let δ_1 be the minumum distance between C and $f^n x_i$, $n > 0$. We assume that $\delta_1 >> \delta$. Let n_0 be the number of iterates the critical orbits of $T_{a,b}$ are required to stay outside of $\mathcal{C}^{(0)}$. We assume n_0 is as large as need be and prespecified. Then there exists $\varepsilon > 0$ such that for all $a \in [a^* - \varepsilon, a^* + \varepsilon]$, the first n_0 iterates of all the critical points of f_a stay $> \frac{\delta_1}{2}$ away from C. Choose b so small that the same holds (with slightly weaker estimates) for all the generation 0 critical points. We let $\Delta_0 = [a^* - \varepsilon, a^* + \varepsilon]$, and let b, which may be shrunk further for other reasons, be fixed in the rest of the discussion.

We first give a rough outline of the inductive process by which the parameter set Δ is chosen. With this outline in mind, we will discuss in greater detail each individual step and then finally estimate the measure of the set of parameters deleted.

Here is the outline. Starting with $N = n_0$, the procedure for going from step N to step $3N$ is as follows: Let $N_0 = [\frac{1}{\theta}]$. Consider first the case $N \leq N_0$. At time N, we are handed a good parameter set Δ_N. For $a \in \Delta_N$, we consider each $z_0 \in \Gamma_0$ separately until time $3N$, making deletions if necessary so that the estimates in (IA2) and (IA4) are obeyed. Let Δ_{3N,z_0} be the set of parameters retained by considering z_0 alone. Then $\Delta_{3N} = \cap_{z_0 \in \Gamma_0} \Delta_{3N,z_0}$. The critical set Γ_1 is created the first time $3N$ exceeds N_0. In general, for $N > N_0$, the procedure is as above with $\Gamma_{\theta N}$ in the place of Γ_0, plus an extra step at the end, namely the creation of $\Gamma_{3\theta N}$ for $a \in \Delta_{3N}$. This process is continued *ad infinitum*, and $\Delta := \cap_N \Delta_N$.

We now begin our detailed discussion. Let $z_0 \in \Gamma_0$ be fixed, and let γ_i denote its associated critical curve. We wish to argue that as we "iterate", γ_i grows long, is roughly horizontal, and the first time it intersects $\mathcal{C}^{(0)}$ nontrivially, the intersection contains at least one of the outermost I_μ. To see that this can be arranged, consider first the case $b = 0$. There, the fact that f_{a^*} is a Misiurewicz map implies that γ_i is "hooked" onto an orbit that remains $> \delta_1$ away from the critical set for all i, giving the desired picture. Fix $n_1 > n_0$ such that γ_i remains outside of $\mathcal{C}^{(0)}$ for $i \leq n_1$ and $|\gamma_{n_1}| >> \delta$. This continues to hold for small $b > 0$. Also, for $b > 0$, it follows from Proposition 6.1 and Lemma 6.3 that there is a time after which w_i and τ_i become comparable both in magnitude and in angle. By choosing n_0 sufficiently large, therefore, we may assume that the first time γ_i meets $\mathcal{C}^{(0)}$, γ_i is a roughly horizontal curve with $|\gamma_i| > \delta$, and the part deleted (in violation of (IA2)) constitutes as small a fraction of Δ_0 as need be.

Let N be the first time when some deletion has taken place for at least one of the z_0. We stated in the outline that we are handed Δ_N, but we actually have more, namely that for each $z_0 \in \Gamma_0$, we have Δ_{N,z_0}, the set of parameters retained by considering z_0 alone up until time N, and a partition $\mathcal{Q}_{N,z_0}$ on Δ_{N,z_0} obtained as in Sect. 6.5. (If no deletion has taken place for this z_0, then $\Delta_{N,z_0} = \Delta_0$ and $\mathcal{Q}_{N,z_0}$ is the trivial partition.) Let $\Delta_N = \cap_{z_0 \in \Gamma_0} \Delta_{N,z_0}$. We describe next how to go from step N to step $3N$.

Fix z_0 and $J \in \mathcal{Q}_{N,z_0}$. If $J \cap \Delta_N = \emptyset$, then we declare J to be "inactive" from here on and do not consider it further. (For purposes of estimating the measure of the set of *deleted* parameters, however, we regard all the inactive intervals as being in Δ_{n,z_0} for all $n \geq N$.) Assume $J \cap \Delta_N \neq \emptyset$. This does not mean $J \subset \Delta_N$, for other critical points may have created some "holes" in J. Let $\hat{\gamma}_i$ denote the critical curve defined on J and let $n > N$ be the first time when part of $\hat{\gamma}_n$ makes a free return to $\mathcal{C}^{(0)}$. In order to continue, we need to verify that the necessary binding structure is available. Since $|\hat{\gamma}_n| < 1$, we have $|J| < \lambda_1^{-1}e^{-cn}$ by Proposition 6.1; and since $\lambda_1^{-1}e^{-cN} << \rho^{6K\alpha N}$, we are guaranteed by Lemma 6.4 that $J \subset \tilde{\Delta}_{3K\alpha N}$. The lemmas in Sect. 6.4 therefore apply to give us a meaningful notion of $d_{\mathcal{C}}(\cdot)$ (see also Sect. 6.5). We subdivide according to $I_{\mu j}$-locations, defining $\mathcal{Q}_{n,z_0}$. For those $J' \in \mathcal{Q}_{n,z_0}$ that do not meet Δ_N, we again declare them to be "inactive", and we track the active ones following the discussions in Sects. 6.4 and 6.5. The process is continued until time $3N$. It is then repeated for each one of the other critical points in Γ_0.

At a generic step N, then, we are handed for each $z_0 \in \Gamma_0$ or $\Gamma_{\theta N}$ a set Δ_{N,z_0}, which has an active part and an inactive part. On the active part there is a partition $\mathcal{Q}_{N,z_0}$. We track the elements of $\mathcal{Q}_{N,z_0}$, declaring some to be "inactive" along the way and making deletions to secure (IA2) and (IA4) as in Sect. 6.5. At time $3N$, we create $\Gamma_{3\theta N}$ if necessary. The newly created critical points are handed the parameter sets and partitions of their parents. The completes the desciption of Δ_N for all $N = 3^i n_0$.

We turn, finally, to the problem of estimating the measure of Δ. From Sect. 6.5, it follows that there exists $\alpha_1 > 0$ such that for each critical point z_0, $|\Delta_{N,z_0} \setminus \Delta_{3N,z_0}| < Ke^{-\alpha_1 N}|\Delta_0|$. (This estimate would not have been valid if we had deleted inactive intervals.) Adding up the deletions from all the critical points, we have

$$|\Delta_0 \setminus \Delta| \ \leq \sum_{i:3^i n_0 \leq N_0} \text{card}(\Gamma_0) Ke^{-\alpha_1 n_0}|\Delta_0| \ + \ \sum_{i=1}^{\infty} \text{card}(\Gamma_{3^i \theta N_0}) Ke^{-\alpha_1 3^i N_0}|\Delta_0|.$$

To estimate $\text{card}(\Gamma_{\theta N})$, let $I_1, \cdots, I_r$ be the monotone intervals of f, and let $K_0 = \max_i\{$ number of I_j counted with multiplicity : $I_j \cap f(I_i) \neq \emptyset\}$.

Lemma 6.10

$$\text{card}(\Gamma_{\theta N}) < K_0^{\theta N}$$

Proof Partition ∂R_k into segments by orbits of critical points of generation $\leq k$. Then each segment has at most one free component, and each free component meets $\leq K_0$ of the monotone intervals, giving rise to $\leq K_0$ new critical points. For more details, see Sect. 9.1. □

We conclude that the fraction of Δ_0 deleted tends to 0 as $n_0 \to \infty$ and $b \to 0$.

In the remainder of this paper, T is assumed to be $T_{a,b}$ where (a,b) is a pair of "good" parameters, i.e. $(a,b) \in \Delta$ where Δ is as in Theorem 1.

7 Nonuniform Hyperbolic Behavior

Recall that Γ is the set to which $\Gamma_{\theta N}$ converges as $N \to \infty$. One of the properties guaranteed by parameter selection is that orbits starting from Γ have some hyperbolic behavior (Theorem 1(2)(ii)). The purpose of this section is to show that this behavior is passed on to a large set of points on the attractor and in the basin, proving Theorem 2 except for the assertion in (1)(iii), the proof of which we postpone to Sect. 10.4.

7.1 Control and hyperbolicity of non-critical orbits

We recapitulate the ideas developed in Sections 3–5 with a view toward proving hyperbolicity for an arbitrary (non-critical) orbit. Given arbitrary $z_0 \in R_0$, we let

$$0 \;\leq\; n_1 \;<\; n_1 + p_1 \;\leq\; n_2 \;<\; n_2 + p_2 \;\leq\; n_3 \;<\; \cdots$$

be such that $z_{n_j} \in \mathcal{C}^{(0)}$ and is bound to a suitable point in Γ, p_j is the ensuing bound period, and n_{j+1} is the first return after $n_j + p_j$. Then:

(1) During its free periods, i.e. between times $n_j + p_j$ and n_{j+1}, the orbit is outside of $\mathcal{C}^{(0)}$, where DT^i is essentially uniformly hyperbolic (Lemma 2.8).

(2) During its bound periods, i.e. between times n_j and $n_j + p_j$, $DT^i(z_{n_j})$ copies the derivative of its guiding orbit from Γ (see (IA6)), which has been guaranteed through parameter selection to have some form of hyperbolicity ((IA4)).

(3) The concatenation of hyperbolic segments, however, need not result in a hyperbolic orbit, for the direction expanded at the end of one segment may be near the contractive direction of the next. Indeed, this happens at times n_j, when there is a "confusion" of stable and unstable directions, leading to a loss of hyperbolicity (see Sect. 3.1).

(4) The properties that guarantee that hyperbolicity is preserved through these concatenations are precisely the *h-relatedness* and *correct splitting* properties at free returns. At time n_j, the correct splitting of an expanded vector limits the magnitude of the loss (Lemma 2.12 and Sect. 3.3.2), while the h-relatedness of z_{n_j} to some $\hat{z} \in \Gamma$ guarantees that the ensuing bound period is long enough for this loss to be compensated (see (IA5)).

In particular, if z_0 has a unit tangent vector w_0 such that (z_0, w_0) is *controlled* by Γ for all $n \geq 0$ in the sense of Definition 4.2, then

$$\limsup_{n\to\infty} \frac{1}{n} \log \|DT^n(z_0)w_0\| \geq c' > 0 \tag{10}$$

where $e^{c'}$ is the minimum of the growth rates of b-horizontal vectors outside of $\mathcal{C}^{(0)}$ and net derivative gains during bound periods. Assuming that the rate of growth outside of $\mathcal{C}^{(0)}$ is $> e^{\frac{c}{3}}$ where c is as in Theorem 1, we may take $c' = \frac{c}{3}$. We remark that in general, the growth of $\|DT^n(z_0)w_0\|$ is not regular: without any assumptions on how close to Γ the free returns are allowed to be, i.e. without a condition in the spirit of (IA2), the loss of hyperbolicity at time n_j can be arbitrarily large; for example, the lim inf in (10) can be negative.

Recall that to establish control of (z_0, w_0), it suffices to look at free returns (Lemmas 4.2 and 4.5). We record below a condition at free returns that enables us to extend control through another bound-free cycle. Lemma 7.1 plays a crucial role in all the results in this section. First, we identify certain locations that are potentially problematic. For $k \geq 0$, let

$$Z^{(k)} := \{z \in \mathcal{C}^{(k)} : \ d_{\mathcal{C}}(z) < b^{\frac{k}{20}}\}.$$

Lemma 7.1 *Let z_0 and w_0 be arbitrary, and suppose that (z_0, w_0) is controlled by Γ up to time $k-1$. Let z_k be a free return. If $z_k \in \mathcal{C}^{(i)} \setminus Z^{(i)}$ for some $i < \frac{5}{4}k$, then w_k splits correctly.*

Proof: The proof of this lemma is virtually identical to that of Proposition 5.2. Let $j = \min\{i, k\}$, so that z_{k-j} makes sense. (The reason we allow i to exceed k has to do with the way this lemma is used.) Claims 5.1-5.3 in Proposition 5.2 continue to be valid becuase they rely only on the fact that (z_0, w_0) is controlled. The proof here differs from that in Section 5 only at the end, where under present conditions we have

$$b^{\frac{j'}{4}} \ \leq \ b^{\frac{j}{12}} \ \leq \ b^{\frac{1}{12}\frac{4}{5}i} \ << \ b^{\frac{i}{20}} \ \leq \ d_{\mathcal{C}}(z_k).$$

$\square$

7.2 Typical derivative behavior in the basin

Let m denote the 2-dimensional Lebesgue measure.

Proposition 7.1 *Assuming the additional regularity condition (**) in Sect. 1.2, we have*

$$m \ \{z_0 \in R_0 : z_k \in Z^{(k)} \text{ infinitely often}\} \ = 0.$$

To prove this result, we need more refined estimates on the width of $Q^{(k)}$ than that given in Lemma 4.1.

Lemma 7.2 *There exists $K > 0$ such that if $Q^{(k)}$ is a component of $\mathcal{C}^{(k)}$, and d_v is the vertical distance between the two horizontal boundaries of $Q^{(k)}$ measured anywhere along the length of $Q^{(k)}$, then*

$$(K^{-1}b)^{k+1} \ < \ d_v \ < \ (Kb)^{\frac{99}{100}k}.$$

Proof: First we prove the lower bound, which relies heavily on the condition (**). Let ω_k be a vertical line segment joining two points in $\partial Q^{(k)}$. For $i < k$, let $\omega_i = T^{-k+i}\omega_k$. If ω_0 connects the two components of ∂R_0, then $d_v > (K^{-1}b)^k \cdot K^{-1}b$ since by (**), $\|DTv\| \geq K^{-1}|\det(DT)| \geq K^{-1}K_1^{-1}b$ for every unit vector v. If not, we will need to rule out the possibility that ω_0 may be extremely short. Let $z_0, z_0' \in \omega_0 \cap \partial R_0$, and let γ_0 be the shorter of the two segments of ∂R_0 between z_0 and z_0'. We consider $\gamma_i := T^i\gamma_0$, and remember that points on ∂R_0 together with their tangent vectors are controlled (Proposition 5.1). Since z_k and z_k' are both free, and they do not lie on a $C^2(b)$-curve, we conclude that a critical point is created on γ_i for some $i < k$. Let i be the first time this happens. If $|z_i - z_i'| > \delta$, then $|\omega_k| > \delta(K^{-1}b)^k$. If not, then both z_i and z_i' are in $\mathcal{C}^{(0)}$. Since both of their bound periods have expired by time k, it follows from (IA5) that $d_{\mathcal{C}}(z_i)$ and $d_{\mathcal{C}}(z_i')$ are $> e^{-K(k-i)}$. We claim that $d_{\mathcal{C}}(z_i) + d_{\mathcal{C}}(z_i')$ is approximately the horizontal distance between these two points (see Lemma 9.1 for more details). This gives $|\omega_k| > 2(e^{-K}K^{-1}b)^k$.

For the upper estimate, we pick an arbitrary $z_k \in \partial Q^{(k)}$, and borrow the argument in the proof of Claim 5.1 with $j = k$, pivoting the line L at $L \cap \{x = \frac{1}{100}k\}$ (instead of $L \cap \{x = k - \frac{1}{3}j\}$) as we rotate clockwise. This gives i_0 with $0 \leq i_0 \leq \frac{1}{100}k$ such that $\|DT^i(z_{i_0})\| > \|DT\|^{-100i}$. Iterating forward once if necessary (and possibly losing a factor of K^{-1} in the last estimate), we may assume that $z_{i_0} \notin \mathcal{C}^{(0)}$, so that it lies on an integral curve γ_0 of e_{k-i_0} which joins the two components of ∂R_0. Note that γ_0 meets ∂R_0 only at its end points. Iterating forward, this curve brings in two segments of ∂R_{k-i_0}. They must lie on the two horizontal boundaries of $Q^{(k-i_0)}(z_k)$ because γ_{k-i_0} passes through z_k and intersects no other point of ∂R_{k-i_0}. This proves that d_v measured at z_k has length at most that of γ_{k-i_0}, which by Lemma 2.3 is $< (\|DT\|^{200}b)^{k-i_0} < (Kb)^{\frac{99}{100}k}$. □

Proof of Proposition 7.1: By the Borel-Cantelli Lemma, it suffices to show that $\sum_k m(T^{-k}Z^{(k)}) < \infty$. We estimate $m(T^{-k}Z^{(k)})$ by

$$\begin{aligned} m(T^{-k}Z^{(k)}) &= \sum m(T^{-k}(Q^{(k)} \cap Z^{(k)})) \\ &\leq \max \frac{m(T^{-k}(Q^{(k)} \cap Z^{(k)}))}{m(T^{-k}Q^{(k)})} \sum m(T^{-k}Q^{(k)}) \end{aligned}$$

where the summations and maximum are taken over all components $Q^{(k)}$ of $\mathcal{C}^{(k)}$. Note also that $\sum m(T^{-k}Q^{(k)}) < 1$. Using Lemma 7.2 and the regularity of $\det(DT)$ in (**), we obtain

$$\begin{aligned} \frac{m(T^{-k}(Q^{(k)} \cap Z^{(k)}))}{m(T^{-k}Q^{(k)})} &\leq K^{2k} \cdot \frac{m(Q^{(k)} \cap Z^{(k)})}{m(Q^{(k)})} \\ &\leq K^{2k} \cdot \frac{(Kb)^{\frac{99}{100}k} \cdot b^{\frac{1}{20}k}}{(\frac{b}{K})^{k+1} \cdot \rho^k} \leq K^{4k}\frac{1}{b} \cdot \frac{b^{\frac{1}{25}k}}{\rho^k} \end{aligned}$$

which decreases geometrically in k as desired. □

Proof of Theorem 2(2): Let $\xi_0 \in R_0$. From the discussion in Sect. 7.1, it follows that
$$\limsup_{n\to\infty} \frac{1}{n} \log \|DT^n(\xi_0)\| \geq \frac{c}{3}$$
holds if we are able to produce $k_0 > 0$ and a vector w_0 such that if $z_0 = \xi_{k_0}$, then (z_0, w_0) is controlled by Γ for all $n \geq 0$. In light of Proposition 7.1, it suffices to consider the following two cases.

Case 1. $\xi_k \notin Z^{(k)}$ for all $k \geq 0$. We take $k_0 = 0$ and let $w_0 = \binom{1}{0}$ if $\xi_0 \notin \mathcal{C}^{(0)}$, $w_0 = \binom{0}{1}$ if $\xi_0 \in \mathcal{C}^{(0)}$. We assume (z_0, w_0) is controlled up to time $k-1$, and let z_k be a free return. The hypothesis of Lemma 7.1 is verified at time k as follows: Let j be the largest integer such that $z_k \in \mathcal{C}^{(j)}$. Then if $j \geq k$, $i = k$ meets the requirements of Lemma 7.1 since $\xi_k \notin Z^{(k)}$; and if $j < k$, then z_k must be in $\hat{Q}^{(j+1)} \setminus Q^{(j+1)}$ for some $Q^{(j+1)}$ since it is in R_k, and so we may take $i = j+1$.

Case 2. $\xi_{k_0} \in Z^{(k_0)}$ for some k_0 and $\xi_k \notin Z^{(k)}$ for all $k > k_0$. Here we let $z_0 = \xi_{k_0}$ and $w_0 = \binom{0}{1}$. There is a critical point $\hat{z}$ in $Q^{(k_0)}(z_0)$ to which z_0 is bound for k_1 iterates. Since $\|DT\|^{k_1} b^{\frac{k_0}{20}} > e^{-\beta k_1}$, we have $k_1 \sim k_0\theta^{-1} >> k_0$. During this period, we may regard (z_0, w_0) as controlled by Γ. For $k \geq k_1$, the situation is identical to that in Case 1 except that $z_k \in R_{k+k_0}$ and we can only guarantee $z_k \notin Z^{(k+k_0)}$. To verify the hypothesis of Lemma 7.1 for z_k, we proceed as above, distinguishing between the cases $j \geq k + k_0$ and $j < k + k_0$ and noting that for $k \geq k_1$, $k + k_0 < (1 + K\theta)k$. $\square$

Remark. The results in this paper that use (**) remain valid if (**) is replaced by

$(**)'$ There exist $\eta \geq 1$ and $K_1, K_2 > 0$ such that for all $z \in R_0$,
$$K_1^{-1} b^\eta \leq |\det(DT)| \leq K_2 b^\eta.$$

To prove this, it suffices to check that Proposition 7.1 is valid under $(**)'$. Observe that the results in Sect. 2.1 are abstract, so that if $\|DT^i(z_0)\| \geq \kappa^i$ for all $i \leq n$, then $\|DT^i e_n\| \leq (Kb^\eta \kappa^{-2})^i$ for all $i \leq n$. Using this and $\|DTv\| \geq K^{-1} b^\eta$ for all $\|v\| = 1$, one checks easily that under $(**)'$, the conclusion of Lemma 7.2 is valid if b is replaced by b^η. Moreover, the number $\frac{99}{100}$ can be replaced by $1 - \varepsilon_0$ for any prespecified $\varepsilon_0 > 0$. Choosing ε_0 such that $\varepsilon_0 \eta < \frac{1}{20}$, we check that the proof of Proposition 7.1 goes through as is.

7.3 Uniform hyperbolicity away from $\mathcal{C}$

Recall that
$$\Omega_\varepsilon = \{z_0 \in \Omega : \quad d_{\mathcal{C}}(z_n) \geq \varepsilon \ \text{ for all } n \in \mathbb{Z}\}.$$

The purpose of this subsection is to prove that Ω_ε is a uniformly hyperbolic invariant set [9] for every $\varepsilon > 0$. This result together with the fact that the strength of hyperbolicity deteriorates as $\varepsilon \to 0$ justifies our identification of $\mathcal{C}$ as the critical set and

[9]Technically, $z^i \to z$ does not imply $d_{\mathcal{C}}(z^i) \to d_{\mathcal{C}}(z)$ when $z^i \notin \mathcal{C}^{(k)}$ and $z \in \mathcal{C}^{(k)}$, but let us assume Ω_ε is closed by taking its closure if necessary.

confirms that $d_{\mathcal{C}}(\cdot)$ is a valid notion of "distance" to the critical set. The approximation of Ω by Ω_ε is a concrete example of the use of uniformly hyperbolic invariant sets to approximate systems that have (weak) hyperbolic properties. See [K] and [P] for results in the same spirit.

Proofs of uniform hyperbolicity often rely on *a priori* knowledge of invariant cones. In our setting, these cones are easily identified for Ω_ε with $\varepsilon > \sqrt{b}$; see Sect. 2.5. As $\varepsilon \to 0$, the situation becomes considerably more delicate: the stable and unstable directions at points in Ω_ε become increasingly confused, both ranging over nearly all possible directions within very small neighborhoods. Our line of proof, which does not rely on *a priori* knowledge of cones, can be formulated as follows:

Let $g : X \to X$ be a self-map of a compact metric space, and let $M : X \to GL(2,\mathbb{R})$ be a continuous map. For $x \in X$ and $n \geq 0$, we define $M^{(n)}(x) = M(g^{n-1}x)\cdots M(gx)M(x)$ and $M^{(-n)}(x) = M(g^{-n}x)^{-1}\cdots M(g^{-1}x)^{-1}$. It is clear what it means for the *cocycle* $(g, M^{(n)})$ to be *uniformly hyperbolic* (think of g as a diffeomorphism and $M(x) = Dg(x)$). Since the condition of interest to us is projective in nature, we will state our result assuming that M takes its values in $SL(2,\mathbb{R})$.

Lemma 7.3 *Let $(g, M^{(n)})$ be as above. If there exist $\lambda > 1$ and $N \in \mathbb{Z}^+$ such that at each $x \in X$, there exists a unit vector $v = v(x)$ such that*

$$\|M^{(n)}(x)v\| \leq \lambda^{-n} \quad \text{for all} \quad n \geq N,$$

then $(g, M^{(n)})$ is uniformly hyperbolic.

Proof: Let $E^s(x)$ be the subspace spanned by $v(x)$, and observe that $M(x)E^s(x) = E^s(gx)$: if not, then there are two linearly independent vectors, $v_1 \in M(x)E^s(x)$ and $v_2 = v(gx)$ such that both $\|M^{(n)}(gx)v_1\|$ and $\|M^{(n)}(gx)v_2\|$ decrease exponentially as $n \to \infty$, contradicting $M \in SL(2,\mathbb{R})$. The continuity of $x \mapsto E^s(x)$ is proved similarly.

Using the uniform contraction of $M^{(N)}$ on vectors in E^s and the fact that $|\det(M)| = 1$, we choose $\delta_0 > 0$ such that for all $x \in X$ and $w \neq 0 \in \mathbb{R}^2$, if $\angle(w, v(x)) < \delta_0$, then $\angle(M^{(N)}(x)w, v(g^N x)) > \frac{1}{2}\lambda^{2N}\angle(w, v(x))$. Let $C^s(x) = \{w : \angle(w, v(x)) < \delta_0\}$ and $C^u(x) = \mathbb{R}^2 \setminus C^s(x)$. We claim that $E^u(x) := \cap_{n=1}^{\infty} M^{(nN)}(g^{-nN}x)C^u(g^{-nN}x)$ is a 1-dimensional subspace. This follows if we show that $\|M^{(-nN)}w\|$ decreases exponentially as $n \to \infty$ for $w \in E^u$. The latter is a consequence of the following: $\angle(M^{(-nN)}(x)w, M^{(-nN)}(x)v(x)) > \delta_0$, $\|M^{(-nN)}(x)v(x)\| \geq \lambda^{nN}$, and $M \in SL(2,\mathbb{R})$. The M-invariance of E^u is checked easily. □

Proposition 7.2 *For every $\varepsilon > 0$, Ω_ε is uniformly hyperbolic with*

$$\|DT^i u\| \;\geq\; K_\varepsilon^{-1} e^{c' i}$$

for all $u \in E^u$. Here K_ε is a constant depending on ε, and c' can be taken to be $\approx \frac{c}{3}$.

Proof: We fix ε and let k_ε be the smallest integer k such that $\varepsilon > b^{\frac{k}{20}}$.

Claim 7.1 *For every $\xi_0 \in \Omega_\varepsilon$, there exists $k(\xi_0) \le 2k_\varepsilon$ and a unit vector w_0 such that if $z_0 = \xi_{k(\xi_0)}$, then for all $i > 0$,*

$$\|DT^i(z_0)w_0\| \;\ge\; e^{\frac{c}{3}i} b^{\frac{k_\varepsilon}{20}} K^{-\frac{k_\varepsilon}{10}}.$$

Proof of Claim 7.1: We consider separately the following cases:

Case 1. $\xi_i \notin \mathcal{C}^{(0)}$ for all $i \le k_\varepsilon$. In this case we let $k(\xi_0) = 0$ and $w_0 = \binom{1}{0}$.

Case 2. $\xi_{i_0} \in \mathcal{C}^{(0)}$ for some $i_0 \le k_\varepsilon$ and $\xi_{i_0+k} \notin Z^{(k)}$ for all $k \ge 0$. We let $k(\xi_0) = i_0$ and $w_0 = \binom{0}{1}$.

Case 3. $\xi_{i_0} \in \mathcal{C}^{(0)}$ for some $i_0 \le k_\varepsilon$ and $\xi_{i_0+k} \in Z^{(k)}$ for some $k \ge 0$. We let k be the last time this happens, and choose $k(\xi_0) = i_0 + k$, $w_0 = \binom{0}{1}$. Note that $k(\xi_0) \le 2k_\varepsilon$.

In each of the three cases, we first show that (z_0, w_0) is controlled by Γ for all $n \ge 0$. This is done by verifying inductively at free returns the hypothesis of Lemma 7.1. The arguments are essentially the same as those for Theorem 2(2).

From the control of (z_0, w_0), it follows that at free returns, $\|w_i\| > e^{\frac{c}{3}i}$. Next we consider the drop in $\|w_i^*\|$ one step later. This is given by $d_{\mathcal{C}}(z_i)$, which by the definition of Ω_ε is $\ge b^{\frac{k_\varepsilon}{20}}$. Further drops at bound returns are exponentially small. For comparisons between w_i^*- and w_i- vectors, since the fold period ℓ initiated at time i is $\le \frac{k_\varepsilon}{10}$, we have, for $j < \ell$, $\|w_{i+j}\| \ge K^{-\frac{k_\varepsilon}{10}}\|w_{i+\ell}\| = K^{-\frac{k_\varepsilon}{10}}\|w_{i+\ell}^*\|$. $\diamondsuit$

Let z_0 be as above. From Claim 7.1, the fields of most contracted directions of sufficiently high orders are defined at z_0, and their uniform contractive estimates are passed on to $e_\infty := \lim_n e_n$ (see Corollary 2.1). Let $v(z_0) = e_\infty(z_0)$. For other $\xi_0 \in \Omega_\varepsilon$, let $v(\xi_0) = DT^{-k(\xi_0)}(\xi_{k(\xi_0)})v(\xi_{k(\xi_0)})$. Using the fact that $k(\xi_0) < 2k_\varepsilon$ and letting $M(z) = \frac{1}{|\det DT(z)|^{1/2}} DT(z)$, we see that the conditions of Lemma 7.3 are satisfied. Uniform hyperbolicity follows.

It remains to prove that a lower bound for $\|DT^i|E^u\|$ is as claimed. In the argument above we have produced for each $\xi_0 \in \Omega_\varepsilon$ a vector u_0 uniformly bounded away from $E^s(\xi_0)$ such that $\|u_i\| \ge K_\varepsilon^{-1} e^{\frac{c}{3}i}$. Since $\angle(u_n, E^u(\xi_n)) \to 0$ uniformly, we have $\|u_n\| \sim \|DT^n|E^u(\xi_0)\|$. The assertion in Theorem 2(i) on periodic points is proved similarly. $\square$

Proof of Theorem 2(1)(ii): We now prove that the deterioration of hyperbolicity on Ω_ε as $\varepsilon \to 0$ is not only a possibility but a fact. To do this, it suffices to produce a point $z \in \Omega_\varepsilon$ with the property that $\angle(E^u(z), E^s(z)) < K\varepsilon$. We can choose this point to be on the unstable manifold $W^u(\hat{z})$ of any $\hat{z} \in \Omega_\delta$. For $\xi_0 \in W^u(\hat{z})$, let τ_0 be its unit tangent vector to $W^u(\hat{z})$,

Claim 7.2 *For all $\xi_0 \in W^u(\hat{z})$, (ξ_0, τ_0) is controlled by Γ for all $n \ge 0$.*

Proof of Claim 7.2: It suffices to prove the result for $\xi_0 \in W^u_{loc}(\hat{z})$. Suppose that (ξ_0, τ_0) is controlled up to time $k-1$, ξ_k is a free return, and $\xi_k \in \mathcal{C}^{(j-1)} \setminus \mathcal{C}^{(j)}$ for some j. Since $\xi_{k-j} \in \Omega$, it follows that $\xi_k \in R_j$, so that $\xi_k \in \hat{Q}^{(j)} \setminus Q^{(j)}$ for some $Q^{(j)}$. If $j \leq k$, then Lemma 7.1 applies directly. If not, we let $z_0 = \xi_{k-j}$ and apply Lemma 7.1 to the orbit of $(z_0, \tau_0(z_0))$. $\diamondsuit$

Let $\gamma = W^u_{\delta/2}(\hat{z})$. We will show that there exists $z \in (T^n\gamma \cap \Omega_\varepsilon)$ for some $n > 0$ such that $d_\mathcal{C}(z) < 2\varepsilon$. As γ is iterated, it gets long and eventually meets the region $\{d_\mathcal{C}(\cdot) < \varepsilon\}$. Let n_0 be the first time this happens, and let $\omega_0 \subset T^{n_0}\gamma$ correspond to some $I_{\mu j}$ in the region $\{\varepsilon \leq d_\mathcal{C}(\cdot) \leq 2\varepsilon\}$. (See the beginning of Sect. 6.1 for notation.) Note that ω_0 is free. We set binding for ω_0 and iterate until it becomes free again at time n_1. We then subdivide the image into segments corresponding to $I_{\mu j}$ (by which we include pieces outside of $\mathcal{C}^{(0)}$), and let ω_1 be the longest of the divided subsegments. We iterate ω_1 until it becomes free again at time n_2. Then divide and choose ω_2 to be the longest of the subsegments etc. Let $z \in \cap_{i\geq 0} T^{-(n_i - n_0)}\omega_i$. Using Corollary 4.3, we verify that $\omega_i \cap \{d_\mathcal{C}(\cdot) < \varepsilon\} = \emptyset$ for all $i \geq 0$, so that $z \in \Omega_\varepsilon$.

It remains to estimate $\angle(E^u(z), E^s(z))$. First, since $\tau(z)$ splits correctly, we have $\angle(E^u(z), \tau(\phi(z))) < \varepsilon_0 d_\mathcal{C}(z) < 2\varepsilon_0\varepsilon$. Note that $\tau(\phi(z)) = e_\infty(\phi(z))$ and $E^s(z) = e_\infty(z)$. We leave it as an easy exercise to show that $\|DT^n(z)\tau_0(z_0)\| \geq 1$ for all $n > 0$ (use Claim 7.2 and Corollary 4.3), so that at both z and $\phi(z)$, $\angle(e_n, e_\infty) = \mathcal{O}(b^n)$. Let n be such that $\lambda^n \sim \varepsilon$ where λ is as in Lemma 2.2. Then $\angle(e_n(z), e_n(\phi(z))) < K\varepsilon$, and $\mathcal{O}(b^n) << \varepsilon$, proving $\angle(\tau(\phi(z)), E^s(z)) < K'\varepsilon$. This completes the proof. $\square$

8 Statistical Properties of SRB Measures

We follow [Y3] and [Y4], which put forward a scheme for obtaining statistical information for general dynamical systems with some hyperbolic properties. In this approach, one constructs reference sets and studies regular returns to these sets. Sufficient conditions in terms of return times are then given for various statistical properties.

In Sect. 8.1, we indicate how this setup is arranged for the class of attractors in question. For technical details on this construction, we refer the reader to [BY2], where a similar construction is carried out for the Hénon maps. SRB measures and their statistical properties are discussed in Sects. 8.2 and 8.4. A feature of the present setting is that depending on the transitivity properties of T, our attractor may admit multiple SRB measures.

Obviously, the method of [Y3] and [Y4] gives information only on orbits that pass through the reference sets constructed. To complete the picture, we prove in Sect. 8.3 that all SRB measures are captured by our reference sets, and Lebesgue-almost every initial condition in the basin is accounted for.

8.1 Positive-measure horseshoes with infinitely many branches and variable return times

In [Y3], a unified way of looking at nonuniformly hyperbolic systems is proposed. This dynamical picture requires that one constructs a reference set and a return map with Markov properties. The purpose of this subsection is to recall this construction in the context of the maps under consideration, and to give a summary of the facts needed in the discussion to follow.

8.1.1 Construction of reference set

Let $\{x_1, \cdots, x_r\}$ be the set of critical points of f. Our reference set Λ is the disjoint union of $2r$ Cantor sets $\Lambda_1^\pm, \cdots, \Lambda_r^\pm$ where Λ_i^+ and Λ_i^- are located in the component of $\mathcal{C}^{(0)}$ containing $(x_i, 0)$, one on each side of $(x_i, 0)$. We define Λ_i^+ (respectively Λ_i^-) by specifying two transversal families of curves $\Gamma_i^{+,s}$ and $\Gamma_i^{+,u}$ and letting

$$\Lambda_i^+ = \{z \in \gamma^u \cap \gamma^s : \gamma^u \in \Gamma_i^{+,u},\ \gamma^s \in \Gamma_i^{+,s}\}.$$

The family $\Gamma_i^{+,s}$ (no relation to the critical set Γ_i in Sections 3–6) is defined as follows. Let $\mathcal{P}$ be the partition in Sect. 6.1 centered at $(x_i, b) \in \partial R_0$. (To simplify notation, ∂R_0 in this section refers to the top boundary of R_0.) Let $\omega_0 \subset \partial R_0$ be the outermost $I_{\mu j}$ on the right, and let $\omega_\infty = \{z_0 \in \omega_0 : d_{\mathcal{C}}(z_n) > \delta e^{-\alpha n}$ for all $n \geq 0\}$. Letting $m_\gamma(\cdot)$ denote the measure on a curve γ induced by arc length, it is proved in Sect. 6.1 that $m_{\omega_0}(\omega_\infty) > 0$. For every $z_0 \in \omega_\infty$, since $\|DT^i(z_0)\tau_0\| \geq \delta e^{\frac{cn}{3}}$ for all $n \geq 0$ (use (IA5) and the definition of ω_∞), there is a stable curve of every order passing through it. These curves converge to a stable curve $\gamma^s(z_0)$ of infinite order (Sect. 2.1). Moreover, $\gamma^s(z_0)$ has slope $> K^{-1}\delta$ and connects the two boundaries of R_0. We define $\Gamma_i^{+,s} := \{\gamma^s(z_0) : z_0 \in \omega_\infty\}$.

To define $\Gamma_i^{+,u}$, we first let $\tilde{\Gamma}_i^{+,u}$ be the set of all free segments γ of ∂R_n, all $n \geq 0$, such that γ is three times as long as ω_0 and has its midpoint vertically aligned with that of ω_0. Let $\Gamma_i^{+,u}$ be the set of curves that are pointwise limits of sequences in $\tilde{\Gamma}_i^{+,u}$. We remark that since the curves in $\tilde{\Gamma}_i^{+,u}$ are $C^2(b)$, their slopes as functions in x form an equicontinuous family. This implies that the curves in $\Gamma_i^{+,u}$ are at least C^{1+Lip}, and that the tangent vectors of curves in $\tilde{\Gamma}_i^{+,u}$ converge uniformly to the tangent vectors of curves in $\Gamma_i^{+,u}$.

Recalling that Λ_i^+ and Λ_i^- are the Cantor sets that straddle x_i, we may, for convenience, choose $\Gamma_i^{-,s}$ and $\Gamma_i^{+,s}$ in such a way that their elements are paired, i.e. the T-image of each element in $\Gamma_i^{-,s}$ lies on a stable curve containing the T-image of an element of $\Gamma_i^{+,s}$, and *vice versa.*

This completes the construction of $\Lambda = \cup_{i=1}^r \Lambda_i^\pm$. A similar construction is carried out for the Hénon maps in [BY2], Sects. 3.1-3.4.

8.1.2 Structure of return map

Next we define a return map $T^R : \Lambda \to \Lambda$ with the following properties: Topologically, $T^R : \Lambda \to \Lambda$ has the structure of an infinite horseshoe. For simplicity of notation, we write $\Lambda_i = \Lambda_i^+$ or Λ_i^-. A set $X \subset \Lambda_i$ is called an *s-subset* of Λ_i if there exists a subcollection of $\Gamma \subset \Gamma_i^s$ such that $X = \{z \in \gamma^s \cap \gamma^u : \gamma^s \in \Gamma,\ \gamma^u \in \Gamma_i^u\}$; *u-subsets* are defined similarly. If X is an s-subset of Λ_i, we say $X = \Lambda_i$ *mod* 0 if $m_{\partial R_0}(\Lambda_i - X) = 0$.

Lemma 8.1 *There is a map $T^R : \Lambda \to \Lambda$ with the following properties: every Λ_i has a collection of pairwise disjoint s-subsets $\{\Lambda_{i,j}\}_{j=1,2,\cdots}$ with $\Lambda_i = \cup_j \Lambda_{i,j}$ mod 0 such that for each j,*

- *$T^R|\Lambda_{i,j} = T^{n_{i,j}}|\Lambda_{i,j}$ for some $n_{i,j} \in \mathbb{Z}^+$;*
- *$T^R(\Lambda_{i,j})$ is a u-subset of Λ_k for some $k = k(i,j)$.*

We stress that the partition of Λ_i into $\{\Lambda_{i,j}\}$ is an infinite one, and that the return times $n_{i,j}$ are not bounded. The **return time function** $R : \Lambda \to \mathbb{Z}^+$ is defined to be $R|\Lambda_{i,j} = n_{i,j}$. As we will see, the tail of this function, that is, the distribution of its large values, plays a crucial role in determining the statistical properties of the system. Note that T^R is not necessarily the first return map; we have settled for possibly larger return times in favor of a Markov structure. Lemma 8.1 corresponds to Proposition A(1) in [BY2]; its proof is given in Sects. 3.4 and 3.5 of [BY2].

8.1.3 Two important analytic estimates

Technical estimates corresponding to (P1)-(P5) in [Y3] or Proposition B of [BY2] are needed. Referring the reader to Section 5 of [BY2] for their precise statements and proofs, we state below two of the most relevant facts.

Lemma 8.2 (Distortion estimate for controlled segments) *There exists $K > 0$ such that the following holds: Let γ_0 be a curve and τ_0 its unit tangent vectors. We assume that*

(i) for all $z_0 \in \gamma_0$, (z_0, τ_0) is controlled up to time $n-1$;
(ii) γ_i is bound or free simultaneously for each i, and γ_i is contained in three contiguous $I_{\mu j}$ at all free returns;
(iii) γ_n is a free return.

Then for all $z_0, z_0' \in \gamma_0$,

$$\frac{1}{K} \le \frac{\|\tau_n(z_0)\|}{\|\tau_n(z_0')\|} \le K.$$

The proof is similar to that of Proposition 6.2 (it is, in fact, a little simpler) and will be omitted. In the construction of $T^R : \Lambda \to \Lambda$, it is important to arrange that γ_0, the shortest subsegment of ∂R_0 that spans $\Lambda_{i,j}$ in the u-direction, satisfies (ii) above up to time $n_{i,j}$.

Let $\cup\Gamma_i^s := \cup\{z \in \gamma^s : \gamma^s \in \Gamma_i^s\}$. If γ and γ' are curves transversal to the elements of Γ_i^s and intersecting them, we define $\psi : \gamma \cap (\cup\Gamma_i^s) \to \gamma'$ by sliding along the curves in Γ_i^s, and say Γ_i^s is *absolutely continuous* if for every pair of $C^2(b)$-curves γ and γ' as above, ψ carries sets of m_γ-measure zero to sets of $m_{\gamma'}$-measure zero. Recall that if γ is the subsegment of ∂R_0 in Γ_i^u, then $m_\gamma(\gamma \cap (\cup\Gamma_i^s)) > 0$; in particular, the definition above is not vacuous.

Lemma 8.3 (Absolute continuity of Γ_i^s) *Γ_i^s is absolutely continuous with*

$$\frac{1}{K} \;<\; \frac{d}{dm_{\gamma'}}\psi_*(m_\gamma | \cup\Gamma_i^s) \;<\; K \qquad \text{on } \gamma' \cap (\cup\Gamma_i^s).$$

Except for one minor technical difference, the proof of Lemma 8.3 is identical to that of Sublemma 10 in Section 5 of [BY2]: in the latter, the transversals are taken to be curves in $\tilde{\Gamma}_i^u$, whereas we need them to be arbitrary $C^2(b)$ curves here. Clearly, it suffices to show that Sublemma 10 of [BY2] is valid with $\gamma \in \tilde{\Gamma}_i^u$ and γ' arbitrary $C^2(b)$, and for that we need distortion estimates for the τ_i-vectors on certain subsegments of γ' (ω' in the proof of Sublemma 10). We have them because these subsegments are connected to subsegments of γ by (temporary) stable curves, and the corresponding τ_i-vectors are comparable.

8.1.4 Tail of return times

Finally we state an estimate on which the statistical properties of T depend crucially. Its proof is identical to that of Proposition A(4) in [BY2].

Lemma 8.4 *There exists K and $\theta_0 < 1$ such that for every Λ_i,*

$$m_{\partial R_0}\{z \in \partial R_0 \cap \Lambda_i : R(z) > n\} < K\theta_0^n.$$

8.2 SRB measures

8.2.1 Construction of SRB measures

We describe below a recipe for constructing SRB measures using the reference sets $\{\Lambda_i^\pm\}$. For the definition of SRB measures, see Sect. 1.3. For more details on the technical justification of the steps below, see [Y3] or [BY2], Sect. 6.2. The construction consists of three steps.

Step 1. Construction of a T^R-invariant measure ν on $\cup\Lambda_k^\pm$ with absolutely continuous conditional measures on the leaves of $\Gamma^u := \cup_k\Gamma_k^{\pm,u}$. We fix some $\Lambda_i = \Lambda_i^+$ or Λ_i^-, and let $m_0 = m_{\partial R_0} \mid (\Lambda_i \cap \partial R_0)$. Let ν be an accumulation point of the sequence of measures

$$\frac{1}{n}\sum_{j=0}^{n-1}(T^R)_*^j m_0, \quad n = 1, 2, \cdots.$$

Then ν is a T^R-invariant measure. By Lemma 8.2, the conditional measures of $(T^R)^j_* m_0$ on the curves of $\tilde\Gamma^u := \cup_k \tilde\Gamma_k^{\pm,u}$ have uniformly bounded densities. From Lemma 8.2 and the Markov property of T^R (see Sect. 8.1.2), it follows that for $\gamma \in \tilde\Gamma^u$, the conditional densities of $(T^R)^j_* m_0$ on γ when restricted to $\gamma \cap (\cup\Gamma^s)$ are uniformly bounded away from 0. These properties are passed on to the conditional measures of ν on the leaves of Γ^u. (The curves in Γ^u are pairwise disjoint except possibly for a countable number of pairs; this is nothing more than a technical nuisance.)

Step 2. Construction of a T-invariant probability measure μ given ν. It follows from the bounded densities of ν, Lemma 8.3 and Lemma 8.4 that $\int_\Lambda R d\nu_0 < \infty$. Let

$$\mu \ = \ \frac{1}{\int R d\nu_0} \sum_{j=0}^{\infty} T_*^j(\nu_0 \mid \{R > j\}).$$

It is straightforward to check that μ is a T-invariant probability measure.

Step 3. Proof of SRB property. Let μ be as in Step 2. First we check that T has a positive Lyapunov exponent μ-a.e.. At $z_0 \in (\cup\Gamma^s) \cap (\cup\tilde\Gamma^u)$, let $\tilde\tau(z_0)$ be a unit tangent vector to $\tilde\Gamma^u(z_0)$, the $\tilde\Gamma^u$-curve through z_0. Just as on ω_∞, we have $\|DT^n(z_0)\tilde\tau\| \geq \delta e^{\frac{cn}{3}}$ for all $n \geq 0$. This uniform growth is passed on to the tangent vectors τ to Γ^u-curves at every $z \in \Lambda = \cup_k \Lambda_k^{\pm}$. The existence of a positive Lyapunov exponent μ-a.e. follows from the fact that the orbit of μ-almost every point passes through Λ. General nonuniform hyperbolic theory (see e.g. [P] or [R4]) then tells us that stable and unstable manifolds exist μ-a.e.

To prove that μ is an SRB measure, we need to show that its conditional measures on unstable manifolds are absolutely continuous. Since μ is the sum of forward images of ν, it suffices to prove this for ν. We know from Step 1 that ν has absolutely continuous conditional measures on the leaves of Γ^u. Thus it remains to prove

Claim 8.1 *For ν-a.e. z_0, $\Gamma^u(z_0)$ is a local unstable manifold, i.e.*

$$\limsup_{n\to\infty} \frac{1}{n} \log \sup_{\xi_0 \in \Gamma^u(z_0)} |\xi_{-n} - z_{-n}| \ < \ 0.$$

Proof of Claim 8.1: From the construction of ν, it follows that that for ν-a.e. $z_0 \in \Lambda$, there is a sequence of $\tilde\Gamma^u$-curves $\{\tilde\gamma_i\}$ such that $\tilde\gamma_i \to \Gamma^u(z_0)$. Let n_i be such that $T^{-n_i}\tilde\gamma_i \subset \partial R_0$. Since $\tilde\gamma_i$ is free, we have that for all tangent vectors $\tilde\tau$ of $\tilde\gamma$, $\|DT^{-n}\tilde\tau\| < e^{-c''n}$ for some $c'' > 0$ and $0 < n \leq n_i$ (Proposition 5.1 and Lemma 4.8). These uniform estimates for backward iterates of T are passed on to all tangent vectors of $\Gamma^u(z_0)$, proving that it is a local unstable manifold of z_0. $\diamondsuit$

8.2.2 Ergodic decomposition of SRB measures

We begin by considering the ergodic decompositions of the T^R-invariant measures constructed in Step 1 in Sect. 8.2.1.

Definition 8.1 *Let $g : X \to X$ be a continuous map of a compact metric space, and let ν be a g-invariant Borel probability measure on X. We say $z \in X$ is* **generic** *or* **future-generic** *with respect to ν if for every continuous function $\varphi : X \to \mathbb{R}$,*

$$\frac{1}{n}\sum_{i=0}^{n-1}\varphi(z_i) \to \int \varphi d\nu.$$

Let $\mathcal{M}(T^R)$ be the set of all normalized invariant measures constructed in Step 1 of Sect. 8.2.1. Let $\nu \in \mathcal{M}(T^R)$, and suppose that $\nu(\Lambda_i) > 0$ for some $\Lambda_i = \Lambda_i^+$ or Λ_i^-. From the positivity of the conditional densities of ν on $\Lambda_i \cap \gamma$, Lemma 8.3, and a standard argument due to Hopf, we know that there is an ergodic component ν^e of ν such that

(i) ν-a.e. $z \in \Lambda_i$ is generic with respect to ν^e;

(ii) for every $C^2(b)$-curve γ, m_γ-a.e. $z \in \gamma \cap (\cup \Gamma_i^s)$ is generic with respect to ν^e.
We abbreviate this by saying ν^e "occupies" Λ_i.

Let $\mathcal{M}_e(T^R)$ denote the set of normalized ergodic components of measures in $\mathcal{M}(T^R)$. Then each Λ_i^+ (resp. Λ_i^-) is occupied by an elememt of $\mathcal{M}_e(T^R)$. Because the stable curves of Λ_i^+ and Λ_i^- are joined, Λ_i^- and Λ_i^+ are in fact occupied by the same element of $\mathcal{M}_e(T^R)$. Thus the cardinality of $\mathcal{M}_e(T^R)$ is $\leq r$.

To further study the structure of $\mathcal{M}_e(T^R)$ we borrow some ideas from finite state Markov chains. Let $\Lambda_i^\pm := \Lambda_i^+ \cup \Lambda_i^-$. We think of each the sets $\Lambda_i^\pm, i = 1, \cdots, r$, as a state, and write "$i \to j$" if $T^R(\Lambda_i^\pm) \cap \Lambda_j^\pm \neq \emptyset$. We say i is *transient* if there exists j such that there is a chain $i \to \cdots \to j$ but no chain with $j \to \cdots \to i$. Non-transient states are called *recurrent*. The following are consequences of simple facts about directed graphs.

(a) The set of recurrent states is partitioned into equivalence classes where $i \sim j$ if there is a chain $i \to \cdots \to j$. On the union of the $\Lambda_i^\pm$ corresponding to the states in each equivalence class is supported exactly one element of $\mathcal{M}_e(T^R)$, which occupies each of these $\Lambda_i^\pm$.

(b) If i is transient, then clearly $\nu(\Lambda_i^\pm) = 0$ for every $\nu \in \mathcal{M}_e(T^R)$. The following claim is a consequence of the structure of T^R (Lemma 8.1) and the fact that for every transient state j, there exists a recurrent k such that $j \to \cdots \to k$.

Claim 8.2 $\Lambda_i^\pm$ *is the mod 0 union of a collection of pairwise disjoint s-subsets $\{\hat{\Lambda}_{i,\ell}\}_{\ell=1,2,\cdots}$ with the property that for each ℓ, there exists $n_\ell > 0$ such that $(T^R)^{n_\ell}\hat{\Lambda}_{i,\ell}$ is a u-subset of some recurrent state.*

The discussion in Sects. 8.2.1 and 8.2.2 are summarized as follows:

Proposition 8.1 *Let r be the number of critical points of f. Then there exist ergodic SRB measures*

$$\mu_1,\ \mu_2,\ \cdots,\ \mu_{r'}, \qquad 1 \le r' \le r,$$

such that for every $C^2(b)$-curve γ, m_γ-a.e. $z \in \gamma \cap (\cup \Gamma^s)$ is generic with respect to some μ_i.

Proof: Let $\mathcal{M}_e(T^R) = \{\nu_1, \nu_2, \cdots, \nu_{r''}\}$. Then the μ_i in this proposition are saturations of the $\nu_j \in \mathcal{M}_e(T^R)$ in the sense of Step 2 in Sect. 8.2.1. Clearly $r' \le r'' \le r$; it may happen that $r' \le r''$ because the saturations of distinct T^R-invariant measures may merge. The genericity assertion is proved as follows. If k is a recurrent state, then it is occupied by some ν_j, and hence m_γ-a.e. $z \in \gamma \cap (\cup \Gamma^s_k)$ is generic with respect to some μ_i. Via Claim 8.2, the same conclusion holds if k is a transient state. □

8.3 A bound on the number of ergodic SRB measures and accounting for almost every initial condition in the basin

Let m denote the Lebesgue measure on R_0. In this subsection we prove

Proposition 8.2 *Let $\{\mu_i\}$ be the ergodic SRB measures in Proposition 8.1. Then for m-a.e. $z_0 \in R_0$, there exists μ_i with respect to which z_0 is generic. We prove this by showing that for some $n > 0$, z_n lies in the local stable curve of a μ_i-typical point.*

The definition of genericity is given in Sect. 1.3. Proposition 8.2 serves two purposes: It accounts for the behavior of Lebesgue-a.e. initial condition in the basin of attraction and proves, at the same time, that the $\{\mu_i\}$ are *all* of the ergodic SRB measures of T (see (ii) below), thereby putting an upper bound on this number. Propositions 8.1 and 8.2 together prove Theorem 3.

The following background information on general nonuniform hyperbolic theory may help put things in perspective:

(i) In general, not all attractors admit SRB measures. This is the case even when there is a great deal of hyperbolicity (see e.g. [HY]). Also, without assumptions of transitivity, the number of ergodic SRB measures on an attractor can, in theory, be countably infinite (see [Led]). For the maps considered in this paper, examples show that multiple SRB measures do occur, and the given bound is achieved.

(ii) Returning to general nonuniform theory, let $B(\mu)$ denote the **measure-theoretic basin** of μ, i.e. the set of points generic with respect to the measure μ. If μ is an ergodic SRB measure with no zero Lyapunov exponents, then $B(\mu)$ has positive Lebesgue measure (see [PS]). This is the reason why SRB measures are important in physics. It is also how we deduce from Proposition 8.2 that we have exhausted our list of ergodic SRB measures.

(iii) In general, without assumptions of transivity, the attractor can be considerably smaller than the union of the supports of its ergodic SRB measures. This happens for the maps we are considering.

(iv) In general, the union of $B(\mu)$ as μ ranges over all ergodic SRB measures need not be a full Lebesgue measure subset of the topological basin of the attractor (meaning the set of all points z with $d(T^i z, \Omega) \to 0$). Measure-theoretic basins can be strictly smaller even when the SRB measure is unique or when $\cup \text{supp}(\mu) = \Omega$. Proposition 8.2 therefore goes beyond general theory to describe a nice property of these attractors.

(v) Finally, when there is more than one ergodic SRB measure, their measure-theoretic basins can be very delicately interwined. For the maps being considered here, we leave it as an exercise to construct examples in which there are arbitrarily small open sets meeting every $B(\mu_i)$ in a set of positive Lebesgue measure.

Proof of Proposition 8.2: Let B be the set of points *not* generic with respect to any of the μ_i. We remark that B is a Borel measurable set, for genericity with respect to a given measure is determined by a countable number of test functions. Let $Z^{(k)}$ be as in Sect. 7.3. Let $Y_0 = \{z_0 \in R_0 : z_k \notin Z^{(k)} \text{ for any } k \geq 0\}$, and for $i \geq 1$, let

$$Y_i = \{z_0 \in R_0 : z_i \in Z^{(i)} \text{ and } z_k \notin Z^{(k)} \text{ for all } k > i\}.$$

Suppose $m(B \cap Y_i) > 0$ for some $i > 0$. Then $m(B \cap T^i Y_i) > 0$, and there is a vertical line γ with $m_\gamma(B \cap T^i Y_i) > 0$. Let $\varepsilon > 0$ be a small number. By the Lebesgue density theorem, there exists a short segment $\gamma_0 \subset \gamma$ with the property that $m_\gamma(B \cap T^i Y_i \cap \gamma_0) > (1-\varepsilon) m_\gamma(\gamma_0)$. We will show in the next paragraphs that points generic with respect to some μ_i make up a definite fraction of γ_0, contradicting our choice of γ_0 if ε is sufficiently small. (The argument we present also works if $m(B \cap Y_0 \cap \mathcal{C}^{(0)}) > 0$. For the case $m(B \cap Y_0 \cap (R_0 \setminus \mathcal{C}^{(0)})) > 0$, use horizontal instead of vertical lines.)

Let τ_0 denote the tangent vectors to γ_0, and let $\gamma_j = T^j \gamma_0$. We regard all of γ_0 (which can be taken to be arbitrarily short) as bound to its nearest critical point, and let n_1 be the first time when part of γ_j makes a free return to $\mathcal{C}^{(0)}$. As before, let $\Lambda_k = \Lambda_k^+$ or Λ_k^-. Let $D(\Lambda_k)$ denote the smallest rectangular region bounded by Γ^u and Γ^s-curves that contains Λ_k. If γ_{n_1} crosses some $D(\Lambda_k)$ with two segments of at least comparable lengths extending beyond the two sides of $D(\Lambda_k)$, we consider the segment $\gamma_{n_1} \cap D(\Lambda_k)$ as having reached its final destination and take it out of circulation. We then divide what remains of γ_{n_1} into $I_{\mu j}$ and delete those subsegments that do not contain a point of $T^{n_1}(B \cap T^i Y_i)$.

Observe that for $z_0 \in \gamma_0 \cap T^i Y_i$, (z_0, τ_0) is controlled through time $n_1 - 1$, and by Lemma 7.1, τ_{n_1} splits correctly (see the proof of Theorem 2(2)). This is true not only for $z_0 \in \gamma_0 \cap T^i Y_i$ but also for $z_0' \in \gamma_0$ such that z_{n_1}' is in the same $I_{\mu j}$ as z_{n_1}. We iterate independently each one of the $I_{\mu j}$-segments that are kept. At the next free return we repeat the same procedure, namely we take out subsegments that cross some $D(\Lambda_k)$, divide the rest into $I_{\mu j}$, delete those that do not contain a point in the image of $B \cap T^i Y_i$, and observe that for the remaining segments control is extended to the next free return.

Let $\gamma_0^d = \{z_0 \in \gamma_0 : z_j$ is deleted at a free return for some $j > 0\}$, and let $\hat{\gamma}_0 = \{z_0 \in \gamma_0 : z_j$ reaches $D(\Lambda_k)$ for some j and k in the required manner$\}$. We note that $m_{\gamma_0}(\gamma_0 \setminus (\gamma_0^d \cup \hat{\gamma}_0)) = 0$. This follows from a sublemma which is the first step in the proof of Lemma 8.4 (see [BY2], Sublemma 4 and its corollary).

Since $(\gamma_0 \cap B \cap T^i Y_i) \cap \gamma_0^d = \emptyset$, we have $(\gamma_0 \cap B \cap T^i Y_i) \subset \hat{\gamma}_0$ mod 0 and that $\hat{\gamma}_0$ is the disjoint union of a countable number of subsegments $\{\omega\}$ with the following properties:

– each ω is mapped under some $T^{n(\omega)}$ onto a $C^2(b)$-curve that connects two Γ^s-sides of some $D(\Lambda_k)$;

– (z_0, τ_0) is controlled up to time $n(\omega)$ for every $z_0 \in \omega$.

From Lemmas 8.2, 8.3 and Proposition 8.1, it follows that there exists $c_1 > 0$ independent of the choice of γ_0 such that for each ω,

$$m_\gamma\{z_0 \in \omega : z_{n(\omega)} \in \cup\Gamma^s \text{ and is generic w.r.t. some } \mu_k\} > c_1 m_\gamma(\omega).$$

This implies that $m_\gamma\{z_0 \in \gamma_0 : z_0$ is generic w.r.t. some $\mu_k\} > c_1 m_\gamma(\hat{\gamma}_0) > c_1(1-\varepsilon)$ $m_\gamma(\gamma_0)$, contradicting our choice of γ_0 if $c_1(1-\varepsilon) > \varepsilon$. □

8.4 Correlation decay and Central Limit Theorem

We indicate how Theorem 4 is proved. The setup $T^R : \Lambda \to \Lambda$ is designed so that the statistical properties in question are easily read off from the tail properties of the return time function R. To use the results in [Y3] or [Y4] directly, however, we need to consider returns to a single recurrent state. Let $\tilde{\mu}$ be one of the μ_j in Proposition 8.1, and let $\tilde{\Lambda}$ be one of the Λ_i such that $\tilde{\mu}(\Lambda_i) > 0$. For $z \in \tilde{\Lambda}$, we define a return time $\tilde{R}(z)$ of z to $\tilde{\Lambda}$ by $\tilde{R}(z) = t_0 + t_1 + \cdots + t_n$, where $t_0 = R(z)$, $t_1 = R(T^R(z))$, $\cdots$, $t_n = R((T^R)^n z)$ and $(T^R)^{n+1} z$ is the first return to $\tilde{\Lambda}$ under T^R. The results in [Y3] or [Y4] allow us to read off information on the statistical properties of $(T, \tilde{\mu})$ via the asymptotics of $m_{\partial R_0}\{z \in \partial R_0 \cap \tilde{\Lambda} : \tilde{R}(z) > n\}$.

Lemma 8.5 *There exists $K > 0$ and $\tilde{\theta}_0 < 1$ such that for every $n > 0$,*

$$m_{\partial R_0}\{z \in \partial R_0 \cap \tilde{\Lambda} : \tilde{R}(z) > n\} < K \tilde{\theta}_0^n.$$

This lemma, which we leave as an exercise, is an easy consequence of Lemma 8.4. The results in [Y3] and [Y4] state that if the quantity estimated in Lemma 8.5 is of order $\mathcal{O}(\frac{1}{n^{2+\varepsilon}})$ for some $\varepsilon > 0$, then the Central Limit Theorem holds in the context of Theorem 4. This condition is evidently satisfied here. They also tell us that if this quantity is exponentially small, then every mixing component of $\tilde{\mu}$ has exponential decay of correlations as asserted.

9 Global Geometry

9.1 Motivation

Nonuniformly hyperbolic attractors have very complicated local structures. The purpose of this section is to develop an understanding of the *coarse geometry* of the attractor Ω for the maps in question, that is to say, to describe in a finite way the approximate shape and complexity of Ω.

To illustrate the idea of coarse geometry, consider the standard solenoid constructed from $z \mapsto z^2$. A good approximation of the attractor is given by the kth forward image of $S^1 \times D_2$, which is a tubular neighborhood of a simple closed curve winding around the solid torus 2^k times. For another example, consider piecewise monotonic maps in 1-dimension. Iterates of these maps continue to be piecewise monotonic and can be understood in terms of their monotone pieces.

Returning to the maps under consideration, the standard solenoid example suggests that R_k may be a good approximation of Ω. In analogy with 1-dimension, one may also guess that R_k is a tubular neighborhood of a simple closed curve whose x-coordinates vary in a piecewise monotonic fashion. The latter is false, as is evident from the following sequence of pictures: Depicted in (a) is a section of R_k lying between two $C^2(b)$-curves; (b) is the image of (a). As (b) is iterated, the horizontal distance between the tips of the two parabolas increases as shown in (c), until at some point they fall on opposite sides of a component of the critical set, resulting in (d). Since this happens to every "turn" that is created, the geometry of R_k for large k is quite complicated.

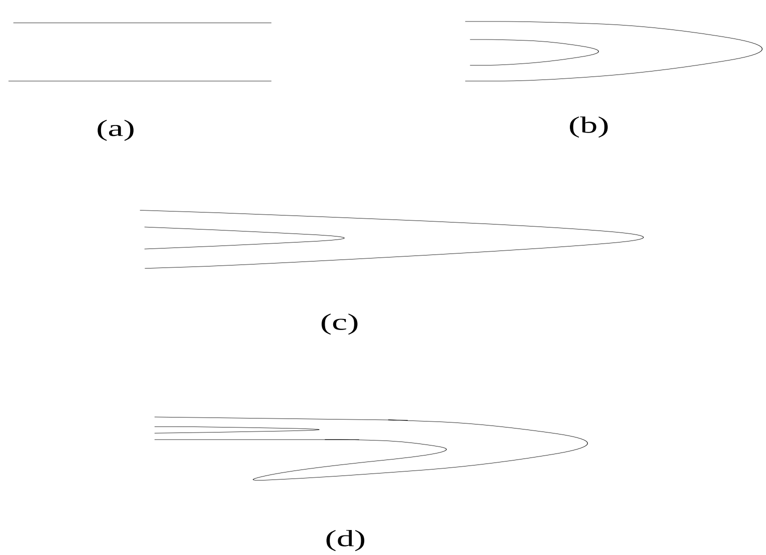


Figure 3 The geometry of R_k

The purpose of this section is to introduce the idea of **monotone branches** as basic building blocks for understanding the global structure of Ω. To each map T we will associate a **combinatorial tree** whose edges correspond to monotone branches, and we will show that Ω has arbitrarily fine neighborhoods made up of unions of finitely many monotone branches. Moreover, the way these branches fit together will tell us exactly how, in finite approximation, T differs from a 1-dimensional map.

9.2 Monotone branches

For $z_0 \in R_0$, let $O_+(z_0) = \{z_1, z_2, z_3, \cdots\}$ denote the positive orbit of z_0, and write $O_+(\Gamma) = \cup_{z_0 \in \Gamma} O_+(z_0)$.

Definition 9.1 *Let γ be a connected subsegment of ∂R_k. We say γ is a* **(maximal) monotone segment** *if*

(i) the two end points of γ are in $O_+(\Gamma)$;

(ii) γ does not intersect $O_+(\Gamma)$ in its interior.

When we say ξ_i is an end point of a monotone segment, it will be understood that ξ_0 is a critical point. We record below some simple facts about monotone segments.

Lemma 9.1 *Let $\gamma \subset \partial R_k$ be a monotone segment. Then:*

(a) All points near the two ends of γ are in their fold periods; the part of γ not in a fold period (respectively bound period), if nonempty, is connected.

(b) If part of γ is free, then its geometry is as follows: γ consists of a relatively long $C^2(b)$-curve connecting two sets of relatively small diameters at the two ends; more precisely, there exists p such that the $C^2(b)$-curve has length $> e^{-\beta p}$ while the diameters of the two small sets are $< b^{\frac{p}{2}}$; also, the curvature of ∂R_k at the end point ξ_i of γ is $> b^{-i}$.

(c) If γ meets Γ in r points, $r \geq 0$, then $T(\gamma)$ is the union of $r+1$ monotone segments joined together at the T-images of these points.

Proof: (c) follows from the definition of a monotone segment. (a) follows from the way monotone segments are created and from the monotonicity of bound and fold periods (see the proof of Lemma 4.10). The first assertion in (b) follows from estimates on the relative sizes of the parts of γ that are in bound versus fold periods; the second follows from the curvature formula in the proof of Lemma 2.4. □

We now begin to study the geometry of certain 2-dimensional objects.

Definition 9.2 *A simply connected region $S \subset R_k$ is called a* **monotone branch** *if it is bounded by two monotone segments γ, $\gamma' \subset \partial R_k$ and two ends E_ξ and E_ζ with the following properties:*

(i) If the end points of γ are ξ_i and ζ_j, then the end points of γ' are ξ'_i and ζ'_j where ξ_0 and ξ'_0 lie on the upper and lower boundaries of the same component $Q^{(k-i)}$ of

$\mathcal{C}^{(k-i)}$, and ζ_0 and ζ_0' are related in the same way.

(ii) $E_\xi = T^i\{z \in Q^{(k-i)}(\xi_0) : |z-\xi_0| < b^{\frac{k-i}{4}}\}$; its **time of creation** *is said to be $k-i$; E_ζ and its time of creation are defined analogously.*

(iii) We define the **age** *of E_ξ to be i and require that $i < \theta^{-1}(k-i+1)$; there is an analogous limit on the age of E_ζ.*

The definitions of E_ξ and E_ζ are quite arbitrary, subject only to the following considerations: We want E_ξ to be large enough to contain all the critical orbits that originate from $Q^{(k-i)}(\xi_0)$. On the other hand, we want it to remain relatively small during the life span of the monotone branch, so that the phenomenon depicted in Figure 3 does not occur. We will assume that for $i < \theta^{-1}(k-i+1)$, $\|DT\|^i b^{\frac{k-i}{4}} < b^{\frac{k-i}{8}} << e^{-\alpha i}$, which is $< d_{\mathcal{C}}(z_i)$ for $z_0 \in \Gamma$ by (IA2) in Section 3; that is to say, if S is a monotone branch of R_k, then its ends are at least a certain distance from $\mathcal{C}^{(k)}$. It is not always easy to visually identify monotone branches, particularly when their boundary segments are in fold periods. When part of γ is free, it follows from Lemma 9.1(b) that S consists of a (relatively long) horizontal strip with two small blobs at the two ends.

Tree structure of a class of monotone branches

Monotone branches can be constructed as follows. First we declare that R_0 is a monotone branch (even though it has no ends). Then if $x_i < x_{i+1}$ are adjacent critical points of the 1-dimensional map f, the T-image of $\{z = (x, y) : x_i - b^{\frac{1}{4}} < x < x_{i+1} + b^{\frac{1}{4}}\}$ is a monotone branch of R_1. In general, let S be a monotone branch of R_k. If one of the ends of S is at its maximum allowed age, then S is "discontinued", meaning we do not iterate it further. If not, $T(S)$ is the union of a finite number of monotone branches of R_{k+1}. More precisely, if $S \cap \mathcal{C}^{(k)} = \emptyset$, then $T(S)$ is a monotone branch. If $S \cap Q^{(k)} \neq \emptyset$, then $S \supset Q^{(k)}$ (in fact, S extends beyond $Q^{(k)}$ by $> e^{-\alpha k}$ in both directions). If S contains r components of $\mathcal{C}^{(k)}$, then $T(S)$ is the union of $r+1$ monotone branches split roughly along the T-images of the middle of each of the $Q^{(k)}$ contained in S (cf. Lemma 9.1(c)).

Let $\mathcal{T} = \cup_k \mathcal{T}_k$ denote the set of all monotone branches inductively constructed this way, with $\mathcal{T}_k$ consisting of branches of R_k. More precisely, $\mathcal{T}_0 = \{R_0\}$, and $\mathcal{T}_{k+1}$ is obtained from $\mathcal{T}_k$ via the procedure described above. **We will be working exclusively with monotone branches in $\mathcal{T}$**, which is a proper subset of the set of all monotone branches in Definition 9.2. The set $\mathcal{T}$ has a natural tree structure: we call the branches obtained by mapping forward and subdividing a given branch its **descendants**. Note that every branch in $\mathcal{T}_k$ has a unique **ancestor** in $\mathcal{T}_i$ for every $i < k$, but not all branches in $\mathcal{T}$ have offsprings: the ones with no offsprings are exactly those one of whose ends has reached its maximum allowed age.

We have elected to discontinue a branch before its geometry "deteriorates". An immediate question that arises is what happens to the part of the attractor contained in a discontinued branch. We will show in the next subsection that branches farther down the tree $\mathcal{T}$ can be used to take its place. We will, in fact, prove the following stronger version of Theorem 5.

Theorem 5′ *One can construct special neighborhoods $\tilde{R}_n$ as in Theorem 5 using only monotone branches from $\mathcal{T}_k$, $n \le k < (1+K\theta)n$.*

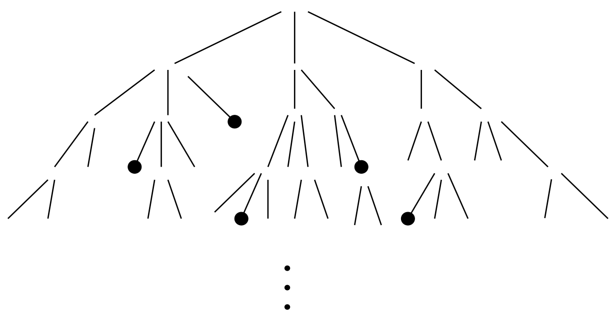


Figure 4 Tree of monotone branches: branches ending in • are discontinued

9.3 Replacement of branches

Let $S \in \mathcal{T}_k$ be a branch whose ends are denoted by E_ξ and E_ζ. In the discussion to follow, we assume that E_ξ is fairly advanced in age, meaning $(k-i) \sim \theta i$ where i is the age of E_ξ and $k-i$ is its time of creation. As we search for replacements for S, the picture we hope to have is the following. There is a finite collection of branches $\{B\} \subset \cup_{k<j\le(1+K\theta)k}\mathcal{T}_j$ such that

(i) the ends of B are contained in those of S; and

(ii) if $S \in \mathcal{S}$ where $\mathcal{S} \subset \mathcal{T}$ is a cover of Ω, then replacing S by $\{B\}$ does not leave any part of Ω exposed.

Let $Q^{(k-i)}$ be the component of $\mathcal{C}^{(k-i)}$ containing $T^{-i}E_\xi$. We hope to show that $T^{-i}S \subset Q^{(k-i)}$, so that the picture described above pulled back to $Q^{(k-i)}$ is as shown in Figure 5.

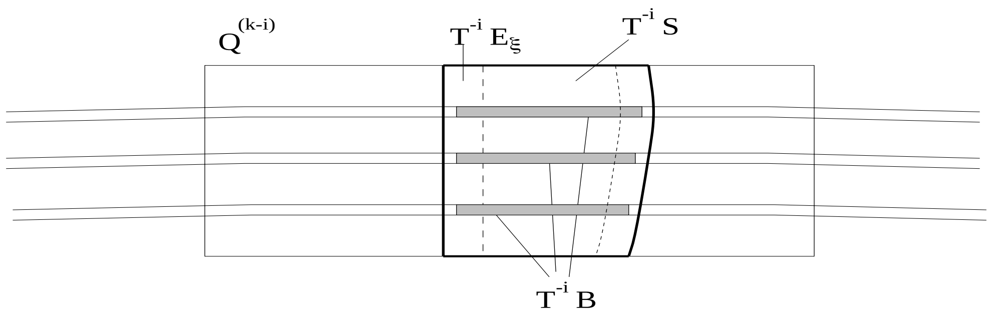


Figure 5 Replacing S by $\{B\}$

We begin to systematically justify this picture. For $j = 0, 1, \cdots, i-1$, let $S_j \in \mathcal{T}_{k-i+j}$ be the ancestor of S, so that S_0 is the monotone branch of R_{k-i} containing $Q^{(k-i)}$. Let E_0 denote the end of S_0 contained in $Q^{(k-i)}$, and let $E_j = T^j E_0$. Let the other end of S_j be called E'_j. Let $t > k-i$, and let $P \in \mathcal{T}_t$ be such that $P \cap Q^{(k-i)}$ is a horizontal strip bounded by two $C^2(b)$-curves stretching all the way across $Q^{(k-i)}$. We think of P as a *pre-branch* with respect to S_0 in the sense that $P \subset S_0$ and it is not yet born when S_0 is created. If P is not discontinued, then we let P_1 be the (unique) child of P with one end in E_1, and assuming P_1 is not discontinued, we let P_2 be the child of P_1 with one end in E_2. Similarly, we define $P_3, P_4, \cdots$ up to P_i if it makes sense.

Lemma 9.2 *There exists K_1 depending on ρ such that*
(i) for all j with $K_1(k-i) < j \leq i$, $T^{-j} S_j \subset Q^{(k-i)}$;
(ii) if P_j is defined for all $j \leq K_1(k-i)$, then it is defined for all $j \leq i$; moreover, for each $j \geq K_1(k-i)$, $P_j \subset S_j$, and the two ends of P_j are contained in the two ends of S_j.

We isolate the following sublemma, the ideas in which are also used elsewhere. See Sect. 6.1 for notation.

Sublemma 9.1 *Let one of the horizontal boundaries of $Q^{(s)}$, any s, be identified with $[-\rho^s, \rho^s]$, with the critical point corresponding to 0. Then for every $I_{\mu_0 j_0} \subset [-\rho^s, \rho^s]$, there exists $n < K|\mu_0|$ such that $T^n I_{\mu_0 j_0}$ traverses completely a component of $\mathcal{C}^{(0)}$.*

Proof: Let $\omega_0 = I_{\mu_0 j_0}$, and let r_0 be the first time when part of ω_0 makes a free return with $T^{r_0}\omega_0$ containing an $I_{\mu j}$ of full length. By Corollary 4.3, either $T^{r_0}\omega_0$ contains one of the outermost $I_{\mu j}$ (which we will call $\tilde{I}$) or it contains some $I_{\mu_1 j_1}$ with $|\mu_1| < K\beta|\mu_0|$. In the latter case, we let $\omega_1 = I_{\mu_1 j_1}$ and continue to iterate until r_1 iterates later when part of $T^{r_1}\omega_1$ is free and contains either $\tilde{I}$ or some $I_{\mu_2 j_2}$ with $|\mu_2| < K\beta|\mu_1|$. After a finite number of iterates, we have $T^{r_q} w_q \supset \tilde{I}$.

From Corollary 4.3, we see that at the end of its bound period, $T^p \tilde{I}$ has length $>> \delta$. Inductively define $\tilde{I}_{p+j} = T(\tilde{I}_{p+j-1}) \setminus \mathcal{C}^{(0)}$ for $j = 1, 2, \cdots$. Then $\tilde{I}_{p+j}$ is a connected $C^2(b)$-curve which grows essentially exponentially – until it crosses completely a component of $\mathcal{C}^{(0)}$. Since $r_i \sim |\mu_i|$ up to the point when $T^{r_q} w_q \supset \tilde{I}$, and the growth is exponential thereafter, we conclude that the end game is reached in a total of $< K|\mu_0|$ iterates. $\square$

Proof of Lemma 9.2:

Claim 9.1 *There exists K_1 (depending on ρ) such that $T^{-K_1(k-i)} S_{K_1(k-i)} \subset Q^{(k-i)}$.*

Proof of Claim 9.1: We identify the upper horizontal boundary of $Q^{(k-i)}$ with the interval $[-\rho^{k-i}, \rho^{k-i}]$, with the critical point corresponding to 0, and let n_1 be the smallest n such that $T^n[0, \frac{1}{2}\rho^{k-i}]$ intersects the horizontal boundary of some $Q^{(k-i+n)}$. From Sublemma 9.1, $n_1 < K_1(k-i)$ for some $K_1 = K(\rho)$. The claim is proved once we show that $T^{-(n_1+1)}S_{n_1+1} \subset Q^{(k-i)}$. Let $[0, \ell]$ be the shortest interval such that $T^{n_1}[0, \ell]$ contains the entire horizontal boundary of a $Q^{(k-i+n_1)}$. Since this boundary is free, $\ell < \frac{1}{2}\rho^{k-i} + e^{-c'n_1}\rho^{k-i+n_1}$, which is $\approx \frac{1}{2}\rho^{k-i}$. Let $\hat{S}_{n_1}$ be the section of R_{k-i+n_1} from E_{n_1} to the middle of $Q^{(k-i+n_1)}$. Since $b^{\frac{k-i}{4}}K^{K_1(k-i)} << \rho^{k-i+K_1(k-i)}$, we have that $T^{n_1}Q^{(k-i)} \supset \hat{S}_{n_1}$. It remains to show $S_{n_1+1} = T(\hat{S}_{n_1})$, for which we need only to check that $T(\hat{S}_{n_1})$ is a monotone branch. To do that, it suffices to show that for $j < n_1$, $T^j[0, \ell]$ does not contain the horizontal boundary of any $Q^{(k-i+j)}$. Suppose it does for some j. By our choice of n_1, this can happen only if $\ell > \frac{1}{2}\rho^{k-i}$ and $| T^j[\frac{1}{2}\rho^{k-i}, \ell] | \geq \rho^{k-i+j}$, which is impossible, for $| T^j[\frac{1}{2}\rho^{k-i}, \ell] | < e^{-c'(n_1-j)}\rho^{k-i+n_1}$. $\diamondsuit$

Suppose we are guaranteed that P_{n_1} exists. We show next that P_{n_1+1} exists and has the properties in Lemma 9.2(ii). Let γ be the part of a horizontal boundary of P that lies below $[0, \ell]$. From the estimates above, we know that $T^{n_1}\gamma$ is C^0 very near $T^{n_1}[0, \ell]$. Let $\hat{P}_{n_1}$ be the section of $T^{n_1}(P \cap Q^{(k-i)})$ that runs from E_{n_1} to the middle of some $Q^{(t+n_1)} \subset Q^{(k-i+n_1)}$. We claim that $P_{n_1+1} = T(\hat{P}_{n_1})$. Clearly, $P_{n_1+1} \subset S_{n_1+1}$. To see that P_{n_1+1} is a monotone branch, it suffices to observe that for $j < n_1$, $T^{-n_1+j}\hat{P}_{n_1} \cap \mathcal{C}^{(t+j)} = \emptyset$, which is an immediate consequence of the fact that $T^{-n_1+j}\hat{S}_{n_1} \cap \mathcal{C}^{(k-i+j)} = \emptyset$.

We are now ready to show that P_j exists for all $j \leq i$. Suppose that P_{j-1} exists. The only reason why P_j may not exist is that one of its ends has reached its maximum allowed age. Of the two ends of P_{n_1+1}, the one contained in E_{n_1+1} is clearly created earlier, which means that of the two ends of P_{j-1}, the one contained in E_{j-1} is created earlier. It suffices therefore to check that this end survives the step from P_{j-1} to P_j. It does, because it is created later than E_{j-1} and has the same age as E_{j-1}, and, by definition, E_{j-1} has not reached its maximum allowed age.

From here on we argue inductively that the relations in Lemma 9.2(ii) between P_j and S_j hold from $j = n_1 + 2$ to $j = i$. Assume this is true for $j - 1$, and that S_{j-1} has more than one child. Then $S_j = T(\hat{S}_{j-1})$ where $\hat{S}_{j-1}$ is the section of S_{j-1} from E_{j-1} to the middle of some $Q^{(k-i+j-1)}$. Since by inductive assumption P_{j-1} has its ends contained in those of S_{j-1}, we are assured that it traverses some $Q^{(t+j-1)} \subset Q^{(k-i+j-1)}$. Letting $\hat{P}_{j-1}$ be the section of P_{j-1} from its end in E_{j-1} to the middle of $Q^{(t+j-1)}$, we see that $P_j = T(\hat{P}_{j-1})$ has the desired properties.

This completes the proof of Lemma 9.2. $\square$

Proof of Theorem $5'$**:** Let $\mathcal{S}_0 = \{R_0\}$, and assume that for each $n \leq m$, a collection of monotone branches $\mathcal{S}_n$ is selected so that $\tilde{R}_n := \cup_{S \in \mathcal{S}_n} S$ is a neighborhood of the attractor, and each $S \in \mathcal{S}_n$ has the following properties:

(i) $S \in \mathcal{T}_k$ for some $n \leq k \leq (1 + 3\theta)n$;
(ii) if an end of S is of age i, i.e. it is created at time $k - i$, then $2\theta i \leq k - i + 1$.

Note that (ii) is a more stringent requirement than the definition of monotone branches.

The collection $\mathcal{S}_{m+1}$ is defined as follows. For each $S \in \mathcal{S}_m$, if the ends of S have not reached their maximum ages as allowed by (ii) above, then we put the children of S in $\mathcal{S}_{m+1}$. If one of its ends has reached this age, then we choose a collection of branches $\{P\}$ to be specified in the next paragraph, construct from each P a monotone branch P_i as in Lemma 9.2, replace S by $\{P_i\}$ and put the children of P_i in $\mathcal{S}_{m+1}$.

Suppose for definiteness that $S \in \mathcal{T}_k$, and its end E has reached age i where

$$2\theta i = k - i + 1. \tag{11}$$

Let $Q^{(k-i)}$ be the component of $\mathcal{C}^{(k-i)}$ containing $T^{-i}E$. Let $\{P\}$ be the subcollection of $\mathcal{S}_{k-i+1}$ with the property that $P \cap Q^{(k-i)} \neq \emptyset$. Observe immediately that by our inductive hypotheses, P is a monotone branch of $R_{\tilde{k}}$ for some $\tilde{k}$ with

$$k - i + 1 \;\leq\; \tilde{k} \;\leq\; (1+3\theta)(k-i+1). \tag{12}$$

Since $e^{-\alpha(1+3\theta)(k-i+1)} >> \rho^{k-i}$, it follows that P intersects $Q^{(k-i)}$ in a horizontal strip bounded by $C^2(b)$ curves. Note also that since the union of the elements of $\mathcal{S}_{k-i+1}$ covers Ω, we have $\cup P \supset (Q^{(k-i)} \cap \Omega)$.

To justify the validity of this replacement procedure, we need to show that

(a) for each P as above, $P_{K_1(k-i)}$ is well defined where K_1 is as in Lemma 9.2;
(b) P_i is a monotone branch of R_j for some $j \leq (1+3\theta)m$.

Suppose that an end of P, which is a branch of $R_{\tilde{k}}$, is of age $\tilde{i}$. Then

$$2\theta\tilde{i} \;\leq \tilde{k} - \tilde{i} + 1. \tag{13}$$

To prove (a), if suffices to verify that this end lasts another $K_1(k-i)$ iterates, i.e.

$$\theta\;[\tilde{i} + K_1(k-i)] \;\leq\; \tilde{k} - \tilde{i} + 1.$$

This is true because $\theta\tilde{i} \leq \frac{1}{2}(\tilde{k} - \tilde{i} + 1)$ by (13), and

$$\begin{aligned} K_1\theta(k-i) \;&\leq\; K_1\theta(\tilde{k}+1) \;=\; K_1\theta[(\tilde{k}-\tilde{i}+1)+\tilde{i}] \\ &\leq\; K_1\theta(\tilde{k}-\tilde{i}+1)(1+\frac{1}{2\theta}) \;<<\; \frac{1}{2}(\tilde{k}-\tilde{i}+1). \end{aligned}$$

The first inequality above is by (12) and the second by (13).

To prove (b), we need to check that the age of the end of P_i that is contained in E, namely $\tilde{k} + i$, is $\leq (1+3\theta)m$. Observe first that $i \leq m$. This is because the replacement procedure described in Lemma 9.2 does not change the ages of the respective ends of the monotone branch in question. (The age of an end is equal to the "age" of the critical orbits it contains.) Thus it remains to check that

$$\tilde{k} \;\leq\; (1+3\theta)(k-i+1) \;=\; (1+3\theta)2\theta i \;<\; 3\theta i \;\leq\; 3\theta m,$$

the first inequality above coming from (12) and the equality from (11). This completes the proof of Theorem $5'$. □

We mention two bonuses of this construction.

First, it can be seen inductively that for every $S \in \mathcal{S}_n$, if S is a branch of R_k, then the two monotone segments of ∂R_k that bound S must necessarily be from different components of ∂R_0. This is used in Sect. 10.6

Second, we claim that if $\deg(f) \neq 0$, then all of our monotone branches $S \in \mathcal{S}_m$ intersect the attractor Ω in an essential way. Let us call a monotone branch S **essential** if every curve connecting the two monotone segments γ and γ' in ∂S meets Ω. Observe first that $R_0 \in \mathcal{S}_0$ is essential if $\deg(f) \neq 0$. If not, then there exists a curve ω connecting the two components of ∂R_0 that does not meet Ω. Since $\Omega = \cap_k R_k$, this implies that for some k, $R_k \cap \omega = \emptyset$, which is absurd since R_k is not contractible. Assuming that $S \in \mathcal{S}_m$ is essential, then clearly all the monotone branches that comprise $T(S)$ are essential if no end replacements are needed in the next step. If an end replacement is required, then since the new branches are the images of parts of earlier essential branches, they are again essential.

9.4 The coarse geometry of Ω

We explain in the following sequence of pictures exactly how, in finite approximation, the geometry of Ω differs from that of a small tubular neighborhood of a single curve. These pictures are justified by Lemma 9.2. Referring back to Figure 3(c), we may think of the region between the parabolas as made up to two ends belonging to adjacent branches. We know from Lemma 9.2 that long before the tips of these parabolas "separate", that is, before the ends in question reach their maximum allowed age, there are *pre-branches* inside running parallel to these parabolas. In Figure 6 below, the pre-branches are shown in grey, and the zig-zagging cut-lines represent pre-images of the critical set. These cut-lines will become "turns" before the ends in question reach their maximum allowed age.

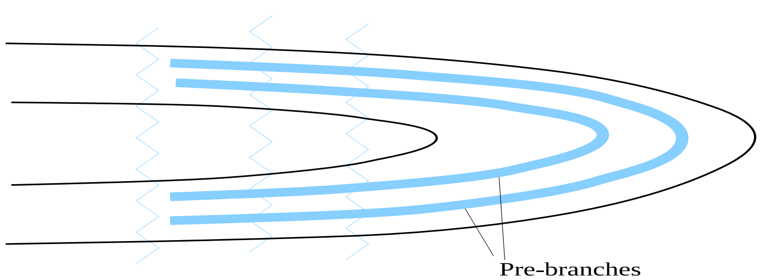
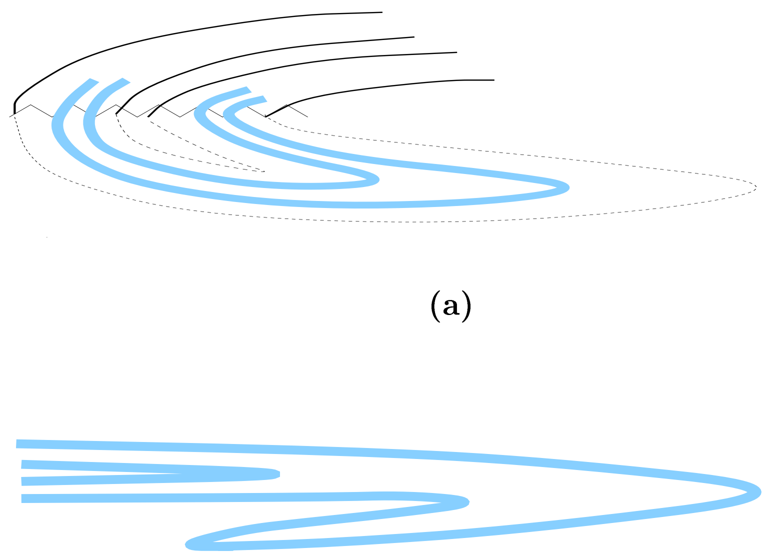


Figure 6 Pre-branches waiting to be released

As this age is reached, the pre-branches are released. Figure 7(a) shows four newly released montone branches grafted onto a branch created earlier. Once released, the new branches evolve independently, resulting possibly in the configuration in Figure 7(b) (cf. Figure 3(d)).



(a)

(b)

Figure 7 Newly released monotone branches evolving independently

The boundaries of every turn (or pair of ends) created every step of the way will in time separate, releasing new branches grafted onto ones born earlier. As the new branches evolve, they create new turns, which again will last for only a finite duration of time. In terms of global geometry, this, in a sense, is the *only* way in which T differs from a 1-dimensional map. Tip replacements are scheduled to take place roughly once every $\sim \log \frac{1}{b}$ iterates, so that in the limit as b tends to 0, no replacement is needed – as it should be for 1-dimensional maps.

10 Symbolic Dynamics and Topological Entropy

The goals of this section are (1) to introduce a natural and unambiguous coding of all points on the attractor Ω for the maps in question, and (2) to use this coding to obtain results on topological entropy and equilibrium states.

10.1 Coding of points on the attractor

Abusing notation slightly, let $x_1 < x_2 < \cdots < x_r < x_{r+1} = x_1$ be the critical points of f in the order in which they appear on the circle, and let $\mathcal{C}_i := \mathcal{C} \cap \mathcal{C}_i^{(0)}$ where $\mathcal{C}_i^{(0)}$ is the component of $\mathcal{C}^{(0)}$ containing x_i. We remark that $\mathcal{C}_i$ can be a fractal set, and that for an arbitrary $z \in R_0$ near $\mathcal{C}_i$, it does not always make sense to think of z as being located on the left or on the right of $\mathcal{C}_i$. The goal of this subsection is to show that points on Ω are special, in that for them this left/right notion is always well defined.

Recall that if $Q^{(k)}$ is a component of $\mathcal{C}^{(k)}$, then $\hat{Q}^{(k)}$ is the component of $R_k \cap \mathcal{C}^{(k-1)}$ containing $Q^{(k)}$. In particular, $\hat{Q}^{(k)} \setminus Q^{(k)}$ has a left and a right component.

Lemma 10.1 *The critical set $\mathcal{C}$ partitions $\Omega \setminus \mathcal{C}$ into disjoint sets $A_1, \cdots, A_r$ as follows:*

- *For $z = (x, y) \notin \mathcal{C}^{(0)}$, $z \in A_i$ if and only if $x_i < x < x_{i+1}$.*
- *For $z \in \mathcal{C}_i^{(0)} \setminus \mathcal{C}_i$, let $Q^{(k)}$ be such that $z \in \hat{Q}^{(k)} \setminus Q^{(k)}$. Then $z \in A_i$ if it lies in the right component of $\hat{Q}^{(k)} \setminus Q^{(k)}$; $z \in A_{i-1}$ if it lies in the left component of $\hat{Q}^{(k)} \setminus Q^{(k)}$.*

Proof: This lemma is an immediate consequence of our description of critical regions (Theorem 1(1)). The sets $\{A_i\}$ are defined by the conditions above. What sets points in Ω apart from arbitrary points in R_0 is that $z \in \Omega$ implies $z \in R_k$ for all k, so that for $z \in \mathcal{C}^{(0)}$, there are only two possibilities: either $z \in \cap_{k \geq 0} \mathcal{C}^{(k)}$, in which case it is a critical point, or there is a largest k such that $z \in \mathcal{C}^{(k-1)}$. In the latter case, it follows from the geometric relation between $\mathcal{C}^{(k)}$ and $\mathcal{C}^{(k-1)}$ that $z \in \hat{Q}^{(k)} \setminus Q^{(k)}$ for some $Q^{(k)}$. □

Lemma 10.1 gives a well defined **address** $a(z)$ for all $z \in \Omega \setminus \mathcal{C}$. We write $a(z) = i$ if $z \in A_i$. Points in $\mathcal{C}$ have two addresses; for example, for $z \in \mathcal{C}_i$, $a(z) =$ both $i - 1$ and i. This in turn allows us to attach to each $z_0 \in \Omega$ with $z_i \notin \mathcal{C}$ for all i an **itinerary** $\iota(z_0) = (\cdots, a_{-1}, a_0, a_1, \cdots)$ where $a_i = a(z_i)$. Orbits that pass through $\mathcal{C}$ have exactly two itineraries as $T^i\mathcal{C} \cap \mathcal{C} = \emptyset$ for all i.

We would like to show that the symbol sequence $\iota(z_0)$ uniquely determines z_0. This may fail in a trivial way: Let $I_i = [x_i, x_{i+1}]$. Then our coding is clearly not unique if for some i, $f(I_i)$ wraps all the way around the circle, meeting some I_j more than once. For simplicity of exposition *we will assume this does not happen.* If it does, it suffices to consider the partition on Ω whose elements correspond to the connected components of $I_i \cap f^{-1} I_j$.

10.2 Coding of monotone branches

Coding of monotone segments of ∂R_k. Observe that points in ∂R_k also have well-defined a-addresses in the spirit of Lemma 10.1: if $z \in \partial R_k \cap \mathcal{C}^{(k)}$, then its location with respect to Γ_k is obvious (except when $z \in \Gamma_k$). This allows us to assign in a unique way a k-block $[a_{-k}, \cdots, a_{-1}]$ to each monotone segment γ of ∂R_k. We write $\iota(\gamma) = [a_{-k}, \cdots, a_{-1}]$.

Coding of monotone branches of R_k. Each $S \in \mathcal{T}_k$, $k > 0$, is associated with a block $\iota(S) = [a_{-k}, \cdots, a_{-1}]$ defined inductively as follows: Let $S \in \mathcal{T}_{k-1}$ be such that

$\iota(S) = [a_{-(k-1)}, \cdots, a_{-1}]$. If $S \cap \mathcal{C}^{(k-1)} = \emptyset$, then it lies between two components of $\mathcal{C}^{(0)}$, say $\mathcal{C}_i^{(0)}$ and $\mathcal{C}_{i+1}^{(0)}$, and $\iota(T(S)) := [a'_{-k}, \cdots, a'_{-1}]$ where $a'_{-1} = i$ and $a'_{-j} = a_{-j+1}$ for $j > 1$. If $S \cap \mathcal{C}^{(k-1)} \neq \emptyset$, then $S = \hat{S}_1 \cup \cdots \cup \hat{S}_n$ where $\hat{S}_1$ is the section of S from one end to the middle of the first $Q^{(k-1)}$ that it meets, $\hat{S}_2$ is the section from the middle of this $Q^{(k-1)}$ to the middle of the next component of $\mathcal{C}^{(k-1)}$ etc., and the a'_{-1}-entry of $\iota(T(\hat{S}_j))$ is defined according to the location of $\hat{S}_j$. Note that this coding of branches in $\mathcal{T}$ is injective, i.e. $S \neq S'$ implies $\iota(S) \neq \iota(S')$, and that if γ and γ' are monotone segments that bound S, then $\iota(\gamma) = \iota(\gamma') = \iota(S)$. Note also that the replacement procedure in Sect. 9.3 corresponds to replacing $[a_k, \cdots, a_{-1}]$ by blocks of the form $[*, \cdots, *, a_{-k}, \cdots, a_{-1}]$.

Coding of arbitrary points in R_0. For points in certain locations of R_0, there is no meaningful way of assigning to it an address as we did in Sect. 10.1. Instead, for each $k \geq 0$, we define the $\tilde{a}^{(k)}$-address(es) of $z \in R_k$ as follows: $\tilde{a}^{(k)}(z)$ has the obvious definition if $z \notin \mathcal{C}^{(k)}$; if $z = (x, y) \in Q^{(k)}$ for some $Q^{(k)} \subset \mathcal{C}_i^{(0)}$, we let $\tilde{a}^{(k)}(z) = i$ if $x > \hat{x} - b^{\frac{k}{4}}$ where $\hat{z} = (\hat{x}, \hat{y})$ is one of the critical points in $\partial Q^{(k)}$; $\tilde{a}^{(k)}(z) = i - 1$ if $x < \hat{x} + b^{\frac{k}{4}}$. Clearly $\tilde{a}^{(k)}$-addresses are not unique: an open set of points in the middle part of each $Q^{(k)} \subset \mathcal{C}_i^{(0)}$ have as their $\tilde{a}^{(k)}$-addresses both $i - 1$ and i.

We further introduce the following notation:

$$\pi_\Omega([a_n, a_{n+1}, \cdots, a_m]) = \{z_0 \in \Omega : a(z_i) = a_i,\ n \leq i \leq m\};$$

$$\pi_{R_0}([a_{-k}, a_{-k+1}, \cdots, a_{-1}]) = \{z_0 \in R_k : \tilde{a}^{(k-i)}(z_{-i}) = a_{-i},\ 1 \leq i \leq k\};$$

"$\tilde{a}^{(k-i)}(z_{-i}) = a_{-i}$" above means a_{-i} is an admissible $\tilde{a}^{(k-i)}$-address of z_{-i}.

Lemma 10.2 *(i) Every $S \in \mathcal{T}_k$, $k \geq 1$, is $= \pi_{R_0}(\iota(S))$ and contains a neighborhood of $\pi_\Omega(\iota(S))$.*
(ii) Given $z_0 \in \Omega$ and $n \in \mathbb{Z}^+$, there exists k with $n \leq k \leq n(1+3\theta)$ and $S = S(z_0, n) \in \mathcal{T}_k$ such that $z_0 \in \pi_\Omega(\iota(S))$.

Proof: That $S = \pi_{R_0}(\iota(S))$ follows inductively from the definitions of these two objects. That S contains a neighborhood of $\pi_\Omega(\iota(S))$ is also obvious inductively. For (ii), we know from Theorem 5′ that there exists $S \in \mathcal{T}_n$ with $z_0 \in S$. The only way one can have $z_0 \notin \pi_\Omega(\iota(S))$ is that at the time S is created, say at time $k - i$, $T^{-i}S$ meets the mid $b^{\frac{k-i}{4}}$-section E of some $Q^{(k-i)}$ and extends to the left of E, while $z_{-i} \in E$ and lies to the "right" of $\Gamma \cap Q^{(k-i)}$ in the sense of Lemma 10.1. Let S_0 be the ancestor of S in $\mathcal{T}_{k-i}$, and let S_1 be the descendant of S_0 that contains the right half of $Q^{(k-i)}$. Our replacement procedure guarantees that there exists $S' \in \mathcal{T}$ that is either a descendant of S_1 or a replacement for a descendent of S_1 which contains z_0. □

Let

$$\Sigma \ := \ \{\mathbf{a} = (a_i)_{i=-\infty}^{\infty} : \iota(z_0) = \mathbf{a} \text{ for some } z_0 \in \Omega\},$$

and let $(\sigma\mathbf{a})_i = (\mathbf{a})_{i+1}$ denote the shift operator. It is easy to check that Σ is a closed subset of $\Pi_{-\infty}^{\infty}\{1, 2, \cdots, r\}$ with $\sigma^{-1}\Sigma \subset \Sigma$. Extending our definition of π_Ω to infinite sequences and writing $\pi = \pi_\Omega$, we have that $\pi(\mathbf{a})$ is the set of all points $z_0 \in \Omega$ with $\iota(z_0) = \mathbf{a}$. The following proposition, whose proof occupies all of the next subsection, completes the proof of Theorem 6.

Proposition 10.1 *For every* $\mathbf{a} \in \Sigma$, $\pi(\mathbf{a})$ *consists of exactly one point, and* $\pi : \Sigma \to \Omega$ *is a continuous mapping.*

Let $B(z_0, \varepsilon)$ denote the ball of radius ε centered at z_0, and let us say $S \in \mathcal{T}_k$ is *compatible* with $\mathbf{a} = (a_i)$ if $\iota(S) = [a_{-k}, \cdots, a_{-1}]$. Proposition 10.1 follows immediately from Lemma 10.2(i) and Proposition 10.1′ below.

Proposition 10.1′ *Given* $\mathbf{a} \in \Sigma$, $z_0 \in \pi(\mathbf{a})$, *and* $\varepsilon > 0$, *there exists* $S \in \mathcal{T}_{n+m}$ *compatible with* $\sigma^n\mathbf{a}$ *such that* $T^{-n}S \subset B(z_0, \varepsilon)$.

10.3 Uniqueness of point in Ω corresponding to each itinerary

We begin with a situation that resembles that in 1-dimension.

Lemma 10.3 *Let* $\mathbf{a}, z_0$ *and* ε *be as in Proposition* 10.1′. *Suppose that for some* k, *the component of* $R_k \cap B(z_0, \varepsilon)$ *containing* z_0, *which we denote by* H, *is bounded by two* $C^2(b)$ *subsegments* γ *and* γ' *of* ∂R_k *cutting across* $B(z_0, \varepsilon)$ *as shown with*

$$\text{Hausdorff distance } (\gamma, \gamma') < \varepsilon^{10}.$$

Then there exists $S \in \mathcal{T}_{n+m}$ *compatible with* $\sigma^n\mathbf{a}$ *such that* $T^{-n}S \subset H$.

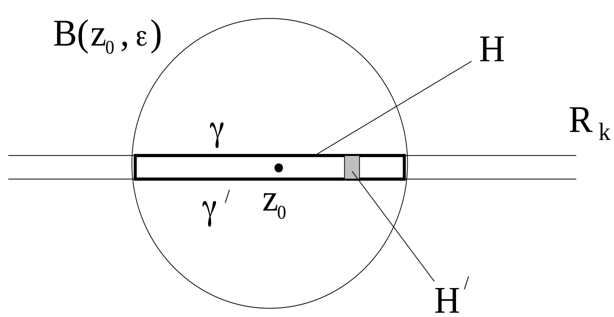


Figure 8 The situation considered in Lemma 10.3

Proof: Our plan of proof is as follows. Assuming $m > k$, so that $T^{-n}S \subset R_k$, we wish to block it from exiting $B(z_0, \varepsilon)$ via, say, the right boundary of H. To this end, we will show that for some section $H' \subset H$ as shown and $k' > 0$, $T^{k'}(H')$ is a component of

$\mathcal{C}^{(k+k')}$, so that the left and right boundaries of H' have incompatible $\tilde{a}^{(k+k')}$-addresses. Assuming $n > k'$, it will follow (using Lemma 10.2) that $T^{-n}S$ cannot meet both the left and right boundaries of H'. Being connected and contained in R_k, $T^{-n}S$ must meet both boundaries of H' in order to exit $B(z_0, \varepsilon)$ from the right. The left boundary of H can be blocked off similarly.

The proof that $T^{k'}H$ crosses a component of $\mathcal{C}^{(k+k')}$ for some k' is similar to that of Sublemma 9.1, but there are two differences: initially at least, we do not know the lengths of $T^j\gamma$ relative to their distances to the critical set, and we must control the shearing between γ and γ' as we iterate. Details of the proof follow.

Consider first the case where $z_0 \notin \mathcal{C}^{(0)}$. Let γ_0 be a subsegment of γ of length $\frac{\varepsilon}{2}$ located half-way between z_0 and the right boundary of H. We first describe how to locate $\gamma_0 \cap H'$. Let n_1 be the first time when $T^i(\gamma_0)$ meets $\mathcal{C}^{(0)}$. If $T^{n_1}\gamma_0$ contains an $I_{\mu j}$ of full length, then we let $\gamma_1 \subset T^{n_1}\gamma_0$ correspond to the longest $I_{\mu j}$ or segment outside of $\mathcal{C}^{(0)}$, whichever is longer. If not, we let $\gamma_1 = T^{n_1}\gamma_0$. In both cases, we let n_2 be the first time when part of $T^{n_2-n_1}\gamma_1$ makes a free return. Choose $\gamma_2 \subset T^{n_2-n_1}\gamma_1$ as before, let n_3 be the first time when part of $T^{n_3-n_2}\gamma_2$ makes a free return, and so on. Using the fact that ∂R_n is controlled (Proposition 5.1), we see that the γ_i increase in length, so that there exists some i_0 such that γ_{i_0} contains an $I_{\mu j}$. From then on, the argument in Sublemma 9.1 produces an i_1 such that $T^{n_{i_1}-n_{i_1-1}}\gamma_{n_{i_1-1}}$ traverses a component of $\mathcal{C}^{(0)}$.

We now proceed to construct H'. Letting τ_0 denote unit tangent vectors to γ, we have that $\|DT^i(\xi_0)\tau_0\| \geq c > 0$ for all $\xi_0 \in \gamma_0$ and $i \leq n_1$. Through each $\xi_0 \in \gamma_0$, therefore, is a stable curve of order n_1 connecting ξ_0 to a point in γ' less than ε^9 away (see Sect. 2.2 and Lemma 2.9). Let H_0 be the region between γ and γ' made up of the union of these stable curves.

Since we do not know how close $T^{n_1}\gamma_0$ gets to the critical set, we cannot continue to claim the expanding property of τ_0 beyond time n_1. Instead, we observe that for $\xi_0 \in \gamma_1$, $\|DT^j(\xi_0)\binom{0}{1}\| \geq 1$ for $j \leq n_2 - n_1$ so that through each $\xi_0 \in \gamma_1$, there is a stable curve of order $n_2 - n_1$. Assuming that these stable curves meet $T^{n_1}\gamma_0'$, we define H_1 to be the region between $T^{n_1}\gamma_0$ and $T^{n_1}\gamma_0'$ spanned by these curves, and check that H_1 can be chosen to be a subregion of $T^{n_1}H_0$.

To justify the last sentence, observe first that if $\gamma_1 \subset I_{\mu_1 j}$, then $e^{-\mu_1} > \varepsilon$. This is true regardless of whether $\gamma_1 = T^{n_1}\gamma_0$. Second, since the contractive field $e_{n_2-n_1}$ near γ_1 makes angles $\sim e^{-\mu_1}$ with $T^{n_1}\gamma_0$ and with $T^{n_1}\gamma_0'$, every point in γ_1 is connected by a stable curve to a point in $T^{n_1}\gamma_0'$ not more than a distance of $(b^{n_1}\varepsilon^9)/e^{-\mu_1} < b^{n_1}\varepsilon^8 << e^{-\mu_1}$ away. This allows us to define H_1. Finally, we may need to trim the edges of H_1 by a length $\sim b^{n_1}\varepsilon^8$ in order to fit it inside $T^{n_1}H_0$. This is easily done since $|\gamma_1| > \min(\varepsilon, \frac{1}{\mu_1^2}e^{-\mu_1})$.

At time n_2, we again do not know how close $T^{n_2-n_1}\gamma_1$ is to the critical set, and so we use $\|DT^j\binom{0}{1}\| \geq 1$ for $j \leq n_3 - n_2$ to construct new stable curves which are then used to construct H_2. Observe that compared to time n_1, the situation has improved:

$|\gamma_2| \geq |\gamma_1|$, and the segments γ_2 and γ_2' are closer than before. We construct H_3, H_4, $\cdots$, until time n_i, when $T^{n_{i_1}-n_{i_1-1}}H_{i_1-1} \supset Q$, a component of $\mathcal{C}^{(k+n_{i_1})}$. Letting $k' = n_{i_1}$ and $H' = T^{-n_{i_1}}(Q)$, the proof for the case $z_0 \notin \mathcal{C}^{(0)}$ is complete.

For $z_0 \in \mathcal{C}^{(0)} \setminus \mathcal{C}$, let j be such that $z_0 \in \hat{Q}^{(j)} \setminus Q^{(j)}$. If k in the statement of the lemma is $\geq j$, repeat the argument above with $n_1 = 0$. If not, replace k by j and ε by $\min(\varepsilon, \frac{1}{2}\rho^j)$ and let $n_1 = 0$. The case of $z_0 \in \mathcal{C}$ is dealt with similarly. $\square$

Recall that for all $z_0 \in \Omega$, at every return to $\mathcal{C}^{(0)}$, z_i is h-related, and bound and fold periods are well defined. (See Section 3 for definitions.)

Proof of Proposition $10.1'$: Let $\mathbf{a} \in \Sigma$ and $z_0 \in \pi(\mathbf{a})$ be given. We wish to arrange for the scenario in Lemma 10.3 at z_0, but it is not possible to do it directly when z_0 is near a "turn". Intuitively, in order for z_0 to be near a "turn", z_{-i} must be near the critical set for some $i > 0$. This motivates the following considerations.

Case 1. There exists arbitrarily large i such that $d_{\mathcal{C}}(z_{-i}) < \rho^k$ for $k \approx K_0(\log \|DT\|)\theta i$ where K_0 is to be specified shortly. Let $\varepsilon > 0$ be given, and let i and k have the relationship above with $\|DT\|^{-i} < \varepsilon^{10}$. Let j be such that $z_{-i} \in \hat{Q}^{(j)} \setminus Q^{(j)}$. Then $j \geq k$. We wish to apply Lemma 10.3 to $z_0' = z_{-i}$ with $\varepsilon' = \|DT\|^{-i}\varepsilon$ and H bounded by ∂R_j. This result transported back to z_0 proves the proposition. To satisfy the hypotheses of Lemma 10.3 at z_0', it suffices to check that the Hausdorff distance between the two horizontal boundaries of $\hat{Q}^{(j)}$ is $< (\|DT\|^{-i}\varepsilon)^{10}$. This is true provided K_0 is chosen to satisfy the inequality

$$b^{\frac{k}{4}} = (b^{\frac{1}{4}K_0\theta \log \|DT\|})^i = (\|DT\|^{\log b^{\frac{1}{4}K_0\theta}})^i < \|DT\|^{-11i} = (\|DT\|^{-i}\varepsilon)^{10}.$$

Case 2. Not Case 1. Note that this means that z_{-i} approaches $\mathcal{C}$ extremely slowly (if at all) as $i \to \infty$. First we observe that with $d_{\mathcal{C}}(z_{-i}) >> b^{\frac{i}{2}}$, z_0 is out of all fold periods from the past. To arrange for the scenario of Lemma 10.3 at z_0, we will show:

(i) there exist $\kappa = \mathcal{O}(1)$ and arbitrarily large i such that $\|DT^j(z_{-i})\binom{0}{1}\| \geq \kappa^j$ for all $j \leq i$;

(ii) the stable curves near z_{-i} when mapped forwards bring with them to z_0 a pair of curves from ∂R_n with z_0 sandwiched in between;

(iii) these curves are $C^2(b)$, they have a minumum length ε_1 independent of i and their Hausdorff distance can be made as small as need be by choosing i large.

We prove (i). Leaving the $\inf_i d_{\mathcal{C}}(z_{-i}) > 0$ case as an exercise, we consider i with $d_{\mathcal{C}}(z_{-i}) \leq d_{\mathcal{C}}(z_{-j})$ for all $0 < j \leq i$. Suppose $d_{\mathcal{C}}(z_{-i}) \approx e^{-\mu}$, so that the ensuing bound period is $> K^{-1}\mu$. Let $w_j = DT^j(z_{-i})\binom{0}{1}$, and let z_{-i+n} be the next free return. Then $\|w_j\| \geq 1$ for $j \leq n$. We argue that w_n splits correctly: If $z_{-i+n} \in \mathcal{C}^{(n)}$, then $d_{\mathcal{C}}(z_{-i+n}) \geq d_{\mathcal{C}}(z_{-i}) \approx e^{-\mu} >> b^{\frac{1}{20}K^{-1}\mu} \geq b^{\frac{1}{20}n}$; if $z_{-i+n} \notin \mathcal{C}^{(n)}$, then it is $\in \hat{Q}^{(j)} - Q^{(j)}$ for some $j < n$. In both cases, Lemma 7.1 applies, and we have $\|w^*_{n+1}\| \geq e^{\frac{cn}{3}}e^{-\mu} \geq e^{(\frac{c}{3}-K)n}$. Since the situation at subsequent free returns is clearly improved ($d_{\mathcal{C}}(\cdot) \geq d_{\mathcal{C}}(z_{-i})$ and the derivative has built up), we have $\|w_j\| \geq e^{(\frac{\varepsilon}{3}-K)j}$ for all $j \leq i$.

To prove (ii), suppose $z_{-i} \in \hat{Q}^{(k)} \setminus Q^{(k)}$ for some k. We consider the stable curve of order i through z_{-i} and let ζ_0 be its intersection with the upper boundary of $\hat{Q}^{(k)}$. A subsegment γ_0 of this upper boundary centered at ζ_0 is constructed by iterating forward i times and trimming whenever necessary so that $T^j\gamma_0$ stays inside three consecutive $I_{\mu\ell}$ for all $j \leq i$. Clearly, stable curves of order i can be constructed through all points in γ_0, and these curves "tie together" the two subsegments of $\partial\hat{Q}^{(k)}$.

We leave it as an exercise to show the existence of ε_1 (which depends only on the slow rate of approach to $\mathcal{C}$ in backward time). The curves brought in are sebsegments of ∂R_{k+i} and they are out of all fold periods. This completes the proof of Proposition 10.1′. □

10.4 Proof of Theorem 2(1)(iii)

We explain how $\Omega = \overline{\cup_{\varepsilon>0}\Omega_\varepsilon}$ follows readily from the ideas in the last two subsections and the surjectivity condition (*) in Sect. 1.2.

In view of Proposition 10.1′, it suffices to show that every $S \in \mathcal{T}$ contains a point in Ω_ε for some $\varepsilon > 0$. Recall the way monotone branches in $\mathcal{T}$ are constructed. Given $S \in \mathcal{T}$, let $\ell > 0$ be the smallest integer such that $T^{-\ell}S \notin \mathcal{T}$. Then $T^{-\ell}S$ contains half of some $Q^{(k)}$. Let H be the middle half of $T^{-\ell}S \cap Q^{(k)}$, with length $\frac{1}{4}\rho^k$. An argument similar to that in Lemma 10.3 but carried on indefinitely in time gives a sequence of domains $H \supset H_1' \supset H_2' \supset \cdots$ and a curve $\omega_0 \subset \cap_{n\geq1}H_n'$ with the following properties:

— ω_0 connects the top and bottom boundaries of $Q^{(k)} \cap T^{-\ell}S$;

— there exists $\varepsilon > 0$ such that $\forall z \in \omega_0,\ d_{\mathcal{C}}(z_n) \geq \varepsilon\ \forall n \geq 0$.

To finish, it suffices to produce $\hat{z}_0 \in \omega_0$ such that $\hat{z}_{-i} \notin \mathcal{C}^{(0)}\ \forall i > 0$. Let D_i be the component of $R_0 \setminus \mathcal{C}^{(0)}$ between the i-th and $(i+1)$-st components of $\mathcal{C}^{(0)}$, and let $\hat{D}_i$ be the union of D_i with the two components of $\mathcal{C}^{(0)}$ adjacent to it. Then we may assume from condition (*) that for every i, there exists j such that $T(D_j) \cap \hat{D}_i$ contains a horizontal strip traversing the full length of $\hat{D}_i$. Suppose $\omega_0 \subset \hat{D}_i$, and let j be as above. Then there is a subsegment $\omega_1 \subset \omega_0$ such that $T^{-1}\omega_1 \subset D_j$ and connects the top and bottom boundaries of D_j. Similarly, we produce for $n = 2, 3, \cdots$ segments $\omega_n \subset \omega_{n-1}$ such that $T^{-n}\omega_n$ is contained in some $D_{j(n)}$ and connects the two horizontal boundaries of $D_{j(n)}$. Let $\hat{z}_0 \in \cap_{n\geq0}\omega_n$. □

10.5 Existence of Equilibrium states

This is a corollary to the symbolic dynamics we have developed. Let $\varphi : R_0 \to \mathbb{R}$ be a continuous function, and let $P(T;\varphi)$ denote the **topological pressure** of T for the potential φ. (See e.g. [Wa], Chapter 9, for definitions and basic facts.) A well known variational principle says that

$$P(T;\varphi) \;=\; sup \;\; P_\nu(T;\varphi)$$

where the supremum is taken over all T-invariant Borel probability measures ν and

$$P_\nu(T;\varphi) := h_\nu(T) + \int \varphi d\nu,$$

where $h_\nu(T)$ denotes the metric entropy of T with respect to ν. An invariant measure for which this supremum is attained is called an **equilibrium state** for $(T;\varphi)$.

Let $\sigma : \Sigma \to \Sigma$ and $\pi : \Sigma \to \Omega$ be as in Theorem 6.

Proof of Corollary 2: Let $\varphi : R_0 \to \mathbb{R}$ be given. We need to prove that there exists ν such that $P_\nu(T;\varphi) = P(T;\varphi)$. Let $\tilde{\varphi}$ be the function on Σ defined by $\tilde{\varphi} = \varphi \circ \pi$. Then $P(T;\varphi) = P(T|\Omega;\varphi|\Omega) \leq P(\sigma;\tilde{\varphi})$. Since $\sigma : \Sigma \to \Sigma$ has a natural finite generator without boundary, $(\sigma,\tilde{\varphi})$ has an equilibrium state which we call $\tilde{\nu}$. Let $\nu = \pi_*\tilde{\nu}$. It suffices to show that $P_\nu(T|\Omega;\varphi|\Omega) = P_{\tilde{\nu}}(\sigma;\tilde{\varphi})$. This follows from the fact that π is one-to-one over $\Omega \setminus \cup T^i\mathcal{C}$, and $\mu(\pi^{-1}(\cup T^i\mathcal{C})) = 0$ for any σ-invariant probability measure μ because $\sigma^i(\pi^{-1}\mathcal{C}) \cap \pi^{-1}\mathcal{C} = \emptyset$ for all $i \in \mathbb{Z}$. $\square$

Since the **topological entropy** of T, written $h_{\text{top}}(T)$, is equal to $P(T;0)$, the discussion above gives immediately

Corollary 10.1 *(i) T has an invariant measure of maximal entropy.*

(ii) Let N_n be the number of distinct blocks of symbols of length n that appear in Σ. Then

$$\lim_{n\to\infty} \frac{1}{n} \log N_n \;=\; h_{\text{top}}(T).$$

10.6 Topological entropy

Topological entropy is, in general, defined in terms of open covers of arbitrarily small diameters, ε-separated or spanning sets. None of the standard definitions is easy to compute with. Corollary 10.1 gives a concrete way to think about this invariant for the class of dynamical systems under consideration. Three other characterizations and estimates of geometric interest are discussed here.

Recall the notion of $\tilde{a}^{(k)}$-addresses for $z \in R_k$ (see Sect. 10.2). For $z_0 \in R_0$, we define its (future) $\tilde{a}$-itinerary to be $(a_i)_0^\infty$ if for each i, $\tilde{a}^{(i)}(z_i) = a_i$. These itineraries are clearly not unique. Let

$$\tilde{N}_n \;=\; \text{the number of } n\text{-blocks appearing in the } \tilde{a}\text{-itineraries of points in } R_0,$$

overcounting whenever ambiguities arise, that is, if an orbit has j different admissible $\tilde{a}$-itineraries of length n, they will be counted as j distinct blocks in $\tilde{N}_n$. Obviously, $N_n \leq \tilde{N}_n$.

Lemma 10.4

$$\limsup_{n\to\infty} \frac{1}{n}\log \tilde{N}_n \;\le\; h_{\rm top}(T).$$

Proof: We fix some arbitrarily small $\varepsilon > 0$, and choose n_0 so that

$$\frac{1}{n_0}\log N_{n_0} < h_{\rm top}(T)+\varepsilon \quad \text{and} \quad \frac{1}{n_0}\log(2n_0) < \varepsilon.$$

Let $n_1 > n_0$ be large enough that $b^{\frac{n_1}{10}}\|DT\|^{n_0} < e^{-\beta n_0}$, so that no orbit segment in R_0 of length $\le n_0$ can pass through the region $D := \{\xi_0 \in \mathcal{C}^{(n_1)} : |\xi_0 - \hat{z}_0| < b^{\frac{n_1}{10}}$ for some $\hat{z}_0 \in \mathcal{C}\cap Q^{(n_1)}(\xi_0)\}$ more than once. For each z_0, let $S_{z_0} = T^{-n_0}S(z_{n_0}, 2n_1)$ where $S(z_{n_0}, 2n_1)$ is as in Lemma 10.2(ii). By part (i) of the same lemma, S_{z_0} is a neighborhood of z_0. Let $n_2 > n_1$ be such that $R_{n_2} \subset \cup_{z_0\in\Omega} S_{z_0}$. Define

$$\begin{aligned}\tilde{N}(n_2, n_2+n_0) \;=\; &\text{ the number of distinct blocks of } [a_{n_2}, \cdots, a_{n_2+n_0-1}]\\ &\text{that appear in the } \tilde{a}\text{-itineraries of all points in } R_0.\end{aligned}$$

Claim 10.1 $\tilde{N}(n_2, n_2+n_0) \le 2n_0 N_{n_0}$

Proof of Claim 10.1: Let $\xi_0 \in R_0$, and let (a_i) be any one of its $\tilde{a}$-itineraries. Let $\xi_{n_2} \in S_{z_0}$ for some $z_0 \in \Omega$, and let $\iota(z_0) = (b_i)$. We compare the two blocks $[a_{n_2}, \cdots, a_{n_2+n_0-1}]$ and $[b_0, \cdots, b_{n_0-1}]$. The ith entry of the first block is an $\tilde{a}^{(n_2+i)}$-address of ξ_{n_2+i}. Since $\xi_{n_2+n_0} \in S = S(z_{n_0}, 2n_1)$, it follows from Lemma 10.2 that the i-th entry of the second block is an $\tilde{a}^{(n(S)-n_0+i)}$-address of ξ_{n_2+i} where $n(S)$ is such that $S \in \mathcal{T}_{n(S)}$. Since the indices in both of these $\tilde{a}$-addresses exceed n_1, they may differ only if $\xi_{n_2+i} \in D$. This can happen at most once in the time period in question. In other words, $[a_{n_2}, \cdots, a_{n_2+n_0-1}]$ and $[b_0, \cdots, b_{n_0-1}]$ can differ in at most one entry, and the difference is either $+1$ or -1. Since $[b_0, \cdots, b_{n_0-1}]$ is one of the sequences counted in N_{n_0}, the claim is proved.

$\diamondsuit$

Similar reasoning shows that $\tilde{N}(n_2+kn_0, n_2+(k+1)n_0) \le 2n_0N_{n_0}$ for all $k \ge 0$, giving

$$\tilde{N}_{n_2+kn_0} \;\le\; K^{n_2}\cdot(2n_0N_{n_0})^k.$$

This combined with the properties we imposed on n_0 at the beginning of the proof gives the desired inequality. $\square$

To complete the proof of Theorem 7(i), recall that P_n is the number of fixed points of T^n in Ω.

Lemma 10.5

$$\lim_{n\to\infty} \frac{1}{n}\log P_n \;=\; h_{\rm top}(T).$$

Proof: Since no point in $\mathcal{C}$ is periodic, there is a one-to-one correspondence between the fixed points of T^n and the periodic symbol sequences of period n in Σ, proving "$\leq$" in the lemma. That

$$\liminf_{n\to\infty} \frac{1}{n} \log P_n \ > \ h_{\rm top}(T) - \varepsilon$$

for every $\varepsilon > 0$ follows from a general theorem of Katok for all C^2 surface diffeomorphisms [K]. □

Perhaps the most concrete geometric quantity of all is the rate of growth of the number of monotone segments of a curve such as ∂R_0. Our next lemma compares this growth rate to the topological entropy of T. Let ∂R_0^+ and ∂R_0^- denote the two components of ∂R_0, and define

$$M_n^{\pm} \ = \ \text{the number of monotone segments in } \partial R_n^{\pm}$$

where "monotone segments" are as defined in Sect. 9.1.

Proof of Theorem 7(ii): First we prove $M_n^{\pm} \leq \tilde{N}_n$. This follows from the fact that for every monotone segment γ in $\partial R_n^{\pm}$, $\iota(\gamma)$ is counted in $\tilde{N}_n$, and the mapping $\gamma \mapsto \iota(\gamma)$ is injective.

To prove the second inequality, we associate to each n-block $[a_{-n}, \cdots, a_{-1}]$ that appears in Σ first a point $z_0 \in \Omega$ with $a(z_{-i}) = a_{-i}$ and then a monotone branch $S = S(z_0, n)$ as in Lemma 10.2. Then $S \in \mathcal{T}_k$ for some k with $n \leq k \leq n(1+\varepsilon_0)$, $\varepsilon_0 = 3(\log \frac{1}{b})^{-1}$. We remarked at the end of Sect. 9.3 that every $S \in \mathcal{T}$ has a boundary component γ^+ in ∂R_k^+ and one in ∂R_k^-. We have thus defined, for each fixed n, a mapping from the set of n-blocks in Σ to the set of monotone segments of ∂R_k^+, $n \leq k \leq n(1+\varepsilon_0)$. This mapping is clearly injective since $\iota(\gamma^+) = \iota(S) = [*, \cdots, *, a_{-n}, \cdots, a_{-1}]$, proving

$$N_n \ \leq \sum_{n\leq k\leq n(1+\varepsilon_0)} M_k^+.$$

From this one deduces easily that

$$\lim \frac{1}{n} \log N_n \ \leq \ (1+\varepsilon_0) \ \liminf \frac{1}{(1+\varepsilon_0)n} \log M^+_{(1+\varepsilon_0)n}.$$

□

Appendix A Examples

A.1 Attractors arising from interval maps including the Hénon attractors

Reduction of Theorem 8 to Theorems 1–7: Let I_0 be a closed interval such that $f(I) \subset \text{int}(I_0) \subset I_0 \subset \text{int }(I)$, and let J_1 and J_2 be the two components of $I \setminus I_0$. Choosing $b_0 << |J_1|, |J_2|$, one obtains easily from the formulas for $T_{a,b}$ in Sect. 1.1 that there exist $K > 0$ and $\hat{\Delta} := [a_0, a_1] \times (0, b_0]$ such that for all $(a,b) \in \hat{\Delta}$, $T_{a,b}$ maps $R := I \times [-Kb, Kb]$ strictly into $I_0 \times [-Kb, Kb]$.

Our plan is to replace $\partial I \times [-Kb, Kb]$ by two curves ω_1 and ω_2 so that each $\omega_i \subset J_i \times [-Kb, Kb]$, joins the top and bottom boundaries of R, and lies on the stable curve of a periodic orbit. We may assume that these periodic orbits stay outside of $\mathcal{C}^{(0)}$. Replacing R by R_0, the subregion of R bounded by ω_1 and ω_2, the situation is now virtually indistinguishable from that of the annlus maps treated in Theorems 1–7: the top and bottom boundaries of R_0 play the role of ∂R_0 in the previous situation, and the left and right boundaries shrink exponentially as we iterate. (There are small differences, such as the existence of monotone branches with one end bounded by images of ω_i. These differences are inessential.)

To produce ω_1 and ω_2, we claim that pre-periodic points of f are dense in I. This claim is justified as follows. First, Misiurewicz maps have no homtervals, so that there is a coding of the orbits of f by a subshift $\sigma : \Sigma \to \Sigma$ with the property that each element of Σ corresponds to the itinerary of exactly one point in I. Second, Σ is the closure of $\cup_n \Sigma_n$ where $\{\Sigma_n\}$ is an increasing sequence of subshifts of finite type, and third, pre-periodic points are dense in shifts of finite type.

To finish, we fix pre-periodic points p_1 and p_2 of f near the middle of J_1 and J_2. Shrinking $\hat{\Delta}$ if necessary, we may assume that for $T_{a,b}$ with $(a,b) \in \hat{\Delta}$, the periodic orbits related to p_1 and p_2 persist and the stable curves through the continuation of p_i have the desired properties. This is possible because the slopes of these stable curves are bounded away from zero (see Sect. ??).

Proof of Corollary 3: For the quadratic family, the transversality condition in Step II in Sect. 1.1 hold at all Misiurewicz points [T]. The nondegeneracy condition in Step IV is obviously satisfied. (To ensure that $f(I) \subset \text{int}(I)$ for some I in the case $a^* = 2$, consider a slightly less than 2.)

A.2 Homoclinic bifurcations

We verify here the conditions in Sect. 1.1 and condition (**) in Sect. 1.2 for homoclinic bifurcations in 2-dimensions, setting the stage to apply Theorems 1–7. See Sect. 1.5 for a more detailed description of the bifurcation in question.

Following [PT], pages 47-51, we assume that linearizing coordinates have been chosen in which g_μ, $\mu \in [0, \mu^*]$, has the following properties:

(i) On $\{|\xi|, |\eta| < 2\}$, g_μ is the linear map

$$g_\mu(\xi, \eta) = (\sigma_\mu \xi, \lambda_\mu \eta)$$

where $0 < \lambda_\mu < 1 < \sigma_\mu$, $\lambda_\mu \sigma_\mu < 1$, and λ_μ, σ_μ depend continuously on μ.

(ii) There exists $N \in \mathbb{Z}^+$ such that g_0^N maps the point $(1,0)$ to $(0,1)$, carrying the unstable curve at $(1,0)$ to a curve making a quadratic tangency with the stable curve at $(0,1)$. Near $(1,0)$, g_μ^N has the form

$$g_\mu^N(\xi, \eta) = (\alpha(\xi - 1)^2 + \beta\eta + \gamma\mu + H_1(\mu, \xi, \eta),\ 1 + H_2(\mu, \xi, \eta)) \tag{14}$$

where $\alpha, \beta, \gamma \neq 0$ are constants. Furthermore, we have that at $(\mu, \xi, \eta) = (0, 1, 0)$, $H_1 = H_2 = 0$, $\partial_\xi H_1 = \partial_\eta H_1 = \partial_\mu H_1 = 0$ and $\partial_{\xi\xi} H_1 = \partial_{\xi\mu} H_1 = \partial_{\mu\mu} H_1 = 0$.

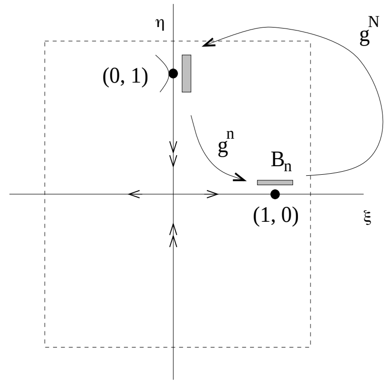


Figure 9 Attractors arising from homoclinic bifurcations

It is not hard to see that for each fixed n, n large, there exist a box B_n (with $\text{diam}(B_n) \to 0$ as $n \to \infty$) and a range of parameters μ (also depending on n) such that $(g_\mu^n \circ g_\mu^N)(B_n) \subset B_n$. The attractors of interest to us have $(n+N)$ components permuted cyclically by g_μ, with one of these components residing in B_n.

To maneuver $g^n \circ g^N$ into the setting in Sect. 1.1, we apply the coordinate transformation $\Phi = \Phi_2 \circ \Phi_1$ where

$$\Phi_1(\xi, \eta) = (\xi - 1, \eta - \lambda^n), \qquad \Phi_2(\xi, \eta) = (-\frac{\sigma^n}{a}\xi, -\frac{\sigma^{2n}}{a}\eta).$$

The purpose of Φ_1 is to shift the center of B_n to the origin. The map Φ_2 magnifies the attractor to unit length; its scaling in the η-direction is chosen with the standard quadratic family in mind. A straightforward computation yields

$$T := \quad \Phi \circ g^n \circ g^N \circ \Phi^{-1} : \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \frac{1}{a}[\sigma^n - \sigma^{2n}(\lambda^n + \mu)] - ax^2 + y - \frac{\sigma^{2n}}{a} H_1(\mu, \Phi^{-1}(x,y)) \\ -\frac{\sigma^{2n}}{a}\lambda^n H_2(\mu, \Phi^{-1}(x,y)) \end{pmatrix}.$$

Letting $a = \Psi(\mu) := \sigma^n - \sigma^{2n}(\lambda^n + \mu)$ and $\tilde{H}_i(a,x,y) := H_i(\mu, \Phi^{-1}(x,y)),\ i = 1,2$, we have

$$T: \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 1 - ax^2 + y - \frac{\sigma^{2n}}{a}\tilde{H}_1(a,x,y) \\ -\frac{\sigma^{2n}}{a}\lambda^n \tilde{H}_2(a,x,y) \end{pmatrix}.$$

Since $\mu = \sigma^{-n} - a\sigma^{-2n} - \lambda^n$, the range of a of interest to us, namely $a \in [1.5, 2)$ (see Appendix A.1), corresponds to a subset of $(0, \mu^*]$ for n large.

What we have so far is a 1-parameter family $\{T_a\}$, which we regard as defined on $U := \{|x|, |y| < 2\}$. The role of $b \to 0$ here is played by $n \to \infty$. Our next task is to choose b (as a function of n) in such a way that $T_{a,b}$ has the form

$$T_{a,b}: \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 1 - ax^2 + y + bu \\ bv \end{pmatrix}$$

where $u = u(a,x,y)$ and $v = v(a,x,y)$ have uniformly bounded C^3-norms. This will put us in the setting of Theorem 8 (see the proof of Corollary 3).

We begin by examing the C^3-norms of $\sigma^{2n}\tilde{H}_1$ and $\sigma^{2n}\lambda^n\tilde{H}_2$. Using the facts that the leading terms in H_1 are $\eta(\xi - 1 + \eta + \mu)$, and that $|\xi| < 3\sigma^{-n}$ and $|\eta| < 3\sigma^{-2n}$ for $(\xi, \eta) \in \Phi^{-1}(U)$, we have $\|\tilde{H}_1\|_{C^0} = \mathcal{O}(\sigma^{-3n})$. Similarly, $\|\tilde{H}_2\|_{C^0} = \mathcal{O}(\sigma^{-n})$. Let $\partial^i, i = 1,2,3$, denote any one of the i-th partial derivatives. Using again the special form of H_1 and the nature of the coordinate transformations Φ and Ψ, we have $\|\partial^i \tilde{H}_1\| = \mathcal{O}(\sigma^{-3n})$ and $\|\partial^i \tilde{H}_2\| = \mathcal{O}(\sigma^{-n})$. Together this gives

$$\|\sigma^{2n}\tilde{H}_1\|_{C^3} < K\sigma^{-n}, \quad \|\sigma^{2n}\lambda^n\tilde{H}_2\|_{C^3} < K(\sigma\lambda)^n.$$

The following choices of b therefore give the desired result:

If $\sigma^2\lambda \leq 1$, let $b = \sigma^{-n}$.
If $\sigma^2\lambda \geq 1$, let $b = (\sigma\lambda)^n$.

This completes the verification of the conditions in Sect. 1.1 for the family $\{T_{a,b}\}$. We finish with the observation that all the results in Section 1 that assume (**) are valid in the present setting: In the case $\sigma^2\lambda \leq 1$, $|\det(DT)| \sim b$, so (**) is satisfied. When $\sigma\lambda \geq 1$, $|\det(DT)| \sim (\sigma\lambda)^n = b^\eta$ where $\sigma^{-1} = (\sigma\lambda)^\eta$. This is condition $(**)'$, a variant of (**) discussed in Sect. 7.2

Appendix B Computational Proofs

B.1 Linear algebra (Sect. 2.1)

Sublemma B.1 *Let e be a unit vector in the most contracted direction of*

$$M = \begin{pmatrix} A & C \\ B & D \end{pmatrix}$$

with $\|Me\| = \lambda^{min}$. Then

$$e = \pm\frac{1}{\rho}(C^2 + D^2 - (\lambda^{min})^2,\ -(AC+BD))\ , \tag{15}$$

$$Me = \pm\frac{1}{\rho}(-A(\lambda^{min})^2 + D\det(M);\ -B(\lambda^{min})^2 - C\det(M)) \tag{16}$$

where ρ is the normalizing constant in (15).

The proof is left as an easy exercise.

Proof of Lemma 2.1: Let O_1 and O_2 be orthogonal matrices such that

$$O_2 M^{(i-1)} O_1 = \begin{pmatrix} \lambda^{min}_{i-1} & 0 \\ 0 & \lambda^{max}_{i-1} \end{pmatrix}.$$

Then the tangent of the angle between e_{i-1} and e_i is given by the slope of the most contracted direction of the matrix

$$M_i O_2^{-1} \begin{pmatrix} \lambda^{min}_{i-1} & 0 \\ 0 & \lambda^{max}_{i-1} \end{pmatrix} := \begin{pmatrix} A & C \\ B & D \end{pmatrix} \begin{pmatrix} \lambda^{min}_{i-1} & 0 \\ 0 & \lambda^{max}_{i-1} \end{pmatrix} = \begin{pmatrix} \lambda^{min}_{i-1}A & \lambda^{max}_{i-1}C \\ \lambda^{min}_{i-1}B & \lambda^{max}_{i-1}D \end{pmatrix}.$$

From Sublemma B.1, we see that the slope in question is equal to

$$\frac{(AC+BD)\lambda^{min}_{i-1}\lambda^{max}_{i-1}}{(C^2+D^2)(\lambda^{max}_{i-1})^2 - (\lambda^{min}_i)^2}\ .$$

This is $\leq \left(\frac{Kb}{\kappa^2}\right)^{i-1}$ because $\lambda_{i-1}^{min}\lambda_{i-1}^{max} = |\det(M^{(i-1)})| < b^{i-1}$, $\lambda_i^{min} < (\frac{b}{\kappa})^i$ and $(C^2+D^2)(\lambda_{i-1}^{max})^2 > K^{-1}\kappa^{2(i-1)}$, the last inequality being a consequence of the fact that $\|M^{(i)}\| > \kappa^i$ and $(A^2+B^2)(\lambda_{i-1}^{min})^2 < K(\frac{b}{\kappa})^{2(i-1)}$. □

Before giving the proof of Corollary 2.2 we state another lemma the proof of which is also a straightforward computation.

Sublemma B.2 *Let*

$$M_i = \begin{pmatrix} A & C \\ B & D \end{pmatrix}, \qquad M^{(j)} = \begin{pmatrix} A_j & C_j \\ B_j & D_j \end{pmatrix}, \qquad j = i-1, i.$$

Then

$$\begin{aligned} \|e_i \times e_{i-1}\| &= \frac{1}{\rho^{(i)}\rho^{(i-1)}} \mid \det(M^{(i-1)})[(AC+BD)(C_{i-1}^2 + D_{i-1}^2) \\ &\quad +(A^2+B^2-C^2-D^2)C_{i-1}D_{i-1}] + \Delta_i \mid \end{aligned} \tag{17}$$

where $\rho^{(i-1)}$ and $\rho^{(i)}$ are the normalizing constants for e_{i-1} and e_i as in Sublemma B.1, and

$$\Delta_i = -(\lambda_i^{min})^2(A_{i-1}C_{i-1} + B_{i-1}D_{i-1}) + (\lambda_{i-1}^{min})^2(A_iC_i + B_iD_i).$$

Observe that each the terms in the numerator of (17) has a factor $|\det(M^{(i-1)})|$, λ_{i-1}^{min} or λ_i^{min}, all of which are $\leq (\frac{b}{\kappa})^{i-1}$. Observe also that if both e_{i-1} and e_i are nearly parallel to the x-axis, then $\rho^{(i)}, \rho^{(i-1)}$ are $> K^{-1}\kappa^{2i}$ (see the proof of Lemma 2.1).

Proof of Corollary 2.2: We begin with some useful derivative estimates. First, we claim that

$$\|\partial^1 M^{(i)}\| < K^i. \tag{18}$$

This is because $\partial^1 M^{(i)}$ is the sum of i terms of the form $M_i \cdots M_{j+1}(\partial^1 M_j)M_{j-1}\cdots M_1$ and the norm of this product is $< K_0^{2i}$. A similar argument gives

$$|\partial^1 \det M^{(i)}| \leq (Kb)^i. \tag{19}$$

Since $\lambda_i^{max} = \|M^{(i)}\|$, it follows from (18) that $|\partial^1\lambda_i^{max}| < K^i$; and since $\lambda_i^{min} = |\det M^{(i)}|/\lambda_i^{max}$, we have $|\partial^1\lambda_i^{min}| < (\frac{Kb}{\kappa^2})^i$.

Pre-composing with a suitable orthogonal matrix as in the proof of Lemma 2.1, we may assume that $\rho^{(i)}, \rho^{(i-1)}$ are $> K^{-1}\kappa^{2i}$. The estimate for $\partial^j\theta_1$ is obtained by differentiating (15). To prove (3), we differentiate (17), and observe using the inequalities above that after differentiation, the numerator is the sum of a finite number of terms each one of which is bounded above by $(\frac{Kb}{\kappa^2})^{i-1}$.

To prove (4), we write

$$\begin{aligned} M^{(i)}e_n &= M^{(i)}e_i + M^{(i)}(e_n - e_i) \\ &= M^{(i)}e_i + \sum_{k=i}^{n-1} M^{(i)}(e_{k+1} - e_k) \end{aligned}$$

and take partial derivative one term at a time. First we have

$$\partial^1 M^{(i)}(e_{k+1} - e_k) = \partial^1 M^{(i)} \cdot (e_{k+1} - e_k) + M^{(i)} \cdot \partial^1(e_{k+1} - e_k).$$

The norm of the first term on the right side is bounded by $(\frac{Kb}{\kappa^2})^k$ because $\|\partial^1 M^{(i)}\| \leq K^i$ and $\|e_{k+1} - e_k\| < (\frac{Kb}{\kappa^2})^k$. The norm of the second term is bounded by $(\frac{Kb}{\kappa^2})^k$ according to (3). It remains to show $\|\partial^1 M^{(i)}e_i\| < (\frac{Kb}{\kappa^2})^i$. This follows by differentiating (16) and using the inequalities above. The proofs for $j = 2, 3$ are similar. □

Sublemma B.3 *Let M_i and M'_i be as in Lemma 2.2, let $m < \frac{n}{2}$, and write*

$$M_{i,m} = M_{i+m} M_{i-1+m} \cdots M_m \ , \quad M'_{i,m} = M'_{i+m} M'_{i-1+m} \cdots M'_m .$$

Then

$$\|M_{i,m} - M'_{i,m}\| < \frac{1}{4}(K\lambda)^m \tag{20}$$

for all i, $0 \le i \le m$.

Proof: Set $\rho_k = \|M_{k,m} - M'_{k,m}\|$. Then

$$\begin{aligned} M_{k+1,m} - M'_{k+1,m} &= M_{k+1+m} M_{k,m} - M'_{k+1+m} M'_{k,m} \\ &= M_{k+1+m}(M_{k,m} - M'_{k,m}) + (M_{k+1+m} - M'_{k+1+m}) M'_{k,m} . \end{aligned}$$

Since $\|M'_{k,m}\| < K_0^k$ and $\|M_{k+1+m} - M'_{k+1+m}\| < \lambda^{k+m}$, we have

$$\rho_{k+1} \le K\rho_k + K^k \lambda^{m+k},$$

which implies (20). □

Proof of Lemma 2.2: ([BC2], p. 108): We prove the assertion for all the indices that are powers of two and leave the rest as an exercise. To prove (b), write $m_j = 2^j$, and let

$$u_j = \frac{w_{m_j}}{\|w_{m_j}\|}, \quad u'_j = \frac{w'_{m_j}}{\|w'_{m_j}\|}$$

where $w_{m_j} = M^{(m_i)} w$ and $w'_{m_j} = M'^{(m_i)} w$. We will show inductively that

$$\|u_j \times u'_j\| < \lambda^{\frac{m_j}{4}}. \tag{21}$$

Assume that (21) is true up to index j. Let

$$A = M_{m_{j+1}-m_j, m_j} \quad \textit{and} \quad A' = M'_{m_{j+1}-m_j, m_j} .$$

Since $\|w_{m_j}\| < K^{m_j}$ and $\|w_{m_{j+1}}\| > \kappa^{m_{j+1}}$, we have

$$\|Au_j\| = \frac{\|w_{m_{j+1}}\|}{\|w_{m_j}\|} > \left(\frac{\kappa^2}{K}\right)^{m_j}, \tag{22}$$

$$\|Au'_j\| \ge \|Au_j\| - \|A\|\|u_j - u'_j\| \ge \left(\frac{\kappa^2}{K}\right)^{m_j} - K^{m_j}\lambda^{\frac{m_j}{4}} \ge \frac{3}{4}\left(\frac{\kappa^2}{K}\right)^{m_j}. \tag{23}$$

Writing $\|A'u'_j\| = \|A'u'_j - A'\hat{u}_j + A'\hat{u}_j - A\hat{u}_j + A\hat{u}_j\|$ where $\hat{u}_j = u_j$ if the angle between u_j and u'_j is smaller than $\frac{\pi}{2}$, $\hat{u}_j = -u_j$ otherwise, we obtain $\|A'u'_j\| \ge \|Au_j\| - \|A\|\|u_j \times u'_j\| - \|A - A'\|$. Using Sublemma B.3 to bound $\|A - A'\|$, we again have

$$\| A'u'_j \| \ge \frac{3}{4}\left(\frac{\kappa^2}{K}\right)^{m_j} . \tag{24}$$

We are now ready to prove (21) for index $j+1$:

$$\begin{aligned} \|u_{j+1} \times u'_{j+1}\| &= \frac{\|Au_j \times A'u'_j\|}{\|Au_j\| \cdot \|A'_j u'_j\|} = \frac{\|Au_j \times (A - A + A')u'_j\|}{\|Au_j\| \cdot \|A'_j u'_j\|} \\ &< \frac{\|Au_j \times Au'_j\|}{\|Au_j\| \cdot \|A'_j u'_j\|} + \frac{\|Au_j \times (A - A')u'_j\|}{\|Au_j\| \cdot \|A'_j u'_j\|}. \end{aligned}$$

The first term is fine since $\|Au_j \times Au'_j\| = |\det(A)|\|u_j \times u'_j\|$ and $\det(A) < b^{m_j}$. To estimate the second term, we use Sublemma B.3 and (22)-(24).

To prove (a), we again let $i = 2^k$. Then for $0 < j \leq k$, we have

$$\|w'_{m_{j+1}}\| = \|w'_{m_j}\|\|A'u'_j\| = \|w'_{m_j}\|\|A'u'_j - Au'_j + Au'_j - A\hat{u}_j + A\hat{u}_j\|,$$

so that

$$\begin{aligned} \frac{\|w'_{m_{j+1}}\|}{\|w'_{m_j}\|} &\geq \|Au_j\| - \|A' - A\|\|u'_j\| - \|A\|\|u'_j - \hat{u}_j\| \\ &= \frac{\|w_{m_{j+1}}\|}{\|w_{m_j}\|}\left(1 - \frac{\|w_{m_j}\|}{\|w_{m_{j+1}}\|}(\|A' - A\| + \|A\|\|u'_j - \hat{u}_j\|)\right). \end{aligned}$$

Using Sublemma B.3 to bound $\|A - A'\|$ and part (b) of this lemma to bound $\|\hat{u}_j - u'_j\|$, we obtain

$$\frac{\|w'_{m_{j+1}}\|}{\|w'_{m_j}\|} \geq \frac{\|w_{m_{j+1}}\|}{\|w_{m_j}\|}(1 - 4^{-m_j}),$$

which implies (a). □

B.2 Stable curves (Sect. 2.2)

On a ball of radius $\frac{\lambda}{2K_0}$ centered at z_0, we have $\|DT\| \geq \frac{\kappa}{2}$ so that e_1, the field of most contracted directions of DT, is well defined. Let γ_1 be the integral curve to e_1 of length $\sim \lambda$ passing through z_0.

To construct γ_2, let B_1 be the $\frac{\lambda^2}{2K_0}$-neighborhood of γ_1. For $\xi \in B_1$, let ξ' be a point in γ_1 with $|\xi - \xi'| < \frac{\lambda^2}{2K_0}$. Then $|T\xi - Tz_0| \leq |T\xi - T\xi'| + |T\xi' - Tz_0| \leq \frac{\lambda^2}{2} + \frac{Kb}{\kappa^2}\lambda < \lambda^2$, so by Lemma 2.2, $\|DT^2\xi\| \geq \frac{\kappa^2}{2}$. This ensures that e_2, the field of most contracted directions for DT^2, is defined on all of B_1. Let γ_2 be the integral curve through z_0 in B_1. We leave it as an exercise to show that the Hausdorff distance between γ_1 and γ_2 is $\mathcal{O}(\frac{b}{\kappa^2}\lambda) << \lambda^2$, so that γ_2 has essentially the same length as γ_1. This uses the fact that e_1 has Lipschitz constant K (Corollary 2.2) and that the angle between e_1 and e_2 is $< \frac{Kb}{\kappa^2}$ (Corollary 2.1).

Next we let B_2 be the $\frac{\lambda^3}{2K_0}$-neighborhood of γ_2 and repeat the argument above to get e_3 and γ_3. Using the Lipschitzness of e_2 and the fact that $\| e_3 \times e_2 \| \leq (\frac{Kb}{\kappa^2})^2$, we conclude again that γ_3 has essentially the same length as γ_2. This process is continued for n steps.

B.3 Curvature estimates (Sect. 2.3)

Recall that

$$k_i(s) = \frac{\|\gamma'_i(s) \times \gamma''_i(s)\|}{\|\gamma'_i(s)\|^3}.$$

Write

$$DT = DT(\gamma_i(s)) = \begin{pmatrix} A & C \\ B & D \end{pmatrix}$$

and

$$X = \begin{pmatrix} <\nabla A, \gamma'_{i-1}> & <\nabla C, \gamma'_{i-1}> \\ <\nabla B, \gamma'_{i-1}> & <\nabla D, \gamma'_{i-1}> \end{pmatrix}$$

where $< \, , \, >$ is the usual inner product. Since $\gamma'_i = DT \cdot \gamma'_{i-1}$ and $\gamma''_i = DT \cdot \gamma''_{i-1} + X \cdot \gamma'_{i-1}$, we have

$$k_i = \frac{1}{\|\gamma'_i\|^3}\|DT \cdot \gamma'_{i-1} \times (DT \cdot \gamma''_{i-1} + X \cdot \gamma'_{i-1})\| \leq \frac{1}{\|\gamma'_i\|^3}(I + II) \tag{25}$$

where

$$I = |\det(DT)| \cdot \|\gamma'_{i-1} \times \gamma''_{i-1}\|, \qquad II = \|DT \cdot \gamma'_{i-1} \times X \cdot \gamma'_{i-1}\|.$$

Term II is degree three homogeneous in γ'_{i-1}. Moreover, the second component of each vector involved in the cross product has a factor b. Thus there exist $K > 0$ such that

$$k_i \le (Kb \cdot k_{i-1} + K \cdot b) \cdot \frac{\|\gamma'_{i-1}\|^3}{\|\gamma'_i\|^3}. \tag{26}$$

Lemma 2.4 follows by recursively applying inequality (26).

B.4 One-dimensional dynamics (Sect. 2.4)

Let δ_0 :=inf$\{d(f^n\hat{x}, C) : \hat{x} \in C, n > 0\}$. We begin with three easy observations:

(i) There exists $k_0 > 0$ such that for all $\delta < \frac{1}{2}\delta_0$, if x is such that $f^n x \in C_\delta$, then $|(f^n)'x| \ge k_0$. This is true because there is an interval (x_1, x_2) containing x on which f^n is monotone and $f^n(x_1, x_2) \supset (\hat{x} - 2\delta, \hat{x} + 2\delta)$ for some $\hat{x} \in C$. It then follows from the negative Schwarzian property that restricted to $f^{-n}(\hat{x} - \delta, \hat{x} + \delta) \cap (x_1, x_2)$, $|(f^n)'| \ge$ some $k_0 > 0$ independent of x.

(ii) There exists $\lambda_0 > 1$ such that for all sufficiently small δ, if $d(x, C) < \delta$, then there exists $p = p(x)$ such that $f^i x \notin C_\delta$ for all $i < p$ and $|(f^p)'x| \ge \lambda_0^p$. This is an easy computation using the fact that the forward critical orbits of f are contained in a uniformly expanding invariant set. Let $\hat{p}(\delta)$ =inf$\{p(x) : d(x, C) < \delta\}$.

(iii) For all sufficiently small δ, there exist $N_1(\delta) \in \mathbb{Z}$ and $\lambda_1(\delta) > 1$ such that if $x, \cdots, f^n x \notin C_\delta$ for some $n > N_1$, then $|(f^n)'x| \ge \lambda_1^n$. This is proved in [M1].

We now prove the assertion in Lemma 2.5. Fix δ_1 sufficiently small for (i)–(iii) above, and with the property that $\lambda_0^{\hat{p}(\delta_1)} >> k_0^{-1}$. Consider $\delta < \delta_1$ and an orbit segment $x, \cdots, f^n x$ with $f^i x \notin C_\delta$ for $i < n$ and $f^n x \in C_\delta$. To estimate $(f^n)'x$, we let n_j be the jth time $f^i x \in C_{\delta_1}$, and let $p_j = p(f^{n_j} x)$. Then $|(f^{p_j})'(f^{n_j} x)| \ge \lambda_0^{p_j}$, and between the times $n_j + p_j$ and n_{j+1}, the derivative is bounded below by $\lambda_1(\delta_1)^{n_{j+1} - (n_j + p_j)}$ if $n_{j+1} - (n_j + p_j) > N_1(\delta_1)$, by k_0 otherwise. The same estimate holds for the initial stretch up to time n_1. Noting that the factor k_0 can be absorbed into $\lambda_0^{p_j}$, we see that $|(f^n)'x| \ge e^{\hat{c}_1 n}$ where $e^{\hat{c}_1}$ can be taken to be slightly smaller than $(\min(\lambda_0, \lambda_1(\delta_1))^{\frac{\hat{p}(\delta_1)}{\hat{p}(\delta_1) + N_1(\delta_1)}}$. Also, $\hat{c}_0$ can be taken to be $k_0 \lambda^{-N_1(\delta_1)}$. This completes the proof of part (ii) of Lemma 2.5.

To prove (i), let $n_q < n$ be the last time $f^i x \in C_{\delta_1}$, and observe that $|(f^{n-n_q})'(f^{n_q} x)| \ge K^{-1} k_0 \delta$ if $n - n_q < N_1(\delta_1)$, $\ge K^{-1} \delta \lambda_1(\delta_1)^{n - n_q}$ otherwise. $\square$

B.5 Critical points inside $\mathcal{C}^{(0)}$ (Sect. 2.6)

Proof of Lemma 2.9: Write

$$\frac{dq_1(s)}{ds} = \partial_x q_1(x, y) \frac{dx(s)}{ds} + \partial_y q_1(x, y) \frac{dy(s)}{ds}. \tag{27}$$

Since γ is b-horizontal, we have $\frac{dx(s)}{ds} \approx 1$ and $|\frac{dy(s)}{ds}| < \mathcal{O}(b) \cdot |\frac{dx(s)}{ds}|$. By (15)

$$q_1(s) = \frac{AC + BD}{C^2 + D^2 - (\lambda^{min})^2}, \tag{28}$$

so

$$\begin{aligned} \partial_x q_1(x, y) &= \frac{A_x C + A C_x + \mathcal{O}(b)}{C^2 + D^2 - (\lambda^{min})^2} - 2\frac{(AC + BD)(CC_x + DD_x + \lambda^{min}\lambda_x^{min})}{(C^2 + D^2 - (\lambda^{min})^2)^2} \\ &:= I + II \end{aligned}$$

where

$$A = F_x + bu_x, \quad C = F_y + bu_y,$$
$$B = bv_x, \quad D = bv_y.$$

We will show that $|I| \geq K^{-1}$ and $|II| = \mathcal{O}(\delta)$. To estimate I, observe that the denominator is $> K^{-1}$, and that for $(x, y) \in \mathcal{C}^{(0)}$, $|AC_x| = \mathcal{O}(\delta)$, while $|A_x C| = |F_{xx} F_y|(1 + \mathcal{O}(b)) \geq K^{-1}$ since $|F_y| > K^{-1}$ (non-degeneracy condition). Term II follows from the fact that its denominator is $\geq K^{-1}$, and $AC + BD = \mathcal{O}(\delta)$. □.

Proof of Lemma 2.10: Using the results in Sect. 2.1 and Lemma 2.9, we have that at $\gamma(s)$ with $|s| < (Kb)^{\frac{m}{2}}$, e_{3m} is defined with $|q_{3m} - q_m| < (Kb)^m$ (Lemma 2.2) and $|\frac{d}{ds} q_m| \geq K^{-1}$ (Corollary 2.2 and Lemma 2.9). Let $\tau(s)$ denote the slope of $\gamma'(s)$, and assume for definiteness that $\frac{d}{ds} q_{3m} > 0$. Then

$$\begin{aligned} q_{3m}((Kb)^{\frac{m}{2}}) - \tau((Kb)^{\frac{m}{2}}) &= (q_{3m}((Kb)^{\frac{m}{2}}) - q_m((Kb)^{\frac{m}{2}})) + (q_m((Kb)^{\frac{m}{2}}) - q_m(0)) \\ &\quad + (q_m(0) - \tau(0)) + (\tau(0) - \tau((Kb)^{\frac{m}{2}})) \\ &\geq -(Kb)^m + K^{-1}(Kb)^{\frac{m}{2}} + 0 - K_1 b (Kb)^{\frac{m}{2}} \\ &\geq \frac{K^{-1}}{2}(Kb)^{\frac{m}{2}}. \end{aligned}$$

Similarly, $q_{3m}(-(Kb)^{\frac{m}{2}}) - \tau(-(Kb)^{\frac{m}{2}}) < 0$, giving a unique critical point of order $3m$ in between. □

Proof of Lemma 2.11: Let $\tau(s)$ be the slope of the tangent vector to γ at $\gamma(s)$, and let $q_m(s)$ be the slope of q_m at $\gamma(s)$. Let $\hat{\tau}(s)$ and $\hat{q}_m(s)$ denote the corresponding quantities at $\hat{\gamma}(s)$. First we claim that

$$|\ \tau(0) - \hat{\tau}(0)\ | \leq 2\sqrt{\varepsilon}. \tag{29}$$

An easy calculation (which we omit) shows that if this was not the case, then γ and $\hat{\gamma}$ would meet at $\gamma(s)$ for some $|s| < \sqrt{\varepsilon}$.

Let $\hat{m}$ be the largest integer $j \leq m$ such that $4K_1\sqrt{\varepsilon} < \|DT\|^{-13j}$. Then by Lemma 2.2, $\|DT^i(\gamma(s))\| > \frac{1}{2}$ for $0 < i < \hat{m}$ and $s \in [-4K_1\sqrt{\varepsilon}, 4K_1\sqrt{\varepsilon}]$. This guarantees that $q_{\hat{m}}$ is defined everywhere on γ and on $\hat{\gamma}$. Let $\hat{\sigma}(s) := \hat{q}_{\hat{m}}(s) - \hat{\tau}(s)$. We have

$$\begin{aligned} |\hat{\sigma}(0)| &\leq |\hat{q}_{\hat{m}}(0) - q_{\hat{m}}(0)| + |q_{\hat{m}}(0) - q_m(0)| + |q_m(0) - \tau(0)| + |\tau(0) - \hat{\tau}(0)| \\ &< K\varepsilon + (Kb)^{\hat{m}} + 0 + 2\sqrt{\varepsilon} \ < \ 3\sqrt{\varepsilon}. \end{aligned}$$

To prove the existence of a critical point of order $\hat{m}$ on $\hat{\gamma}$, we will compare the signs of $\hat{\sigma}$ at the two end points of $\hat{\gamma}$. First,

$$\hat{\sigma}(4K_1\sqrt{\varepsilon}) = \hat{q}_{\hat{m}}(0) + \frac{d}{ds} q_{\hat{m}}(s_1) \cdot 4K_1\sqrt{\varepsilon} - \hat{\tau}(0) - \frac{d}{ds}\hat{\tau}(s_2) \cdot 4K_1\sqrt{\varepsilon}$$

for some $s_1, s_2 \in [0, 4K_1\sqrt{\varepsilon}]$. This is

$$= \hat{\sigma}(0) + (\frac{d}{ds} q_{\hat{m}}(s_1) + \mathcal{O}(b)) \cdot 4K_1\sqrt{\varepsilon}.$$

Since the second term has absolute value $> (K_1^{-1} - \mathcal{O}(b)) \cdot 4K_1\sqrt{\varepsilon} > |\hat{\sigma}(0)|$, it follows that $\hat{\sigma}(4K_1\sqrt{\varepsilon})$ has the same sign as $\frac{d}{ds} q_1$. An analogous computation shows that $\hat{\sigma}(-4K_1\sqrt{\varepsilon})$ has the opposite sign as $\frac{d}{ds} q_1$. □

B.6 Growth of w_i and w_i^* (Sect. 4.2)

Sublemma B.4 *Let z_0 be h-related to $\hat{z}_0 \in \Gamma_{\theta N}$ with bound period $p < \frac{2}{3}N$, and let $w_0 = \binom{0}{1}$. Then for $i \leq p$, $\|w_i^*\| > K^{-1}e^{c''i}$ for some $c'' \approx c$.*

Proof: Let $\hat{w}_i^*$ be as defined in (IA6). Then (IA4) and (IA6) together imply that $\|\hat{w}_i^*\| > \frac{c_0}{2}e^{ci}$. The only difference between $\hat{w}_i^*$ and w_i^* is that contractive fields of order $\ell(\hat{z}_i)$ are used for splitting for the former and $\ell(z_i)$ the latter at returns to $\mathcal{C}^{(0)}$. By Lemma 4.2, $\ell(z_i) = \ell(\hat{z}_i) \pm 1$, so that recombination times may differ by one. This is clearly of no consequence. Assuming these times are synchronized, we observe next that w_i^* has the same direction as $\hat{w}_i^*$. This can be seen inductively (using the nested property of fold periods). Finally, a vector split using a field of order ℓ or $\ell+1$ may differ in length by a factor of $1 \pm \mathcal{O}(b^\ell)$. Thus $\|w_i^*\| \geq (1-\mathcal{O}(b))^i\|\hat{w}_i^*\|$. □

Proof of Lemma 4.6: We may assume z_i is in a fold period, otherwise there is nothing to prove. Let $i_1 < i \leq i_2$ be the longest fold period containing i. By Lemma 4.4, which applies also to controlled orbits satisfying $d_{\mathcal{C}}(z_j) > e^{-\alpha j}$, we have $i_2 - i_1 \leq \varepsilon i$. Let $w_{i_1} = Ae + B\binom{0}{1}$ be the usual splitting. Then

$$\|w_i^*\| \leq K^{i-i_1}|B| \leq K^{i-i_1}\|w_{i_2}^*\| = K^{i-i_1}\|w_{i_2}\| \leq K^{i-i_1}(K^{i_2-i}\|w_i\|) \leq K^{\varepsilon i}\|w_i\|.$$

The first "$\leq$" uses the fact that $\frac{\|w_{j+1}^*\|}{\|w_j^*\|} \leq$ some K, the second uses Sublemma B.4, and the third $\|DT\| \leq K$. The reverse estimate follows from $\|w_i\| \leq K^{i-i_1}\|w_{i_1}\| \leq K^{i-i_1}d_{\mathcal{C}}(z_{i_1})^{-1}\|w_i^*\|$ and $d_{\mathcal{C}}(z_{i_1}) > e^{-\alpha i}$. □

Proof of Lemma 4.7: We give a proof in the case where j exists; the other case is simpler. Let $k \leq i_1 < i_1 + p_1 \leq i_2 < i_2 + p_2 \leq \cdots \leq i_r = j < n$ be defined as follows: we let i_1 be the first return to $\mathcal{C}^{(0)}$ at or after time k, p_1 the bound period of z_{i_1}, i_2 the first return after $i_1 + p_1$, and so on until $i_r = j$. Writing $k = i_0 + p_0$, we have that $\frac{\|w_n^*\|}{\|w_k^*\|}$ is a product of factors of the following three types:

$$I := \frac{\|w_{i_{s+1}}^*\|}{\|w_{i_s+p_s}^*\|}, \qquad II := \frac{\|w_{i_s+p_s}^*\|}{\|w_{i_s}^*\|} \qquad \text{and} \qquad III := \frac{\|w_n^*\|}{\|w_j^*\|}.$$

First we prove the lemma assuming that no fold periods initiated before time k expires between times k and n. By Lemma 2.8, $I \geq c_0 e^{c_1(i_{s+1}-(i_s+p_s))}$. Since $w_{i_s}^*$ splits correctly, we have, by (IA5), $II \geq K^{-1}e^{\frac{c}{3}p_s}$. Moreover, we may assume that c_0 and K above can be absorbed into the exponential estimate for the bound period $[i_s, i_s+p_s]$. For III, let ℓ be the fold period initiated at time j. If $\ell > n-j$, then $III \geq K^{-1}d_{\mathcal{C}}(z_j)e^{c''(n-j)}$ by Sublemma B.4. If not, we split w_j^* into $w_j^* = Ae_{n-j} + B\binom{0}{1}$, noting that e_{n-j} is defined at z_j by Sublemma B.4 and Lemma 4.6. Then $III \geq K^{-1}d_{\mathcal{C}}(z_j)e^{c''(n-j)} - (Kb)^{n-j}$. The last term is negligible because $d_{\mathcal{C}}(z_j) \sim b^{\frac{\ell}{2}} >> (Kb)^{n-j}$. Altogether, this gives $\frac{\|w_n^*\|}{\|w_k^*\|} \geq K^{-1}d_{\mathcal{C}}(z_j)e^{c'(n-k)}$ for some $c' > 0$ as claimed.

In the rest of the proof, we view contributions from fold periods initiated before time k as perturbations of the estimates above, and verify that they are inconsequential. For I, we claim that for each t in question,

$$\frac{\|w_t^*\|}{\|w_{t-1}^*\|} = (1 \pm \mathcal{O}(\sqrt{b}))\frac{\|DT(z_{t-1})w_{t-1}^*\|}{\|w_{t-1}^*\|},$$

so that I has the same estimate as before with possibly a slightly smaller c_1. This claim follows from the fact that when a fold period initiated ℓ steps earlier expires at time t, the vector to rejoin the main term has magnitude $\|DT(z_{t-1})w_{t-1}^*\|\mathcal{O}(b^{\frac{\ell}{2}})$. (See Sect. 2.7)

Next we turn to III, which is similar to and a little more complicated than II. Given z_t and a vector u, we let $u,\ T_*^1(z_t)u,\ T_*^2(z_t)u,\ \cdots$ denote the vectors given by the splitting algorithm for the orbit segment beginning at z_t with initial vector u – neglecting recombinations from fold periods initiated before time t. Then

$$w_n^* = T_*^{n-j}(z_j)w_j^* + \sum_{t=j+1}^{n} T_*^{n-t}(z_t)E_t$$

where E_t is the sum of the vectors to be rejoined at time t. For fixed t, let ℓ be the shortest fold period initiated before k to expire at time t. From Sect. 2.7, we have $\|E_t\| \le (Kb)^{\frac{\ell}{2}}\|w_t^*\|$. Also, since this fold period contains the one initiated at j, we have, by Sect. 4.1, $K\alpha\ell > (n-j)$. Together this gives

$$\|T_*^{n-t}(z_t)E_t\| \le K^{n-t}(Kb)^{\frac{\ell}{2}}\|w_t^*\| \le (Kb^{\frac{1}{K\alpha}})^{n-j}\|w_t^*\|.$$

Assuming inductively that the assertion in the lemma has been proved for shorter time intervals, we have $\|w_t^*\| \le Kd_{\mathcal{C}}(z_{j_t})^{-1}\|w_n^*\|$ where j_t is a return between times t and n. Thus

$$\sum_{t=j+1}^{n} \|T_*^{n-t}(z_t)E_t\| < (n-j)(Kb^{\frac{1}{K\alpha}})^{n-j}e^{\alpha(n-j)}\|w_n^*\| << \|w_n^*\|,$$

which together with our earlier estimate on $\|T_*^{n-j}(z_j)w_j^*\|$ gives the disired result. □

Proof of Lemma 4.8: The case where z_k is not in a fold period is contained in Lemma 4.7. Let $j < k$ be the point in time when the largest fold period covering z_k is initiated. Splitting $w_j = A\begin{pmatrix}0\\1\end{pmatrix} + Be_\ell$ as usual, and noting that $\ell \ge k-j$, we have

$$\|w_k\| \le b^{\frac{\ell}{2}}\|DT\|^{k-j}\|w_j\| + (Kb)^{j-k}\|w_j\| << \|w_j\|.$$

The assertion again follows from Lemma 4.7. □

Proof of Lemma 4.5: The proof proceeds inductively. Consider a bound return z_i, and assume that the w_j^*-vectors split as desired at *all* returns prior to time i. Let $\angle(\cdot,\cdot)$ denote the angle between two vectors, and let $u = \begin{pmatrix}0\\1\end{pmatrix}$.

Case 1. z_i is in a fold period. Let $j < i$ be the largest integer such that the fold period initiated at time j remains in effect at time i, and let $\hat{z}_0 = \phi(z_j)$. We will prove that

$$\angle(DT^{i-j}(z_j)u,\ \tau(\phi(z_i))\ < \frac{3}{2}\varepsilon_0 d_{\mathcal{C}}(z_i).$$

We compare this inequality to

$$\angle(DT^{i-j}(\hat{z}_0)u,\ \tau(\phi(\hat{z}_{i-j}))\ < \varepsilon_0 d_{\mathcal{C}}(\hat{z}_{i-j}),$$

which we know to be true by (IA3). Suppose $\hat{z}_{i-j} \in \mathcal{C}^{(k)}$. Then

- $\angle(DT^{i-j}(z_j)u,\ DT^{i-j}(\hat{z}_0)u) << e^{-\beta(i-j)} << d_{\mathcal{C}}(z_i)$ by (IA6);
- $\angle(\tau(\phi(z_i)),\ \tau(\phi(\hat{z}_{i-j})) < b^{\frac{k-1}{4}} << d_{\mathcal{C}}(z_i)$ by Lemma 4.1;
- $|d_{\mathcal{C}}(z_i) - d_{\mathcal{C}}(\hat{z}_{i-j})| < e^{-\beta(i-j)} + b^{\frac{k-1}{4}} << d_{\mathcal{C}}(z_i)$.

Case 2. z_i is not in any fold period. In this case let $j < i$ be the last free return, so that the bound period initiated at j remains in effect at i and $w_i^* = DT^{i-j}(z_j)w_j^*$. We split

$$w_j^*(z_0) = Ae_{i-j} + Bu;$$

$e_{i-j}(z_j)$ is defined (even though $i-j > \ell(z_j)$) by (IA6) and Lemma 4.6. We proceed as in Case 1 to estimate the angle of splitting for $DT^{i-j}(z_j)u$ at z_i. It remains to check that adding $A \cdot DT^{i-j}(z_j)e_{i-j}$ will only change the angle of $B \cdot DT^{i-j}(z_j)u$ by $<< e^{-\alpha(i-j)} < d_{\mathcal{C}}(z_i)$. This is true because

$$\|A \cdot DT^{i-j}(z_j)e_{i-j}\| < \frac{|B|}{\angle(e, w_j^*)} b^{i-j} < |B| b^{\frac{i-j}{2}} < b^{\frac{i-j}{2}} \|B \cdot DT^{i-j}(z_j)u\|.$$

□

B.7 Distortion during bound periods (Sect. 4.3)

The notation and context are those of Lemma 4.9.

Sublemma B.5

$$\sum_{i=1}^{\mu} K \frac{\Delta_i}{d_{\mathcal{C}}(z_i)} << 1.$$

Proof: Since $|\xi_s - z_s| < e^{-\beta s}$ for all $s < \mu$, we have $\Delta_i < 2e^{-\beta i}$. Let h_0 be large enough that

$$\sum_{i=h_0+1}^{\mu} K \frac{\Delta_i}{d_{\mathcal{C}}(z_i)} < K \sum_{i=h_0+1}^{\infty} \frac{1}{\delta_0} e^{-(\beta-\alpha)i} < K \frac{1}{\delta_0} \frac{e^{-(\beta-\alpha)h_0}}{1 - e^{-(\beta-\alpha)}} << 1$$

and assume δ is small enough that

$$\sum_{i=1}^{h_0} K \frac{\Delta_i}{d_{\mathcal{C}}(z_i)} < (\sum_{i=1}^{h_0} K \frac{1}{\delta_0} (e^{\alpha} \|DT\|)^i) \delta << 1.$$

□

Proof of Lemma 4.9: (cf. [BC2], Lemma 7.8) Assuming the lemma for all $i < \mu$, we give the proof of (7) for step μ; the bound in (8) is proved similarly.

Case 1 No fold period expires at z_μ and $\mu - 1$ is not a return time. In this case $w_\mu^*(\cdot) = DT(\cdot) w_{\mu-1}^*(\cdot)$. Writing $C = DT(z_{\mu-1})$, $C' = DT(\xi_{\mu-1})$,

$$u = \frac{w_{\mu-1}^*(z_0)}{\|w_{\mu-1}^*(z_0)\|} \quad \textit{and} \quad u' = \frac{w_{\mu-1}^*(\xi_0)}{\|w_{\mu-1}^*(\xi_0)\|},$$

we have

$$\begin{aligned} \frac{M'_\mu}{M_\mu} &= \frac{M'_{\mu-1}}{M_{\mu-1}} \cdot \frac{\|C'u'\|}{\|Cu\|} \le \frac{M'_{\mu-1}}{M_{\mu-1}} \left(1 + \frac{\|C'u' - Cu\|}{\|Cu\|}\right) \\ &\le \frac{M'_{\mu-1}}{M_{\mu-1}} \left(1 + \frac{\|C' - C\|}{\|Cu\|} + \frac{\|C(u-u')\|}{\|Cu\|}\right). \end{aligned}$$

Since $\|Cu\| > K^{-1}\delta$, $\|C - C'\| < K|\xi_{i-1} - z_{i-1}|$ and $\|u - u'\| \sim |\theta'_{\mu-1} - \theta_{\mu-1}| < K b^{\frac{1}{2}} \Delta_{\mu-2}$, we have

$$\frac{M'_\mu}{M_\mu} \le \frac{M'_{\mu-1}}{M_{\mu-1}} \cdot \left(1 + K \frac{\Delta_{\mu-1}}{d_{\mathcal{C}}(z_{\mu-1})}\right).$$

Case 2 $\mu - 1$ is a return time. Then

$$w^*_{\mu-1}(z_0) = A(z_{\mu-1}) \cdot e(z_{\mu-1}) + B(z_{\mu-1}) \cdot w_0.$$

Let

$$A_0 = \frac{A(z_{\mu-1})}{\|w^*_{\mu-1}(z_0)\|}; \qquad B_0 = \frac{B(z_{\mu-1})}{\|w^*_{\mu-1}(z_0)\|}.$$

Then since $w^*_\mu(z_0) = B(z_{\mu-1}) \cdot DT(z_{\mu-1})w_0$, we have

$$\frac{M'_\mu}{M_\mu} = \frac{M'_{\mu-1}}{M_{\mu-1}} \cdot \frac{|B'_0|}{|B_0|} \cdot \frac{\|C'w_0\|}{\|Cw_0\|}.$$

Also with $|B_0| \sim d_{\mathcal{C}}(z_{\mu-1})$ and $|B'_0 - B_0| \leq |\theta'_{\mu-1} - \theta_{\mu-1}| + \|e - e'\|$, we get

$$\left| \frac{B'_0}{B_0} - 1 \right| < K \frac{\Delta_{\mu-1}}{d_{\mathcal{C}}(z_{\mu-1})}. \tag{30}$$

For the last ratio,

$$\frac{\|C'w_0\|}{\|Cw_0\|} \leq 1 + K|\xi_{\mu-1} - z_{\mu-1}|.$$

This finishes the computation for the magnitude. We record also the estimate

$$|A_0 - A'_0| < K\Delta_{\mu-1}$$

for use in Case 3.

Case 3 There exists a return time j whose fold period expires at time μ. In this case

$$w^*_\mu(z_0) = B(z_j) \cdot DT^{\mu-j}(z_j)w_0 + A(z_j) \cdot DT^{\mu-j}(z_j)e(z_j).$$

Let

$$B_0 = \frac{B(z_j)}{\|w^*_j(z_0)\|}, \qquad A_0 = \frac{A(z_j)}{\|w^*_j(z_0)\|},$$
$$C = DT^{\mu-j}(z_j)w_0, \qquad Y = DT^{\mu-j}(z_j)e(z_j).$$

As before, all the corresponding quantities for ξ_0 carry a prime. Then

$$\frac{M'_\mu}{M_\mu} = \frac{M'_j}{M_j} \cdot \frac{\|B'_0C' + A'_0Y'\|}{\|B_0C + A_0Y\|} \leq \frac{M'_j}{M_j} \cdot \frac{\|C'\|}{\|C\|} \cdot \frac{|B'_0|}{|B_0|} \cdot \left(1 + \frac{\left\| \frac{C'}{\|C'\|} - \frac{C}{\|C\|} + \frac{A'_0Y'}{B'_0\|C'\|} - \frac{A_0Y}{B_0\|C\|} \right\|}{\left\| \frac{C}{\|C\|} + \frac{A_0Y}{B_0\|C\|} \right\|} \right)$$

Since $\frac{|A_0|}{|B_0|} \sim \frac{1}{d_{\mathcal{C}}(z_j)}$ and $\frac{\|Y\|}{\|C\|} \leq d^2_{\mathcal{C}}(z_j)$, it follows that $\frac{\|A_0Y\|}{\|B_0C\|} << 1$, giving

$$\frac{M'_\mu}{M_\mu} \leq \frac{M'_j}{M_j} \cdot \frac{\|C'\|}{\|C\|} \cdot \frac{|B'_0|}{|B_0|} \cdot \left(1 + 2\left\| \frac{C'}{\|C'\|} - \frac{C}{\|C\|} \right\| + 2\left\| \frac{A'_0Y'}{B'_0\|C'\|} + \frac{A_0Y}{B_0\|C\|} \right\| \right).$$

Since both $\{z_s\}^{\mu}_{s=j}$ and $\{\xi_s\}^{\mu}_{s=j}$ are bound to a critical segment $\{\eta_s\}^{\mu-j}_{s=0}$, $\eta_0 \in \Gamma_{\theta N}$, we have

$$\frac{\|C'\|}{\|C\|} \leq 1 + K \sum_{s=1}^{\mu-j-1} \frac{\hat{\Delta}_s}{d_{\mathcal{C}}(\eta_s)} \leq 1 + K \sum_{s=1}^{\mu-j-1} \frac{\Delta_{s+j}}{d_{\mathcal{C}}(z_{s+j})} = 1 + K \sum_{i=j+1}^{\mu-1} \frac{\Delta_i}{d_{\mathcal{C}}(z_i)}$$

where

$$\hat{\Delta}_s = \sum_{j=1}^{s} (Kb)^{\frac{j}{4}} |z_{s-j} - \xi_{s-j}|. \tag{31}$$

The factor $\frac{|B_0'|}{|B_0|}$ is estimated in (30). This term has no cumulative effect because it is a one-time addition to the exponent in the distortion formula for any given return. Next

$$\left\| \frac{C'}{\|C'\|} - \frac{C}{\|C\|} \right\| < \hat{\theta}$$

where $\hat{\theta}$ is the angle between C and C', which is smaller than $\hat{\Delta}_{\mu-j-1}$. Now

$$\left\| \frac{A_0'Y'}{B_0'\|C'\|} - \frac{A_0Y}{B_0\|C\|} \right\| \leq \frac{|A_0|}{|B_0|} \cdot \frac{\|Y'-Y\|}{\|C\|} + \left\| \frac{A_0}{B_0\|C\|} - \frac{A_0'}{B_0'\|C'\|} \right\| \|Y'\|.$$

For the first term we have

$$\frac{|A_0|}{|B_0|} \sim \frac{1}{d_{\mathcal{C}}(z_j)}, \qquad \|Y'-Y\| \leq (Kb)^{\mu-j}|\xi_j - z_j|,$$

and $\|C\| > 1$, where the estimate on $\|Y'-Y\|$ is from (4) in Corollary 2.2. For the second term,

$$\begin{aligned} \left\| \frac{A_0}{B_0\|C\|} - \frac{A_0'}{B_0'\|C'\|} \right\| \|Y'\| &\leq (Kb)^{\mu-j} \frac{|A_0'|}{|B_0'|} \cdot \frac{1}{\|C\|} \cdot \left(\left| \frac{A_0}{A_0'} \cdot \frac{B_0'}{B_0} - 1 \right| + \left| 1 - \frac{\|C\|}{\|C'\|} \right| \right) \\ &\leq \frac{(Kb)^{\mu-j}}{d_{\mathcal{C}}(z_j)} \left(\frac{|A_0|}{|A_0'|} \left| \frac{B_0'}{B_0} - 1 \right| + \left| \frac{A_0}{A_0'} - 1 \right| + \left| 1 - \frac{\|C\|}{\|C'\|} \right| \right) \end{aligned}$$

We again estimate term by term: For the first term,

$$\frac{(Kb)^{\mu-j}}{d_{\mathcal{C}}(z_j)} \cdot \frac{|A_0|}{|A_0'|} \cdot \left| \frac{B_0'}{B_0} - 1 \right| \leq \frac{(Kb)^{\mu-j}}{d_{\mathcal{C}}(z_j)} \cdot \frac{\Delta_j}{d_{\mathcal{C}}(z_j)} < \frac{(Kb)^{\frac{\mu-j}{2}}}{d_{\mathcal{C}}(z_j)} \cdot \Delta_j$$

because $b^{\frac{\mu-j}{2}} < d_{\mathcal{C}}(z_j)$ by the definition of fold period. For the second term,

$$\frac{(Kb)^{\mu-j}}{d_{\mathcal{C}}(z_j)} \cdot \left| \frac{A_0 - A_0'}{A_0'} \right| \leq \frac{(Kb)^{\mu-j}}{d_{\mathcal{C}}(z_j)} \Delta_j.$$

Finally, for the third term, we have

$$\frac{(Kb)^{\mu-j}}{d_{\mathcal{C}}(z_j)} \left| 1 - \frac{\|C'\|}{\|C\|} \right| \leq \frac{(Kb)^{\mu-j}}{d_{\mathcal{C}}(z_j)} \sum_{s=1}^{\mu-j-1} \frac{\hat{\Delta}_s}{d_{\mathcal{C}}(z_{s+j})}$$

where $\hat{\Delta}_s$ is as in (31). We also have $b^{\frac{\mu-j}{2}} < d_{\mathcal{C}}(z_{s+j})$, for no fold period starting at time $s+j$ extends beyond index μ. Also, $b^{\frac{\mu-j}{2}} \cdot \hat{\Delta}_s \leq \Delta_j$ for all s, $0 < s < \mu - j$. Therefore the third term is again bounded by $K\frac{\Delta_j}{d_{\mathcal{C}}(z_j)}$.

Observe further that if we replace z_0 by another point ξ_0' which is bounded to z_0, the same argument above continues to work with $\Delta_i(\xi_0, z_0)$ replaced by $\Delta_i(\xi_0, \xi_0')$. This completes the proof. □

B.8 Quadratic behavior (Sect. 4.3)

Let $\xi_0(s)$ and z_0 be as in Lemma 4.11 We begin with the following

A priori estimate on $\xi_\mu(s) - z_\mu$ **:** (cf. [BC2], p.144-147) Let $t_0(s)$ be a unit vector to γ at $\xi_0(s)$, and let $t_\mu = DT^\mu t_0$. We split t_0 using e_μ to get

$$t_0 = A_0 e_\mu + B_0 \binom{0}{1}, \quad \text{so that} \quad t_\mu = A_0 DT^\mu e_\mu + B_0 w_\mu.$$

Writing

$$w_\mu = w_\mu(0) + (w_\mu - w_\mu(0)) = w_\mu(0) + (w_\mu^* - w_\mu^*(0)) + (E_\mu - E_\mu(0))$$

where

$$E_\mu = \sum_{j \in S_\mu} A_j DT^{\mu-j} e_{\ell_j}$$

and S_μ is the collection of j such that the fold period begun at time j extends beyond time μ, we have

$$\xi_\mu(s) - z_\mu \;=\; \int_0^s t_\mu(u)du \;=\; w_\mu(0)\int_0^s B_0(u)du \;+\; I \;+\; II \;+\; III \tag{32}$$

where

$$I = \int_0^s A_0 DT^\mu e_\mu, \quad II = \int_0^s B_0(w_\mu^* - w_\mu^*(0)), \quad III = \int_0^s B_0(E_\mu - E_\mu(0)).$$

Since $A_0 \approx 1$, $\|I\| \le (Kb)^\mu s$. We claim that

$$\|II\|,\ \|III\| \ \le K e^{2\alpha\mu} \|w_\mu^*(0)\| \int_0^s u\ (\sup_{i \le \mu} |z_i - \xi_i(u)|)\ du.$$

The norm of II is estimated using the distortion estimate in Lemma 4.9. To estimate $\|III\|$, we have, for each $j \in S_\mu$,

$$\begin{aligned} &\|A_j DT^{\mu-j}(\xi_j)e(\xi_j) - A_j(0) DT^{\mu-j}(z_j)e(z_j)\| \\ \le\ & (Kb)^{\mu-j}|A_j - A_j(0)| + |A_j(0)|\ \|DT^{\mu-j}(\xi_j)e(\xi_j) - DT^{\mu-j}(z_j)e(z_j)\|. \end{aligned}$$

From the distortion estimate in appendix B.7,

$$|A_j - A_j(0)| < K\|w_j^*(0)\| e^{2\alpha j} \sup_{i \le j} |z_i - \xi_i|.$$

For the second term we have $|A_j(0)| \le \|w_j^*(0)\| e^{\alpha j}$ because $w_j^*(0)$ splits correctly at time j, and $\|w_j^*(0)\| \le \frac{1}{\delta} e^{\alpha\mu} \|w_\mu^*(0)\|$ by Lemma 4.7. Finally, $\|DT^{\mu-j}(\xi_j)e(\xi_j) - DT^{\mu-j}(z_j)e(z_j)\| \le (Kb)^{\mu-j}|\xi_j - z_j|$ by Corollary 2.2, and $B_0(s) \approx 2K_1 s$.

Proof of Lemma 4.11: We will show that for the μ and s in question, the first term in (32) is the dominating one. For those s with $p = p(\xi_0(s)) \le n_0$, the entire action takes place at a distance $> \frac{1}{2}\delta_0$ from the critical set (see the remark after Definition 3.5). This case is straightforward and is left to the reader. We consider here only those s with $p > n_0$.

Define

$$U_\mu \ := \ K e^{4\alpha\mu} \sup_{j \le \mu} \|w_j^*(0)\|$$

where K is the constant in the bound for $\|II\|$ and $\|III\|$ above. Let μ_0 be large enough that $e^{7\alpha\mu_0} e^{-\beta\mu_0} << 1$. By taking n_0 sufficiently large, we may assume $U_{\mu_0} s^2 << 1$ for all the s in question. We will show inductively first the weaker statement

(i) $|\xi_j(s) - z_j| < U_j s^2$
and then the stronger statement
(ii) $|\xi_j(s) - z_j| = K_1(1 \pm \varepsilon_1)\|w_j(0)\|s^2$.

Assume that (i) and (ii) have been proved for all $j < \mu$. To prove them for $j = \mu$, we need the preliminary estimate $U_\mu s^2 << 1$. It suffices to consider the case $\mu > \mu_0$. Observe first that for $\mu \le n_0$, one has the trivial estimate

$$\sup_{j\le\mu} \|w_j^*(0)\| \le K\|w_{\mu-1}(0)\|,$$

and if $\mu > n_0$, then using Lemmas 4.7, 4.6 and the fact that $\frac{1}{\delta} < e^{\alpha\mu}$, one again has

$$\sup_{j\le\mu} \|w_j^*(0)\| < e^{\alpha\mu}\|w_\mu^*(0)\| \le e^{2\alpha\mu}\|w_\mu(0)\| \le Ke^{2\alpha\mu}\|w_{\mu-1}(0)\|.$$

This combined with (ii) for step $\mu - 1$ gives

$$U_\mu s^2 \le Ke^{4\alpha\mu}\ (Ke^{2\alpha\mu}\|w_{\mu-1}(0)\|)s^2 \approx K^2 e^{6\alpha\mu} K_1^{-1}|\xi_{\mu-1}(s) - z_{\mu-1}|$$
$$< K^2 e^{6\alpha\mu} K_1^{-1} e^{-\beta(\mu-1)} << 1.$$

Noting that $\int_0^s B_0 \approx K_1 s^2$, we see from our *a priori* estimate that

$$|\xi_\mu(s) - z_\mu| \ \le\ \|w_\mu(0)\|K_1 s^2 \ +\ (Kb)^\mu s \ +\ 2U_\mu \int_0^s Ku\ [\sup_{i<\mu} |z_i - \xi_i(u)|]\ du.$$

With the quantity inside square brackets being $< U_\mu u^2$ by (i) from the previous step, this is

$$< (U_\mu s^2)e^{-\alpha\mu} + (Kb)^{\frac{\mu}{2}} s^2 + K(U_\mu s^2)^2 \ <\ U_\mu s^2.$$

The proof of (ii) for step μ now follows immediately. □

B.9 Proof of Lemma 6.2 (Sect. 6.2)

We begin with a scenario for which one sees easily that the assertion in this lemma holds: Suppose for $1 \le j \le i - s$, $\|DT^j(z_s)\| > \kappa^j$ for some $\kappa >> b^{\frac{1}{2}}$, and that z_s is bounded away from $\mathcal{C}^{(0)}$. Then $e_{i-s}(z_s)$ is well defined and has slope $> K^{-1}$. Suppose, in addition, that z_s is out of all fold periods, so that w_s is a b-horizontal vector. Then

$$\|DT^{i-s}(z_s)\|\ \|w_s\| \le K\|DT^{i-s}(z_s)w_s\| = \|w_i\|.$$

This together with $\|w_s\| > c^{c''s}$ (which follows from $\|w_s^*\| > e^{cs}$) gives the desired estimate.

Now, intuitively, the behavior of $\|DT^j(z_s)\|$ is a little different just before or after a return to $\mathcal{C}^{(0)}$. This motivates the following definition: If t is a return time to $\mathcal{C}^{(0)}$ for z_0, let ℓ_t denote its fold period and let $I_t := (t - 5\ell_t, t + \ell_t)$.

Claim B.1 *By modifying I_t slightly to $\tilde I_t = (t - (5 \pm \varepsilon)\ell_t,\ t + (1 \pm \varepsilon)\ell_t)$, we may assume they have a nested structure.*

Proof of Claim B.1: We consider $t = 0, 1, 2, \cdots$ in this order, and determine, if t is a return time, what $\tilde I_t$ will be. The right end point of $\tilde I_t$ is determined by the following algorithm: Go to $t + \ell_t$, and look for the largest t' *inside the bound period initiated at time* t with the property that $t' - 5\ell_{t'} < t + \ell_t$. If no such t' exists, then $t + \ell_t$ is the right end point of $\tilde I_t$. If t' exists, then the new candidate end point is $t' + \ell_{t'}$, and the search continues. For the same reasons as in Sect. 4.1, the increments in length are exponentially small and the process teminates.

As for the left end point of $\tilde{I}_t$, it is possible that $t - 5\ell_t \in \tilde{I}_{t'}$ for some t' the bound period initiated at which time does not extend to time t. This means that $\ell_{t'} << \ell_t$, and since we assume a nested structure has been arranged for $\tilde{I}_{t'}$ for all $t' < t$, we simply extend the left end of $\tilde{I}_t$ to include the largest $\tilde{I}_{t'}$ that it meets. $\diamondsuit$

Let us assume this nested structure and write I_t instead of $\tilde{I}_t$ from here on.

Claim B.2 *For $s \notin \cup I_t$, we have, for all j with $1 \leq j < i - s$,*

$$\|w_{s+j}\| \geq b^{\frac{j}{9}} \|w_s\|.$$

Proof of Claim B.2: We fix j and let r be such that z_r makes the deepest return between times s and $s+j$. Let j' be the smallest integer $\geq j$ such that $z_{s+j'}$ is outside of all fold periods. Then from Sect. 4.2 , it follows that

$$\|w_{s+j}\| \geq K^{-(j'-j)} \|w_{s+j'}\| \geq K^{-K(j'-j)} d_{\mathcal{C}}(z_r) \|w_s\| \approx K^{-K(j'-j)} b^{\frac{\ell_r}{2}} \|w_s\|. \tag{33}$$

Case 1. $s + j \notin I_r$. In this case, $6\ell_r < j$ since I_r is sandwiched between s and $s + j$, and $j' - j \leq \ell_r$ because r is the deepest return. The rightmost quantity in (33) is therefore $> K^{-\ell_r} b^{\frac{\ell_r}{2}} \|w_s\| > b^{\frac{j}{9}} \|w_s\|$.

Case 2. $s + j \in I_r$. The argument is as above, except we only have $5\ell_r < j$.
This completes the proof of the claim. $\diamondsuit$

As noted in the first paragraph, Claim B.2 implies the assertion in Lemma 6.2 for $s \notin \cup I_t$ provided z_s is bounded away from $\mathcal{C}^{(0)}$. For $s \leq n_0$, this is always the case. For $s > n_0$, the slope of the contractive vector is only guaranteed to be $> K^{-1}\delta$, which in principle introduces a copy of $\frac{1}{\delta}$ to the right side of the inequality in Lemma 6.2. This factor, however, can be absorbed into the exponential by taking n_0 sufficiently large.

It remains to prove the lemma for $s \in \cup I_t$. Let I_r be the maximal I_t-interval containing s. Observe that $6\ell_r < K\alpha\theta s$ (recall that z_0 obeys (IA2)) and $\|w_i\| > e^{c'' i}$ for some $c'' > 0$. If $i \in I_r$, then $\|DT^{i-s}(z_s)\| < K^{6\ell_r} << e^{\frac{1}{2}c'' i} < e^{-\frac{1}{2}c'' s} e^{c'' i} < e^{-\frac{1}{2}c'' s} \|w_i\|$. If $i \notin I_r$, let $s' = r + \ell_r$. Then $s' \notin \cup I_t$, and

$$\|DT^{i-s}(z_s)\| \leq \|DT^{s'-s}(z_s)\| \cdot \|DT^{i-s'}(z_{s'})\| \leq K^{6\ell_r} \cdot K e^{-c' s'} \|w_i\|.$$

B.10 Initial data for critical curves (Sect. 6.3)

Proof of Lemma 6.4: Let $J_i := [\hat{a} - \rho^{2i}, \hat{a} + \rho^{2i}]$. Assume for all $i < n$ that the following has been proved:

(i) $J_i \subset \tilde{\Delta}_i$;

(ii) $\Gamma_{i,i}(\hat{a})$ has a smooth continuation on J_i and $\mathcal{C}^{(i)}$ deforms continuously;

(iii) for all $z \in \Gamma_{i,i}$, $\|\frac{dz}{da}\| \leq K^i$.

We now prove (i)–(iii) for $i = n$.

First we verify that for all $a \in J_n$ and $z_0 \in \Gamma_{n-1,n-1}$, (IA2) and (IA4) hold up to time n. This is true for $a = \hat{a}$. For $a \in J_n$, $|z_0(a) - z_0(\hat{a})| < \rho^{2n} K^{n-1}$, so that $|z_j(a) - z_j(\hat{a})| < \rho^{2n} K^{2n}$ for all $j \leq n$. We may assume that ρK is $<< 1$. It then follows from the discussion at the beginning of Sect. 6.3.1 that $\Gamma_{n,n}(a)$ is well defined, proving (i).

To prove (ii), we fix an arbitrary $\tilde{a} \in J_n$, a component $Q^{(n-1)}$ of $\mathcal{C}^{(n-1)}$, and show that every segment of $\partial R_n(\tilde{a}) \cap Q^{(n-1)}(\tilde{a})$ has a continuation to a segment of $\partial R_n(a) \cap Q^{(n-1)}(a)$. Let $\tilde{\omega}$ be a segment of this kind, and let $\omega(a) := T_a^n(2T_{\tilde{a}}^{-n}\tilde{\omega})$ where $2T_{\tilde{a}}^{-n}\tilde{\omega}$ refers to the segment in ∂R_0 with the same midpoint as $T_{\tilde{a}}^{-n}\tilde{\omega}$ and two times as long. Observe that as we vary our parameter from $\tilde{a}$ to a, the segment $\omega(a)$ cannot intersect the horizontal boundaries of $Q^{(n-1)}(a)$. Thus the only way $\omega(a)$ can fail to traverse fully $Q^{(n-1)}(a)$ is that it has moved sufficiently far from $\omega(\tilde{a})$ in the *horizontal* direction. We know this cannot happen because $|T_{\tilde{a}}^n - T_a^n| \le \rho^{2n}K^n$ which is $<< \rho^n$. This proves (ii).

It remains to prove (iii). Consider $\bar{z}(a) = (\bar{x}(a), \bar{y}(a)) \in \Gamma_{n,n}(a)$, and let $y = \psi(x, a)$ denote the $C^2(b)$-curve in $\partial R_n(a)$ containing $\bar{z}(a)$. Then

$$q_n(\bar{x}(a), \psi(\bar{x}(a), a), a) = \partial_x \psi(\bar{x}(a), a)$$

where $q_n(x, y, a)$ is the slope of the contractive vector of order n at $z = (x, y)$. Taking derivative with respect to a on both sides of the last equation, we have

$$\partial_x q_n \cdot \frac{d\bar{x}}{da} + \partial_y q_n \cdot (\partial_x \psi \cdot \frac{d\bar{x}}{da} + \partial_a \psi) + \partial_a q_n = \partial_{xx}\psi \cdot \frac{d\bar{x}}{da} + \partial_{ax}\psi.$$

This implies

$$\frac{d\bar{x}}{da} = \frac{\partial_{xa}\psi - \partial_y q_n \cdot \partial_a \psi - \partial_a q_n}{\partial_x q_n + \partial_y q_n \cdot \partial_x \psi - \partial_{xx}\psi}. \tag{34}$$

Since $\partial_x \psi$, $\partial_{xx}\psi = \mathcal{O}(b)$, $|\partial_x q_n| > K_1$ and $|\partial_y q_n| < K$ (see Corollary 2.2 and Lemma 2.9), the denominator on the right-hand side is bounded away from zero. In the numerator, we have $|\partial_y q_n|, |\partial_a q_n| < K$, and we need to estimate $\partial_a \phi(x, a)$ and $\partial_{ax}\phi(x, a)$.

For this purpose we write the horizontal curve $y = \psi(x, a)$ in parametric form $x = X(t, a), y = Y(t, a)$ where t is the x-coordinate of $T_a^{-n}(x, y)$, i.e., $(t, \pm b) \in \partial R_0$ and

$$(X(t, a), Y(t, a)) = T_a^n(t, \pm b).$$

Let $t = t(x, a)$ be defined by $\psi(x, a) = Y(t(x, a), a)$. Then

$$\partial_a \psi = \partial_t Y(t, a) \cdot \partial_a t(x, a) + \partial_a Y(t, a).$$

Clearly, $|\partial_t Y(t, a)| < K^n b$ and $|\partial_a Y(t, a)| < K^n$. One way to bound $\partial_a t(x, a)$ is to write it as

$$\partial_a t(x, a) = -\frac{\partial_a X(t, a)}{\partial_t X(t, a)}.$$

Since $|\partial_t X(t, a)| > 1$ (recall that $T_a^n|\partial R_0$ is controlled), this term is also $< K^n$. Similar considerations yield $|\partial_{ax}\psi(x, a)| < K^n$. We have proved $\frac{d\bar{x}}{da} < K^n$. The corresponding estimate for $\frac{d\bar{y}}{da}$ follows immediately since

$$\frac{d\bar{y}}{da} = \partial_x \psi \frac{d\bar{x}}{da} + \partial_a \psi.$$

We record an estimate needed in the proof of Lemmas 6.5 and 6.6. Taking derivatives with respect to a one more time on both sides of (34) and estimating corresponding terms (using again Corollary 2.2 and Lemma 2.9), we obtain $|\frac{d^2\bar{x}}{da^2}| < K^n$. This estimate requires that $T_{a,b}$ be C^3.

Recall the following lemma due to Hadamard:

Lemma B.1 (Hadamard) *Let $g \in C^2(0, L)$ be such that $|g| \le M_0$ and $|g''| < M_2$. If $4M_0 < L^2$, then*

$$|g'| \le \sqrt{M_0}(1 + M_2).$$

Proof of Lemma 6.5: Let $z^{(n)} = (x^{(n)}, y^{(n)})$. For our pusposes, let $g(a) = x^{(n)}(a) - x^{(n-1)}(a)$ and $L = 2\rho^{2n}$. Then $M_0 = b^{\frac{n}{4}}$ and $M_2 = K^n$. Thus $|\frac{dx^{(n)}}{da}| < b^{\frac{n}{8}} K^n < b^{\frac{n}{9}}$. A similar estimate holds for $y^{(n)}$. □

Proof of Lemma 6.6: Let $z^n(a) = (x^n(a), y^n(a))$, $z^m(a) = (x^m(a), y^m(a))$, and let $y = \psi(x,a)$ be the $C^2(b)$-curve segment in ∂R_n containing both $z^n(a)$ and $z^m(a)$. Arguments similar to those used to prove $|\partial_{xa}\psi| < K^n$ can also be used to prove that the C^3-norm of ψ is $< K^n$.

Let P_n and Q_n be the numerator and denominator on the right hand side of (34), and similarly for P_m and Q_m. Then

$$\frac{dx^n}{da} - \frac{dx^m}{da} = \frac{P_n Q_m - P_m Q_n}{Q_n Q_m} = \frac{(P_n - P_m)Q_n + (Q_m - Q_n)P_n}{Q_n Q_m}.$$

As observed in the proof of Lemma 6.4, $|Q_m|, |Q_n| > K^{-1}$. $|Q_m|, |P_n| < K^n$. It remains therefore to estimate $|Q_m - Q_n|$ and $|P_m - P_n|$. Let q_n and q_m denote the slopes of e_n and e_m respectively. Fixing a and omitting it in the arguments of the functions below, we have

$$\begin{aligned} |Q_m - Q_n| \;&\le\; |\partial_x q_m(z^m) - \partial_x q_n(z^n)| \;+\; |\partial_y q_m(z^m)\cdot\partial_x\psi(z^m) - \partial_y q_n(z^n)\cdot\partial_x\psi(z^n)| \\ &\quad + |\partial_{xx}\psi(z^m) - \partial_{xx}\psi(z^n)|. \end{aligned}$$

The second difference, for example, is

$$\begin{aligned} \le\; & |\partial_y q_n(z^n)|\,|\partial_x\psi(z^m) - \partial_x\psi(z^n)| \;+\; |\partial_x\psi(z^m)|\,|\partial_y q_m(z^m) - \partial_y q_n(z^m)| \\ & + |\partial_x\psi(z^m)|\,|\partial_y q_n(z^m) - \partial_y q_n(z^n)|. \end{aligned}$$

This is $< (Kb)^{\frac{n}{4}}$ since $\|\psi\|_{C^3} < K^n$, $|\partial_{xx}q|, |\partial_x\partial_y q| < K$ (Corollary 2.2), $|q_m(z^n) - q_n(z^n)| < (Kb)^n$ (Lemma 2.1) and $|z^m - z^n| < (Kb)^{\frac{n}{4}}$ (Lemma 2.10). The other terms in $|Q_m - Q_n|$ and $|P_m - P_n|$ are estimated similarly. □

B.11 Dynamics of critical curves (Sect. 6.4)

Proof of Lemma 6.8: Let $\hat{z}_0$ be an arbitrary critical point. First we observe that as functions of a, $z_i(a)$ and $\hat{z}_0(a)$ move at very different speeds: $\|\frac{d}{da}z_i(a)\| \sim \|w_i(a)\| > e^{ci}$ by Proposition 6.1, whereas from Sect. 6.3 we have $\|\frac{d}{da}\hat{z}_0(a)\| < K$.

Next we consider $z_i(a) \in Q^{(k-1)}(a) \setminus Q^{(k)}(a)$ for some $k << i$, so that $\phi_a(z_i(a)) \in \partial Q^{(k-1)}(a)$, and study the relative movements of z_i, $\phi(z_i)$ and the relevant critical regions as a varies. For definiteness, let us assume that z_i is in the right component of $(Q^{(k-1)} \cap R_k) \setminus Q^{(k)}$ (which we call A), and that it moves left as a increases. (See Fig. 1 in Sect. 1.2.) In horizontal distance, it follows from the first paragraph that *relative to* $\phi(z_i)$, z_i is moving left at a speed $> K^{-1}e^{ci} - K$, which we assume to be $>> 1$. We do not have analytic estimates on the relative vertical movements of $\phi(z_i)$ and z_i, but note that since $z_i \notin \partial R_k$, it must enter A through its right vertical boundary and exit through the left. As z_i meets these vertical boundaries, it crosses them instantaneously due again to the horizontal speed differential between z_i and the critical points which determine these regions.

What we have shown is that the function $a \mapsto \phi_a(z_i(a))$ is continuous except at a discrete set of points corresponding to when $z_i(a)$ crosses a vertical boundary of some $Q^{(k)}$. If a and a' are the entry and exit parameters for $Q^{(k-1)} \setminus Q^{(k)}$ as above, we have $|a - a'| < \rho^{k-1}(K^{-1}e^{ci} - K)^{-1}$, and consequently $|\phi_a(z_i) - \phi_{a'}(z_i)| < K'\rho^{k-1}e^{-ci}$. As z_i crosses the vertical boundary into $Q^{(k)}$, a jump in $\phi_a(\cdot)$ occurs due to our rule for selecting binding points; this jump is $< b^{\frac{k-1}{4}}$.

As we continue to move toward the cricial set, either γ_i ends or we enter the "last" $Q^{(k)}$ available at step i, with $k \sim \theta i$. Let $\bar{a}$ be the parameter that corresponds to the end point of γ_i or where $d_{\mathcal{C}(\bar{a})}(z_i(\bar{a})) = e^{-\frac{\alpha i}{2}}$, whichever is reached first, and let $\bar{z} = \phi_{\bar{a}}(z_i(\bar{a}))$. We will use $\bar{z}$ as our "binding point" for γ_i. The "error" in this choice for $z_i(a)$, i.e. $|\,|z_i(a)-\bar z|_h - d_{\mathcal{C}(a)}(z_i(a))\,|$, is less than the total variation of $a \mapsto \phi_a(z_i(a))$ between a and $\bar a$. We have proved that this is $< Ke^{-ci}d_{\mathcal{C}(a)}(z_i(a)) + Kb^{\frac{k-1}{4}}$ where $\rho^k \sim d_{\mathcal{C}(a)}(z_i(a))$. □

Proof of Lemma 6.9: Let $\tilde p = \min\ \{p_a(z_i(a)) : z_i(a) \in I_{\mu j}\}$. Then by Corollary 4.2(a) and the last lemma, $\tilde p < K|\mu|$. Assertion (a) in Lemma 6.9 is obvious for $j \le \ell$ where ℓ is the common fold period. For $\ell < j \le \tilde p$, we have:

$$|z_{i+j}(a) - z_{i+j}(a')| \ \le\ \text{length}(\omega_j)\ =\ \int_{\omega_0} \frac{\|\tau_{i+j}\|}{\|\tau_i\|}\ \sim\ \int_{\omega_0} \frac{\|w_{i+j}\|}{\|w_i\|}.$$

Since z_i is outside of fold periods and w_i splits correctly, we have, for $j > \ell$, $\|w_{i+j}\|/\|w_i\| \sim e^{-\mu}\|w_j(z_i)\|$. Furthermore, if $\hat z_0 = \phi(z_i)$ and $\tilde p = p(z_i(\tilde a))$, then

$$e^{-\mu}\|w_j(z_i)\| \sim e^{-\mu}\|w_j(\hat z_0)\| \sim e^{-\mu}\|w_j(\hat z_0(\tilde a))\|,$$

the first $\sim$ coming from Lemma 4.9, and the second from the fact that $|\phi_a(z_i(a)) - \phi_{\tilde a}(z_i(\tilde a))| < e^{-ci} << K^j$ for $j < K|\mu| < K\alpha i$. Thus using the distance formula (9) in Lemma 4.11 for $T_{\tilde a}$, we have

$$\int_{\omega_0} \frac{\|\tau_{i+j}\|}{\|\tau_i\|}\ \sim\ e^{-2\mu}\frac{1}{\mu^2}\|w_j(\hat z_0(\tilde a))\|\ \sim\ \frac{1}{\mu^2}|z_{i+j}(\tilde a) - \hat z_j(\tilde a)|\ <\ \frac{1}{\mu^2}e^{-\beta j}.$$

This completes the proof of (a); (c) follows from $\|w_{\tilde p}(z_i(a))\| \sim \|w_{\tilde p}(z_i(\tilde a))\|$ and Proposition 6.1.

It remains to prove (b). From (a) we have that $z_{i+\tilde p}$ is out of all fold periods whenever $z_{i+\tilde p}(\tilde a)$ is. To show that the slopes of $\tau_{i+\tilde p}$ are $< K(\delta)$, we use Lemma 6.3: the $w_{i+\tilde p}$-vectors are b-horizontal, so it suffices to show that $\|w_s\| \le K\frac{1}{\delta}\|w_{i+\tilde p}\|$ for all $s < i+\tilde p$. For $s \ge i$, this is true by comparison with $a = \tilde a$; for $s < i$, $\|w_s\| \le \|w_i\|$ because z_i is a free return. Finally, the small slope of $\omega_{\tilde p}$ allows us to reverse the inequalities displayed above to conclude that $|\omega_{\tilde p}| \ge \frac{1}{\mu^2}e^{-\beta\tilde p}$. □

B.12 Distortion estimate for critical curves (Sect. 6.4)

Let J be a parameter interval satisfying all the assumptions made in Proposition 6.2, and let $a, a' \in J$. Assume that $z_i(a)$ and $z_i(a')$ are free returns, and that they lie in the same $I_{\mu j}$ with $\mu < \alpha i$. Write $\xi_0(a) = z_i(a)$ and $w_k(\xi_0(a)) = DT_a^k(\xi_0(a))\binom{0}{1}$. Let $\tilde p = p(z_i(\tilde a))$ be the bound period and $\hat z_0(\tilde a) = \phi(z_i(\tilde a))$ the binding point in the proof of Lemma 6.9. For $k < \tilde p$, let $\{w_k^*(\xi_0(a))\}$ be given by the splitting algorithm taken with respect to the orbit segment $\{\hat z_k(\tilde a)\}_{k=0}^{\tilde p}$, and write $w_k^*(\xi_0(a)) = M_k e^{i\theta_k(\xi_0(a))}$. The corresponding quantities for $\xi_0(a') = z_i(a')$ are defined analogously.

Sublemma B.6 *For $k < \tilde p$,*

$$\frac{M_k(\xi_0(a'))}{M_k(\xi_0(a))},\quad \frac{M_k(\xi_0(a))}{M_k(\xi_0(a'))} \le \exp\{K\sum_{j=1}^{k-1}\frac{\Delta_j(a,a')}{d_{\mathcal{C}}(\hat z_j(a))}\}$$

and

$$|\theta_k(\xi_0(a)) - \theta_k(\xi_0(a'))| < (Kb)^{\frac12}\Delta_{k-1}(a,a')$$

where

$$\Delta_j(a,a') = \sum_{s=0}^{j}(Kb)^{\frac{s}{4}}(|\xi_{j-s}(a) - \xi_{j-s}(a')| + |a-a'|).$$

Proof: The computation is similar to that in Appendix B.7, modulo the following adaptations to accommodate for the fact that different parameter values are involved in the present situation:

(i) Replace $|DT(\xi) - DT(\xi')| < K|\xi - \xi'|$ by

$$|DT_a(\xi(a)) - DT_{a'}(\xi(a'))| < K(|\xi(a) - \xi(a')| + |a - a'|).$$

(ii) Replace $|e - e'| < K|\xi - \xi'|$ by

$$|e(a) - e(a')| < K(|\xi(a) - \xi(a')| + |a - a'|).$$

(iii) Replace $|Y - Y'| < (Kb)^{\mu - j}|\xi - \xi'|$ by

$$|Y(a) - Y(a')| < (Kb)^{\mu - j}(|\xi(a) - \xi(a')| + |a - a'|).$$

□

Next we prove a version of Sublemma B.6 with $\binom{0}{1}$ replaced by $u_i(a) := \frac{w_i(z_0)(a)}{\|w_i(z_0)(a)\|}$.

Sublemma B.7

$$\frac{\|DT_a^{\tilde{p}}(\xi_0(a))u_i(a)\|}{\|DT_{a'}^{\tilde{p}}(\xi_0(a'))u_i(a')\|} < \exp\{K\frac{|\xi_0(a) - \xi_0(a')|}{e^{-\mu}}\}$$

Proof: The proof uses the fact that both $u_i(a)$ and $u_i(a')$ split correctly. Writing

$$u_i(a) = A(a)e(a) + B(a)\binom{0}{1},$$

we have

$$DT_a^{\tilde{p}}(\xi_0(a))u_i(a) = A(a)DT_a^{\tilde{p}}(\xi_0(a))e(a) + B(a)w_{\tilde{p}}(\xi_0(a)).$$

The proof is similar to that of Case 3 of Lemma 4.9, and Sublemma B.6 is used to compare $w_p(\xi_0(a))$ and $w_p(\xi_0(a'))$. □

Proof of Proposition 6.2: In view of Proposition 6.1, it suffices to show that there exists a constant $K > 0$ such that

$$\frac{1}{K} < \frac{|\ w_n(z_0(a))\ |}{|\ w_n(z_0(a'))\ |} < K.$$

Divide the time interval $(1, n)$ into bound and free period according to Lemma 6.9. As usual we denote free return times as t_k, $1 \le k < q$, and the bound period at t_k as p_{t_k}. Write

$$\log\frac{\|w_n(z_0(a))\|}{\|w_n(z_0(a'))\|} = \sum_{k<q} S'_k + \sum_{k<q} S''_k$$

where

$$S'_k = \log\frac{\|DT_a^{p_k}(z_{t_k}(a))u_{t_k}(a)\|}{\|DT_{a'}^{p_k}(z_{t_k}(a'))u_{t_k}(a')\|}, \qquad S''_k = \log\frac{\|DT_a^{t_{k+1}-p_k}(z_{t_k+p_k}(a))u_{t_k+p_k}(a)\|}{\|DT_{a'}^{t_{k+1}-p_k}(z_{t_k+p_k}(a'))u_{t_k+p_k}(a')\|}.$$

First we prove that $\sum_{k<q} S''_k < K$. Since $\gamma_j \cap \mathcal{C}^{(0)} = \emptyset$ for $t_k + p_k \le j \le t_{k+1}$, it is straightforward to see using Sublemma B.6 that

$$S''_k < \frac{K}{\delta}\sum_{j=t_k+p_k}^{t_{k+1}}(|z_j(a) - z_j(a')| + |a - a'|).$$

The effect of $|a-a'|$ can be ignored since $|a-a'| < e^{-cn}$. By Lemma 6.7, the slopes of γ_j are uniformly bounded and the length of γ_j grows exponentially, so

$$\sum_{j=t_k+p_k}^{t_{k+1}} |z_j(a)-z_j(a')| < K|\gamma_{t_{k+1}}|.$$

Again by Lemma 6.9(b), $|\gamma_{t_{k+1}}| > K|\gamma_{t_k}|$. Therefore $\sum_{k<q} S''_k < K$.

To estimate $\sum_{k<q} S'_k$ we apply Sublemma B.7. The effect of the term $|a-a'|$ can again be ignored, so that

$$\sum_{k<q} S'_k \leq K \sum_{k=1}^{q-1} \frac{\gamma_{t_k}}{e^{-\mu_k}}$$

where $\gamma_{t_k} \in I_{\mu_k j_k}$. To estimate this sum, let $m(\mu) = max\{t_k : \mu_k = \mu\}$ for each μ. Using the fact that $\mid \gamma_{t_{k+1}} \mid \geq K|\gamma_{t_k}|$, we conclude that

$$\sum_{k<q} S'_k < K \sum_{k<q} \frac{\mid \gamma_{t_k} \mid}{e^{-\mu_k}} < K \sum_{\mu} \frac{\mid \gamma_{m(\mu)} \mid}{e^{-\mu}} < K \sum_{\mu} \frac{1}{\mu^2}.$$

This completes the proof. □

References

[Bal] V. Baladi, *Dynamical zeta functions*, Real and Complex Dynamical Systems (B. Branner and P. Hjorth, ed.), Kluwer Academic Publisher, 1995a.

[Bar] M. Barge, *Prime end rotation numbers associated with the Hénon maps*, Continuum Theory and Dynamical Systems (T. West, ed.,), Lecture Notes in Pure and Applied Math. **149** (1993), Dekker, New York, 15-34.

[BC1] M. Benedicks and L. Carleson, *On iterations of* $1 - ax^2$ *on* $(-1, 1)$, Ann. Math. **122** (1985), 1-25.

[BC2] M. Benedicks and L. Carleson, *The dynamics of the Hénon map*, Ann. Math. **133** (1991), 73-169.

[BV] M. Benedicks and M. Viana, *Solutions of the basin problem for certain non-uniformly hyperbolic attractors*, in preparation.

[BY1] M. Benedicks and L.-S. Young, *Sinai-Bowen-Ruell measure for certain Hénon maps*, Invent. Math. **112** (1993), 541-576.

[BY2] M. Benedicks and L.-S. Young, *Markov extensions and decay of correlations for certain Hénon maps*, Asterisque (1999).

[Bi] G.D. Birkhoff, *Dynamical Systems*, Amer. Math. Soc., Providence, RI, 1927.

[Bo] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Springer (Lecture Notes in Math., Vol. **470**), Berlin, 1975.

[BSC1] L. A. Bunimovich, Ya. G. Sinai and N. I. Chernov, *Markov partitions for two-dimensional hyperbolic billiards*, Russ. Math. Survey **45** (1990), 105-152

[BSC2] L. A. Bunimovich, Ya. G. Sinai and N. I. Chernov, *Statistical properties of two-dimensional hyperbolic billiards*, Russ. Math. Survey **46** (1991), 47-106.

[CE] P. Collet and J.P. Eckmann, *Positive Liapunov exponents and absolute continuity for maps of the interval*, Ergodic Theory and Dynamical Systems **3** (1983), 13-46.

[CL] P. Collet and Y. Levy, *Ergodic properties of the Lozi mappings*, Comm. Math. Phys. **93** (1984), 461-481.

[C] P. Cvitanović, *Periodic orbits as the skeleton of classical and quantum chaos*, Physica D **51** (1991), 138-151.

[DRV] L.J. Diaz, J. Rocha and M. Viana, *Strange attractors in saddle-node cycles: prevalence and globality*, Invent. Math. **125** 37-74.

[G] J. Guckenheimer, *A strange strange attractor*, Hopf Bifurcation and its Application (J.E. Marsden and M. McCracken, ed.), Springer-Verlag: New York, Heidelberg, Berlin, 368-381.

[GH] J. Guckenheimer and P. Holmes, *Nonlinear oscillators, dynamical systems and bifurcations of vector fields*, Springer-Verlag, Appl. Math. Sciences **42** (1983).

[H] M. Hénon, *A two-dimensional mapping with a strange attractor*, Comm. Math. Phys. **50** (1976), 69-77.

[HY] H. Hu and L.-S. Young, *Nonexistence of SRB measures for some systems that are "almost Anosov"*, Ergodic Theory and Dynamical Systems **15** (1995), 67-76.

[I1] Y. Ishii, *Towards a kneading theory for Lozi mappings I: A solution of the pruning front conjecture and the first tangency problem*, Nonlinearity **10** (1997), 731-747.

[I2] Y. Ishii, *Towards a kneading theory for Lozi mappings II: Monotonicity of the topological entropy and Hausdorff dimension of attractors*, Commun. Math. Phys. **190** (1997), 375-394.

[J] M. Jakobson, *Absolutely continues invariant measures for one-parameter families of one-dimensional maps*, Comm. Math. Phys. **81** (1981), 39-88.

[Led] F. Ledrappier, *Propriétés ergodiques des mesures de Sinai*, Publ. Math., Inst. Hautes Etud. Sci. **59** (1984), 163-188.

[Lev] M. Levi, *Qualitative analysis of periodically forced relaxation oscillations*, Mem. AMS, **214** (1981), 1-147.

[K] A. Katok and B. Hasselblatt, *Introduction to the modern dynamical systems*, Cambridge University Press, Cambridge, 1995.

[Ma] R. Mañé, *Ergodic theory and differentiable dynamics*, Springer Verlag, 1987.

[dMvS] W. de Melo and S. van Strien, *One-dimensional Dynamics*, Springer-Verlag, (1993).

[MS] M. Misiurewicz and W. Szlenk, *Entropy of piecewise monotone mappings*, Studia Math., **67**, (1980), 45-67.

[M1] M. Misiurewicz, *Absolutely continues invariant measures for certain maps of an interval*, Publ. Math. IHES. **53** (1981), 17-51.

[M2] M. Misiurewicz, *The Lozi mapping has a strange attractor*, Nonlinear Dynamics, (R.H.G. Helleman, ed.), New York Academy of Sciences: New York, 1980, 348-358.

[Mil] J. Milnor, *On the concept of attractor*, Comm. Math. Phys., **99** (1985), 177-195.

[MT] J. Milnor and W. Thurston, *On iterated maps of the interval I and II*, 1977 preprint. Published in "Dynamical Systems: Proc. Univ. of Maryland 1986-87", (1988), Lect. Notes in Math., vol. 1342. Springer, Berlin New York, pp. 465-563.

[MV] L. Mora and M. Viana, *Abundance of strange attractors*, Acta. Math. **171** (1993), 1-71.

[NS] T. Nowicki and S. Van Strien, *Absolutely continuous invariant measures under a summability condition*. Invent. Math. **105** (1991), 123-136.

[P] Ja. B. Pesin, *Characteristic Lyapunov exponents and smooth ergodic theory*, Russ. math. Surv. **32.4** (1977), 55-114.

[PP] W. Parry and M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Société Mathématique de France (Astérisque, vol. **187-188**), Paris, 1990.

[PS] C. Pugh and M. Shub, *Ergodic attractors*, Trans. A. M. S. **312** (1989) 1-54.

[PT] J. Palis and F. Takens, *Hyperbolicity & sensitive chaotic dynamics at homoclinic bifurcations*, Cambridge studies in advanced mathematics, **35** Cambridge University Press, Cambridge, (1993)

[R1] D. Ruelle, *A measure associated with Axiom A attractors* Amer. J. Math. **98** (1976), 619-654.

[R2] D. Ruelle, *Thermodynamic Formalism*, Addison Wesley, Reading, MA, 1978.

[R3] D. Ruelle, *Dynamical Zeta functions for piecewise monotone maps of interval*, CRM Monograph Series, Vol. **4** Amer. Math. Soc., Providence, NJ, 1994.

[R4] D. Ruelle, *Ergodic theory of differentiable dynamical systems*. Publ. Math., Inst. Hautes Étud. Sci. **50** (1979), 27-58.

[Ro] C. Robinson, *Homoclinic bifurcation to a transitive attractor of Lorenz type*, Non-linearity **2** (1989), 495-518.

[Ry] M. Rychlik, *Lorenz attractors through Sil'nikov-type bifurcation, Part I*, Ergodic Theory and Dynamical Systems **10** (1990), 793-822.

[S1] Y. G. Sinai, *Dynamical systems with elastic reflections: Ergodic properties of dispersing billiards*, Russ. Math. Surveys **25** (1970), 137-189.

[S2] Y. G. Sinai, *Gibbs measure in ergodic theory*, Russian Math. Surveys **27** (1972), 21-69.

[Sm] S. Smale, *Differentiable dynamical systems*, Bull. Amer. Math. Soc. **73** (1967), 747-817.

[T] M. Tsujii, *A simple proof for monotonicity of entropy in the quadratic family*, Ergodic Theory and Dynamical Systems, to appear

[TTY] P. Thieullen, C. Tresser and L.-S. Young, *Positive exponent for generic 1-parameter families of unimodal maps*, C.R. Acad. Sci. Paris, t. 315 Série I (1992), 69-72; *J'Analyse* **64** (1994), 121-172.

[V] M. Viana, *Strange attractors in higher dimensions*, Bull. Braz. Math. Soc. **24** (1993), 13-62.

[W1] R. Williams, *Classification of one-dimensional attractors*, Global Analysis, Proc. Sym. in Pure Math., vol **14** (1970), 341-361.

[W2] R. Williams, *The structure of Lorenz attractors*, Turbulence Seminar Berkeley 1996/97, (P. Bernard and T. Ratiu, ed.), 1977, Springer-Verlag: New York, Heidelberg, Berlin, 1977, 94-112.

[Wa] P. Walters, *An introduction to ergodic theory*, Grad. Texts in Math., Springer-Verlag (1981).

[Y1] L.-S. Young, *A Bowen-Ruelle measure for certain piecewise hyperbolic maps*, Trans. A.M.S. **287** (1985), 41-48.

[Y2] L.-S. Young, *Ergodic theory of differentiable dynamical systems*, Real and Complex Dynamical Systems, (B. Branner and P. Hjorth, ed.), Kluwer Acad. Press, 1995.

[Y3] L.-S. Young, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. of Math. **147** (1998), 585-650.

[Y4] L.-S. Young, *Recurrence time and rate of mixing*, Israel J. of Math. (1999).