Toward a Theory of Rank One Attractors

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

January 2004/Revised March 2005

Contents

[1]Dept. of Math., University of Arizona, Tucson, AZ 85721, email dwang@math.arizona.edu. This research is partially supported by a grant from the NSF

[2]Courant Institute of Mathematical Sciences, 251 Mercer St., New York, NY 10012, email lsy@cims.nyu.edu. This research is partially supported by a grant from the NSF

Introduction

This paper is about a class of strange attractors that have the dual property of occurring naturally and being amenable to analysis. Roughly speaking, a *rank one* attractor is an attractor that has some instability in one direction and strong contraction in $m-1$ directions, m here being the dimension of the phase space.

The results of this paper can be summarized as follows. Among all maps with rank one attractors, we identify inductively subsets $\mathcal{G}_n,\ n = 1, 2, 3, \cdots$, consisting of maps that are "well-behaved" up to the nth iterate. The maps in $\mathcal{G} := \cap_{n>0}\mathcal{G}_n$ are then shown to be nonuniformly hyperbolic in a controlled way and to admit natural invariant measures called SRB measures. This is the content of Part II of this paper. The purpose of Part III is to establish existence and abundance. We show that for large classes of 1-parameter families $\{T_a\}$, $T_a \in \mathcal{G}$ for positive measure sets of a.

Leaving precise formulations to Section 1, we first put our results into perspective.

A. In relation to hyperbolic theory

Axiom A theory, together with its extension to the theory of systems with invariant cones and discontinuities, has served to elucidate a number of important examples such as geodesic flows and billiards (see e.g. [Sm],[A],[Si1],[B],[Si2],[W]). The invariant cones property is quite special, however. It is not enjoyed by general dynamical systems.

In the 1970s and 80s, an abstract nonuniform hyperbolic theory emerged. This theory is applicable to systems in which hyperbolicity is assumed only asymptotically in time and almost everywhere with respect to an invariant measure (see e.g. [O],[P],[R],[LY]). It is a very general theory with the potential for far-reaching consequences.

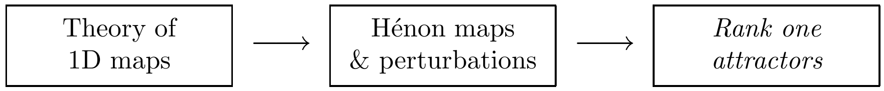
Yet using this abstract theory in concrete situations has proved to be difficult, in part because the assumptions on which this theory is based, such as the positivity of Lyapunov exponents or existence of SRB measures, are inherently difficult to verify. At the very least, the subject is in need of examples. To improve its utility, better techniques are needed to bridge the gap between theory and application. The project of which the present paper is a crucial component (see B and C below) is an attempt to address these needs.

We exhibit in this paper large numbers of nonuniformly hyperbolic attractors with controlled dynamics near every 1D map satisfying the well-known Misiurewicz condition. A detailed account of the mechanisms responsible for the hyperbolicity is given in Part II.

With a view toward applications, we sought to formulate conditions for the existence of SRB measures that are verifiable in concrete situations. These conditions cannot be placed on the map directly, for in the absence of invariant cones, to determine whether a map has this measure requires knowing it to infinite precision. We resolved this dilemma for the systems in question by identifying checkable conditions on 1-parameter families. These conditions guarantee the existence of SRB measures *with positive probability*, i.e. for positive measure sets of parameters. See Section 1.

B. In relation to one dimensional maps

In terms of techniques, this paper borrows heavily from the theory of iterated 1D maps, where much progress was made in the last 25 years. Among the works that have influenced us the most are [M],[J],[CE],[BC1] and [TTY]. The first breakthrough from 1D to a family of strongly dissipative 2D maps is due to Benedicks and Carleson, whose paper [BC2] is a *tour de force* analysis of the Hénon maps near the parameters $a = 2$, $b = 0$. Much of the local phase-space analysis in this paper is a generalization of their techniques, which in turn have their origins in 1D. Based on [BC2], SRB measures were constructed for the first time in [BY] for a (genuinely) nonuniformly hyperbolic attractor. The results in [BC2] were generalized in [MV] to small perturbations of the same maps. These papers form the core material referred to in the second box below.



All of the results in the second box depend on the formula of the Hénon maps. In going from the second to the third box, our aim is to take this mathematics to a more general setting, so that it can be leveraged in the analysis of attractors with similar characteristics (see below). Our treatment of the subject is necessarily more conceptual as we replace the equation of the Hénon maps by geometric conditions. A 2D version of these results was published in [WY1].

We believe the proper context for this set of ideas is m dimensions, $m \geq 2$, where we retain the rank one character of the attractor but allow the number of stable directions to be arbitrary. We explain an important difference between this general setup and 2D: For strongly contractive maps T with $T(X) \subset X$, by tracking $T^n(\partial X)$ for $n = 1, 2, 3, \cdots$, one can obtain a great deal of information on the attractor $\cap_{n \geq 0} T^n(X)$. This is because the area or volume of $T^n(X)$ decreases to zero very quickly. Since the boundary of a 2D domain consists of 1D curves, the study of planar attractors can be reduced to tracking a finite number of curves in the plane. This is what has been done in 2D, implicitly or explicitly. In $D > 2$, both the analysis and the geometry become more complex; one is forced to deal directly with higher dimensional objects. The proofs in this paper work in all dimensions including $D = 2$.

C. Further results and applications

We have a fairly complete dynamical description for the maps $T \in \mathcal{G}$ (see the beginning of this introduction), but in order to keep the length of the present paper reasonable, we have opted to publish these results separately. They include (1) a bound on the number of ergodic SRB measures, (2) conditions that imply ergodicity and mixing for SRB measures, (3) almost-everywhere behavior in the basin, (4) statistical properties of SRB measures such as correlation decay and CLT, and (5) coding of orbits on the attractor, growth of periodic points, etc. A 2D version of these results is published in [WY1]. Additional work is needed in higher dimensions due to the increased complexity in geometry.

We turn now to applications. First, by leveraging results of the type in this paper, we were able to recover and extend – by simply checking the conditions in Section 1 – previously known results on the Hénon maps and homoclinic bifurcations ([BC2],[MV],[V]).

The following new applications were found more recently: Forced oscillators are natural candidates for rank one attractors. We proved in [WY2],[WY3] that *any* limit cycle, when periodically kicked in a suitable way, can be turned into a strange attractor of the type studied here. It is also quite natural to associate systems with a single unstable direction with scenarios following a loss of stability. This is what led us to the result on the emergence of strange attractors from Hopf bifurcations in periodically kicked systems [WY3]. Finally, we mention some work in preparation in which we, together with K. Lu, bring some of the ideas discussed here including strange attractors and SRB measures to the arena of PDEs.

About this paper: This paper is self-contained, in part because relevant results from previously published works are inadequate for our purposes. The table of contents is self-explanatory. We have put all of the computational proofs in the Appendices so as not to obstruct the flow of ideas, and recommend that the reader omit some or all of the Appendices on first pass. This suggestion applies especially to Section 3, which, being a toolkit, is likely to acquire context only through subsequent sections. That having been said, we must emphasize also that the Appendices are an integral part of this paper; our proofs would not be complete without them.

1 Statement of Results

We begin by introducing $\mathcal{M}$, the class of one-dimensional maps of which all maps studied in this paper are perturbations. In the definition below, I denotes either a closed interval or a circle, $f : I \to I$ is a C^2 map, $C = \{f' = 0\}$ is the critical set of f, and C_δ is the δ-neighborhood of C in I. In the case of an interval, we assume $f(I) \subset \text{int}(I)$, the interior of I. For $x \in I$, we let $d(x, C) = \min_{\hat{x} \in C} |x - \hat{x}|$.

Definition 1.1 *We say $f \in \mathcal{M}$ if the following hold for some $\delta_0 > 0$:*
(a) Critical orbits: *for all $\hat{x} \in C$, $d(f^n(\hat{x}), C) > 2\delta_0$ for all $n > 0$.*
(b) Outside of C_{δ_0}: *there exist $\lambda_0 > 0, M_0 \in \mathbb{Z}^+$ and $0 < c_0 \leq 1$ such that*
(i) for all $n \geq M_0$, if $x, f(x), \cdots, f^{n-1}(x) \notin C_{\delta_0}$, then $|(f^n)'(x)| \geq e^{\lambda_0 n}$;
(ii) if $x, f(x), \cdots, f^{n-1}(x) \notin C_{\delta_0}$ and $f^n(x) \in C_{\delta_0}$, any n, then $|(f^n)'(x)| \geq c_0 e^{\lambda_0 n}$.
(c) Inside C_{δ_0}: *there exists $K_0 > 1$ such that for all $x \in C_{\delta_0}$,*
(i) $f''(x) \neq 0$;
(ii) $\exists p = p(x)$, $K_0^{-1} \log \frac{1}{d(x,C)} < p(x) < K_0 \log \frac{1}{d(x,C)}$, such that $f^j(x) \notin C_{\delta_0}$ $\forall j < p$ and $|(f^p)'(x)| \geq c_0^{-1} e^{\frac{1}{3}\lambda_0 p}$.

This definition may appear a little technical, but the properties are exactly those needed for our purposes. The class $\mathcal{M}$ is a slight generalization of the maps studied by Misiurewicz in [M].

Assume $f \in \mathcal{M}$ is a member of a one-parameter family $\{f_a\}$ with $f = f_{a^*}$. Certain orbits of f have natural *continuations* to a near a^*: For $\hat{x} \in C$, $\hat{x}(a)$ denotes the corresponding critical point of f_a. For $q \in I$ with $\inf_{n \geq 0} d(f^n(q), C) > 0$, $q(a)$ is the unique point near q whose symbolic itinerary under f_a is identical to that of q under f. For more detail, see Sects. 2.1 and 2.4.

Let $X = I \times D_{m-1}$ where I is as above and D_{m-1} is the closed unit disk in $\mathbb{R}^{m-1}$, $m \geq 2$. Points in X are denoted by (x, y) where $x \in I$ and $y = (y^1, \cdots, y^{m-1}) \in D_{m-1}$. To $F : X \to I$ we associate two maps, $F^\# : X \to X$ where $F^\#(x, y) = (F(x, y), 0)$ and $f : I \to I$ where $f(x) = F(x, 0)$. Let $\| \cdot \|_{C^r}$ denote the C^r norm of a map. A one-parameter family $F_a : X \to I$ (or $T_a : X \to X$) is said to be C^3 if the mapping $(x, y; a) \mapsto F_a(x, y)$ (respectively $(x, y; a) \mapsto T_a(x, y)$) is C^3.

Standing Hypotheses *We consider embeddings $T_a : X \to X$, $a \in [a_0, a_1]$, where $\|T_a - F_a^\#\|_{C^3}$ is small for some F_a satisfying the following conditions:*

(a) *There exists $a^* \in [a_0, a_1]$ such that $f_{a^*} \in \mathcal{M}$.*

(b) *For every $\hat{x} \in C = C(f_{a^*})$ and $q = f_{a^*}(\hat{x})$,*

$$\frac{d}{da} f_a(\hat{x}(a)) \neq \frac{d}{da} q(a) \text{ [3]} \qquad \text{at} \quad a = a^*. \tag{1}$$

(c) *For every $\hat{x} \in C$, there exists $j \leq m - 1$ such that*

$$\frac{\partial F(\hat{x}, 0; a^*)}{\partial y^j} \neq 0. \tag{2}$$

A T-invariant Borel probability measure ν is called an **SRB measure** if (i) T has a positive Lyapunov exponent ν-a.e.; (ii) the conditional measures of ν on unstable manifolds are absolutely continuous with respect to the Riemannian measures on these leaves.

[3] Here $q(a)$ is the continuation of $q(a^*)$ viewed as a point whose orbit is bounded away from C; it is not to be confused with $f_a(\hat{x}(a))$.

Theorem *In addition to the Standing Hypotheses above, we assume that* $\|T_a - F_a^\#\|_{C^3}$ *is sufficiently small depending on* $\{F_a\}$*. Then there is a positive measure set* $\Delta \subset [a_0, a_1]$ *such that for all* $a \in \Delta$, $T = T_a$ *admits an SRB measure.*

Notation For $z_0 \in X$, let $z_n = T^n(z_0)$, and let X_{z_0} be the tangent space at z_0. For $v_0 \in X_{z_0}$, let $v_n = DT^n_{z_0}(v_0)$. We identify X_z freely with $\mathbb{R}^m$, and work in $\mathbb{R}^m$ from time to time in local arguments. Distances between points in X are denoted by $|\cdot - \cdot|$, and norms on X_z by $|\cdot|$. The notation $\|\cdot\|$ is reserved for norms of maps (e.g. $\|T_a\|_{C^3}$ as above, $\|DT\| := \sup_{z\in X} \|DT_z\|$).

For definiteness, our proofs are given for the case $I = S^1$. Small modifications are needed to deal with the case where I is an interval. This is discussed in Sect. 3.9 at the end of Part I.

PART I PREPARATION

2 Relevant Results from One Dimension

The attractors studied in this paper have both an m-dimensional and a 1-dimensional character, the first having to do with how they are embedded in m-dimensional space, the second due the fact that the maps in question are perturbations of 1D maps. In this section, we present some results on 1D maps that are relevant for subsequent analysis. When specialized to the family $f_a(x) = 1 - ax^2$ with $a^* = 2$, the material in Sects. 2.2 and 2.3 is essentially contained in [BC2]; some of the ideas go back to [CE]. Part of Sect. 2.4 is a slight generalization of part of [TTY], which also contains an extension of [BC1] and the 1D part of [BC2] to unimodal maps.

2.1 More on maps in $\mathcal{M}$

The maps in $\mathcal{M}$ are among the simplest maps with nonuniform expansion. The phase space is divided into two regions: C_{δ_0} and $I \setminus C_{\delta_0}$. Condition (b) in Definition 1.1 says that on $I \setminus C_{\delta_0}$, f is essentially (uniformly) expanding. (c) says that every orbit from C_{δ_0}, though contracted initially, is not allowed to return to C_{δ_0} until it has regained some amount of exponential growth.

An important feature of $f \in \mathcal{M}$ is that its Lyapunov exponents outside of C_δ are bounded below by a strictly positive number independent of δ. Let δ_0, λ_0, M_0 and c_0 be as in Definition 1.1.

Lemma 2.1 *For* $f \in \mathcal{M}$, $\exists c_0' > 0$ *such that the following hold for all* $\delta < \delta_0$*:*
(a) if $x, f(x), \cdots, f^{n-1}(x) \notin C_\delta$*, then* $|(f^n)'(x)| \geq c_0' \delta e^{\frac{1}{3}\lambda_0 n}$*;*
(b) if $x, f(x), \cdots, f^{n-1}(x) \notin C_\delta$ *and* $f^n(x) \in C_{\delta_0}$*, any* n*, then* $|(f^n)'(x)| \geq c_0 e^{\frac{1}{3}\lambda_0 n}$*.*

Obviously, as we perturb f, its critical orbits will not remain bounded away from C. The expanding properties of f outside of C_δ, however, will persist in the manner to be described. Note the order in which ε and δ are chosen in the next lemma.

Lemma 2.2 *Let* f *and* c_0' *be as in Lemma 2.1, and fix an arbitrary* $\delta < \delta_0$*. Then there exists* $\varepsilon = \varepsilon(\delta) > 0$ *such that the following hold for all* g *with* $\|g - f\|_{C^2} < \varepsilon$*:*
(a) if $x, g(x), \cdots, g^{n-1}(x) \notin C_\delta$*, then* $|(g^n)'(x)| \geq \frac{1}{2} c_0' \delta e^{\frac{1}{4}\lambda_0 n}$*;*
(b) if $x, g(x), \cdots, g^{n-1}(x) \notin C_\delta$ *and* $g^n(x) \in C_{\delta_0}$*, any* n*, then* $|(g^n)'(x)| \geq \frac{1}{2} c_0 e^{\frac{1}{4}\lambda_0 n}$*.*

Lemmas 2.1 and 2.2 are proved in Appendix A.1

2.2 A larger class of 1D maps with good properties

We introduce next a class of maps more flexible than those in $\mathcal{M}$. These maps are located in small neighborhoods of $f_0 \in \mathcal{M}$. They will be our model of controlled dynamical behavior in higher dimensions.

For the rest of this subsection, we fix $f_0 \in \mathcal{M}$, and let δ_0, λ_0, M_0 and c_0 be as in Definition 1.1. The letter $K \geq 1$ is used as a generic constant that is allowed to depend only on f_0. (By "generic", we mean K may take on different values in different situations.) We fix also $\lambda < \frac{1}{5}\lambda_0$ and $\alpha << \min\{\lambda, 1\}$.

Let $\delta > 0$, and consider f with $\|f - f_0\|_{C^2} << \delta$. Let C be the critical set of f. We assume that for all $\hat{x} \in C$, the following hold for all $n > 0$:

(G1) $d(f^n(\hat{x}), C) > \min\{\delta, e^{-\alpha n}\}$; [4]

(G2) $|(f^n)'(f(\hat{x}))| \geq \hat{c}_1 e^{\lambda n}$ for some $\hat{c}_1 > 0$.

Proposition 2.1 *Let $\delta > 0$ be sufficiently small depending on f_0. Then there exists $\varepsilon = \varepsilon(f_0, \lambda, \alpha, \delta) > 0$ such that if $\|f - f_0\|_{C^2} < \varepsilon$ and f satisfies (G1) and (G2), then it has properties (P1)–(P3) below.*

(P1) **Outside of C_δ:** There exists $c_1 > 0$ such that the following hold:
(i) if $x, f(x), \cdots, f^{n-1}(x) \notin C_\delta$, then $|(f^n)'(x)| \geq c_1 \delta e^{\frac{1}{4}\lambda_0 n}$;
(ii) if $x, f(x), \cdots, f^{n-1}(x) \notin C_\delta$ and $f^n(x) \in C_{\delta_0}$, any n, then $|(f^n)'(x)| \geq c_1 e^{\frac{1}{4}\lambda_0 n}$.

For $\hat{x} \in C$, let $C_\delta(\hat{x}) = (\hat{x} - \delta, \hat{x} + \delta)$. We now introduce a partition $\mathcal{P}$ on I: For each $\hat{x} \in C$, $\mathcal{P}|_{C_\delta(\hat{x})} = \{I_{\mu j}^{\hat{x}}\}$ where $I_{\mu j}^{\hat{x}}$ are defined as follows: For $\mu \geq \log \frac{1}{\delta}$ (which we may assume is an integer), let $I_\mu^{\hat{x}} = (\hat{x} + e^{-(\mu+1)}, \hat{x} + e^{-\mu})$; for $\mu \leq \log \delta$, let $I_\mu^{\hat{x}}$ be the reflection of $I_{-\mu}^{\hat{x}}$ about $\hat{x}$. Each $I_\mu^{\hat{x}}$ is further subdivided into $\frac{1}{\mu^2}$ subintervals of equal length called $I_{\mu j}^{\hat{x}}$. We usually omit the superscript $\hat{x}$ in the notation above, with the understanding that $\hat{x}$ may vary from statement to statement. For example, "$x \in I_{\mu j}$ and $f^n(x) \in I_{\mu' j'}$" may refer to $x \in I_{\mu j}^{\hat{x}}$ and $f^n(x) \in I_{\mu' j'}^{\hat{x}'}$ for $\hat{x} \neq \hat{x}'$. The rest of I, i.e. $I \setminus C_\delta$, is partitioned into intervals of length $\approx \delta$.

(P2) **Partial derivative recovery for $x \in C_\delta(\hat{x})$:** For $x \in C_\delta$, let $p(x)$, the *bound period* of x, be the largest integer such that $|f^i x - f^i \hat{x}| \leq e^{-2\alpha i}$ $\forall j < p(x)$. Then
(i) $K^{-1} \log \frac{1}{|x - \hat{x}|} \leq p(x) \leq K \log \frac{1}{|x - \hat{x}|}$.
(ii) $|(f^{p(x)})'(x)| > e^{\frac{\lambda}{3} p(x)}$.
(iii) If $\omega = I_{\mu j}$, then $|f^{p(x)}(I_{\mu j})| > e^{-K\alpha|\mu|}$ for all $x \in \omega$.

The idea behind (P1) and (P2) is as follows: By choosing ε sufficiently small depending on δ, we are assured that there is a neighborhood $\mathcal{N}$ of f_0 such that all $f \in \mathcal{N}$ are essentially expanding outside of C_δ. Non-expanding behavior must, therefore, originate from *inside* C_δ. We hope to control that by imposing conditions (G1) and (G2) on C, and to pass these properties on to other orbits starting from C_δ via (P2).

(P2) leads to the following view of an orbit:

Returns to C_δ and ensuing bound periods: For $x \in I$ such that $f^i(x) \notin C$ for all $i \geq 0$, we define (free) return times $\{t_k\}$ and bound periods $\{p_k\}$ with

$$t_1 < t_1 + p_1 \leq t_2 < t_2 + p_2 \leq \cdots$$

as follows: t_1 is the smallest $j \geq 0$ such that $f^j(x) \in C_\delta$. For $k \geq 1$, p_k is the bound period of $f^{t_k}(x)$, and t_{k+1} is the smallest $j \geq t_k + p_k$ such that $f^j(x) \in C_\delta$. Note that an orbit may return to C_δ during its bound periods, i.e. t_i are not the only return times to C_δ.

[4] We will, in fact, assume f is sufficiently close to f_0 that $f^n(\hat{x}) \notin C_{\delta_0}$ for all n with $e^{-\alpha n} > \delta$.

The following notation is used: If $P \in \mathcal{P}$, then P^+ denotes the union of P and the two elements of $\mathcal{P}$ adjacent to it. For an interval $Q \subset I$ and $P \in \mathcal{P}$, we say $Q \approx P$ if $P \subset Q \subset P^+$. For practical purposes, P^+ containing boundary points of C_δ can be treated as "inside C_δ" or "outside C_δ".[5] For an interval $Q \subset I^+_{\mu j}$, we define the bound period of Q to be $p(Q) = \min_{x \in Q}\{p(x)\}$.

(P3) is about comparisons of derivatives for nearby orbits. For $x, y \in I$, let $[x, y]$ denote the segment connecting x and y. We say x and y have *the same itinerary* (with respect to $\mathcal{P}$) *through time* $n-1$ if there exist $t_1 < t_1 + p_1 \le t_2 < t_2 + p_2 \le \cdots \le n$ such that for every k, $f^{t_k}[x, y] \subset P^+$ for some $P \subset C_\delta$, $p_k = p(f^{t_k}[x, y])$, and for all $i \in [0, n) \setminus \cup_k [t_k, t_k + p_k)$, $f^{t_k}[x, y] \subset P^+$ for some $P \cap C_\delta = \emptyset$.

(P3) **Distortion estimate:** There exists K (independent of δ, x, y or n) such that if x and y have the same itinerary through time $n-1$, then

$$\left| \frac{(f^n)'(x)}{(f^n)'(y)} \right| \le K.$$

We remark that the partition of I_μ into $I_{\mu j}$-intervals is solely for purposes of this estimate. A proof of Proposition 2.1 is given in Appendix A.1.

2.3 Statistical properties of maps satisfying (P1)–(P3)

We assume in this subsection that f satisfies the assumptions of Proposition 2.1, so that in particular (P1)–(P3) hold. Let $\omega \subset I$ be an interval. For reasons to become clear later, we write $\gamma_i = f^i$, i.e. we consider $\gamma_i : \omega \to I$, $i = 0, 1, 2, \cdots$.

Lemma 2.3 *For $\omega \approx I_{\mu_0 j_0}$, let n be the largest j such that all $s \in \omega$ have the same itinerary up to time j. Then $n \le K|\mu_0|$.*

We call $n+1$ the *extended bound period* for ω. The next result, the proof of which we leave as an exercise, is used only in Lemma 8.2.

Lemma 2.4 *For $\omega \approx I_{\mu_0 j_0}$, there exists $n \le K|\mu_0|$ such that $\gamma_n(\omega) \supset C_\delta(\hat{x})$ for some $\hat{x} \in C$.*

The results in the rest of this subsection require that we track the evolution of γ_i to infinite time. To maintain control of distortion, it is necessary to divide ω into shorter intervals. The increasing sequence of partitions $\mathcal{Q}_0 < \mathcal{Q}_1 < \mathcal{Q}_2 < \cdots$ defined below is referred to as a *canonical subdivision by itinerary* for the interval ω: $\mathcal{Q}_0$ is equal to $\mathcal{P}|_\omega$ except that the end intervals are attached to their neighbors if they are strictly shorter than the elements of $\mathcal{P}$ containing them. We assume inductively that all $\hat{\omega} \in \mathcal{Q}_i$ are intervals and all points in $\hat{\omega}$ have the same itinerary through time i. To go from $\mathcal{Q}_i$ to $\mathcal{Q}_{i+1}$, we consider one $\hat{\omega} \in \mathcal{Q}_i$ at a time.

- If $\gamma_{i+1}(\hat{\omega})$ is in a bound period, then $\hat{\omega}$ is automatically put into $\mathcal{Q}_{i+1}$. (Observe that if $\gamma_{i+1}(\hat{\omega}) \cap C_\delta \neq \emptyset$, then $\gamma_{i+1}(\hat{\omega}) \subset I^+_{\mu' j'}$ for some μ', j', i.e. no cutting is needed during bound periods. This is an easy exercise.)
- If $\gamma_{i+1}(\hat{\omega})$ is not in a bound period, but all points in $\hat{\omega}$ have the same itinerary through time $i+1$, we again put $\hat{\omega} \in \mathcal{Q}_{i+1}$.
- If neither of the last two cases hold, then we partition $\hat{\omega}$ into segments $\{\hat{\omega}'\}$ that have the same itineraries through time $i+1$ *and* with $\gamma_{i+1}(\hat{\omega}') \approx P$ for some $P \in \mathcal{P}$. (If, for example, a segment appears that is strictly shorter than the $I_{\mu j}$ containing it, then it is attached to a neighboring segment.) The resulting partition is $\mathcal{Q}_{i+1}|_{\hat{\omega}}$.

[5]In particular, if $I_{\mu_0 j_0}$ is one of the outermost $I_{\mu j}$ in C_δ, then $I^+_{\mu_0 j_0}$ contains an interval of length δ just outside of C_δ.

For $s \in \omega$, let $\mathcal{Q}_i(s)$ be the element of $\mathcal{Q}_i$ to which s belongs. We consider the *stopping time* S on ω defined as follows: For $s \in \omega$, let $S(s)$ be the smallest i such that $\gamma_i(\mathcal{Q}_{i-1}(s))$ is not in a bound period and has length $> \delta$.

Lemma 2.5 *Assume δ is sufficiently small, and let $\omega \approx I_{\mu_0 j_0}$. Then*

$$|\{s \in \omega : S(s) > n\}| \ < \ e^{-\frac{1}{2}K^{-1}n} \ |\omega| \quad \text{for all} \quad n > K|\mu_0|.$$

Here K is the constant in the statement of Lemma 2.2.

Corollary 2.1 *There exists $\hat{K} > 0$ such that for any $\omega \subset I$ with $\delta < |\omega| < 3\delta$,*

$$|\{s \in \omega : S(s) > n\}| < e^{-\hat{K}^{-1}n}|\omega| \qquad \text{for} \quad n > \hat{K} \log \delta^{-1}.$$

For $\hat{\delta} < \delta$, $s \in \omega$ and $n \geq 0$, let $B_n(s)$ be the number of $i \leq n$ such that $\gamma_i(s)$ is in a bound period initiated from a visit to $C_{\hat{\delta}}$.

Proposition 2.2 *Given any $\sigma > 0$, there exists $\varepsilon_1 > 0$ such that for all $\hat{\delta} > 0$ sufficiently small, the following holds for all $\omega \approx I_{\mu_0 j_0}$:*

$$|\{s \in \omega : B_n > \sigma n\}| < \ e^{-\varepsilon_1 n} \ |\omega| \qquad \text{for all } n \geq \sigma^{-1} K \mu_0.$$

Proofs of all the results in this subsection are given in Appendix A.2 except that of Lemma 2.4, which is left to the reader as an exercise.

Remark The main use of Proposition 2.2 in this paper is in parameter estimates. When used in that context, it will be necessary for us to stop considering certain elements ω' of $\mathcal{Q}_i$ corresponding to deletions. Without going further into parameter considerations, we introduce the following notation. Let $*$ be the "garbage symbol". At step i, we may, in principle, choose to set $\gamma_i = *$ on any collection of elements of $\mathcal{Q}_i$. Once we set $\gamma_i|_{\omega'} = *$, it follows automatically that $\gamma_j|_{\omega'} = *$ for all $j \geq i$, i.e. we do not iterate ω' forward from time i on. We leave it as an (easy) exercise to verify that Proposition 2.2 remains valid in this slightly more general setting if we count only those i for which $\gamma_i(s) \neq *$ in the definition of $B_n(s)$.

2.4 Parameter transversality

We begin with a description of the structure of $f \in \mathcal{M}$ in terms of its symbolic dynamics. Let $\mathcal{J} = \{J_1, \cdots, J_q\}$ be the components of $I \setminus C$. For $x \in I$ such that $f^i x \notin C$ for all $i \geq 0$, let $\phi(x) = (\iota_i)_{i=0,1,\cdots}$ be given by $\iota_i = k$ if $f^i x \in J_k$.

Lemma 2.6 *For $f \in \mathcal{M}$, there exists an increasing sequence of compact sets $\Lambda^{(n)}$ with $\cup_n \Lambda^{(n)}$ dense in I such that the following hold:*
(a) $\Lambda^{(n)} \cap C = \emptyset$, $f(\Lambda^{(n)}) \subset \Lambda^{(n)}$, and $f|_{\Lambda^{(n)}}$ is conjugate to a shift of finite type;
(b) if $\inf_{i>0} d(f^i(x), C) > 0$, then $f(x) \in \Lambda^{(n)}$ for some n.

Our next result, which is a corollary of Lemmas 2.2 and 2.6, guarantees that continuations of the type in Standing Hypothesis (b) are well defined.

Corollary 2.2 *Let $f \in \mathcal{M}$, and let $q \in f(I)$ be such that $\delta_1 := \inf_{n \geq 0} d(f^n(q), C) > 0$. Then for all g with $\|g - f\|_{C^2} < \varepsilon$ where $\varepsilon = \varepsilon(\delta_1)$ is as in Lemma 2.2, there is a unique point $q_g \in I$ with $\phi_g(q_g) = \phi_f(q)$.*

Let $\{f_a\}$ be as in Section 1, with $f_{a^*} \in \mathcal{M}$. We fix $\hat{x} \in C(f_{a^*})$, and let $q = f_{a^*}(\hat{x})$. Let ω be an interval containing a^* on which $\hat{x}(a)$ and $q(a)$ (as given by Corollary 2.2) are well defined. We write $\hat{x}_k(a) = f_a^k(\hat{x}(a))$.

Proposition 2.3 *(i)* $a \mapsto q(a)$ *is differentiable;*
(ii) as $k \to \infty$,

$$Q_k(a^*) \ := \ \frac{\frac{d\hat{x}_k}{da}(a^*)}{(f_{a^*}^{k-1})'(\hat{x}_1(a^*))} \quad \to \quad \frac{d\hat{x}_1}{da}(a^*) - \frac{dq}{da}(a^*) = \sum_{i=0}^{\infty} \frac{\partial_a f_a(\hat{x}_i(a^*))|_{a=a^*}}{(f_{a^*}^i)'(\hat{x}_1(a^*))}.$$

A proof of this proposition, which is a slight adaptation of a result in [TTY], is given in Appendix A.3. Hypothesis (b) states that the expression on the right is nonzero. This condition, which can be viewed as a transversality condition for one-parameter families in the space of C^2 maps, is open and dense among the set of all 1-parameter families f_a passing through a given $f \in \mathcal{M}$. The proof in [TTY] is easily adapted to the present setting.

3 Tools for Analyzing Rank One Maps

This section is a toolkit for the analysis of maps $T : X \to X$ that are small perturbation of maps from X to $I \times \{0\}$. More conditions are assumed as needed, but detailed structures of the maps in question are largely unimportant. The purpose of this section is to develop basic techniques for use in the rest of the paper.

Notation The following rules on the use of constants are observed throughout:

- Two constants, $K_0 \geq 1$ and $0 < b << 1$, are used to bound the sizes of the objects being studied; they appear in assumptions.
- K is used as a generic constant; it appears in statements of results. In Sects. 3.1–3.4, K depends only on K_0 and m, the dimension of X; from Sect. 3.5 on, it depends on an additional object to be specified.
- b is assumed to be as small as need be; it is shrunk a finite number of times as we go along. Under no conditions is K allowed to depend on b.

For small angles, θ is often confused with $|\sin\theta|$.

3.1 Stability of most contracted directions

Most contracted directions on planes

Consider first $M \in \mathrm{L}(2,\mathbb{R})$ and assume $M \neq cO$ where O is orthogonal and $c \in \mathbb{R}$. Then there is a unit vector e, uniquely defined up to sign, that represents *the most contracted direction* of M, i.e. $|Me| \leq |Mu|$ for all unit vectors u. From standard linear algebra, we know $e^\perp$ is the most expanded direction, meaning $|Me^\perp| \geq |Mu|$ for all unit vectors u, and $Me \perp Me^\perp$. The numbers $|Me|$ and $|Me^\perp|$ are the *singular values* of M.

Next let $M \in \mathrm{L}(m,\mathbb{R})$ for $m \geq 2$, and let $S \subset \mathbb{R}^m$ be a 2D linear subspace. Then the ideas in the last paragraph clearly apply to $M|_S$, and we say $e = e(S)$ is a most contracted direction of M *restricted to* S if $|Me| \geq |Mu|$ for all unit vectors $u \in S$. We let f denote one of the two unit vectors in S orthogonal to e, i.e. f represents the most expanded direction in S, and $|Mf| = \|M|_S\|$, the norm of M restricted to S.

Two notions of stability for most contracted directions

For $M_1, M_2, \cdots \in \mathrm{L}(m,\mathbb{R})$, we let $M^{(i)}$ denote the composition $M_i \cdots M_2 M_1$.

(1) Let $S \subset \mathbb{R}^m$ be as above, and let $e_i(S)$ be the most contracted direction of $M^{(i)}|_S$ assuming that is well defined. It is known that if $M^{(i)}|_S$, $i = 1, 2, \cdots$, has two distinct Lyapunov exponents

as $i \to \infty$, then $e_i(S)$ converges to some $e_\infty(S)$ as $i \to \infty$. We are interested in the speed of this convergence.
(2) For parametrized families of linear maps $M_i(s)$ and plane fields $S(s)$ where $s = (s_1, \cdots, s_q)$ is a q-tuple of numbers, control of $\partial^k e_i$ and $\partial^k M^{(n)} e_i$ represents another form of stability for e_i. Here ∂^k denotes any one of the kth partial derivatives in s.

Main results

The ideas above are used to study the relation between pairs of vectors under the action of DT^n. To accommodate the many situations in which this analysis will be applied, we formulate our next lemma in terms of abstract linear maps. For motivation, the reader should think of M_i as $DT_{z_{i-1}}$ where $z_0 \in X$ and $T : X \to X$ is as in Sect. 1.1. For (H2), consider $z_0(s) \in X, S(s) \subset X_{z_0(s)}$, and $M_i(s) = DT_{z_{i-1}(s)}$.

(H1) Let $M_i = (\hat{M}_i^1, \cdots, \hat{M}_i^m) \in$L$(m, \mathbb{R})$, i.e. $\hat{M}_i^j : \mathbb{R}^m \to \mathbb{R}$. Then for all $i \geq 1$,
  (i) $\|\hat{M}_i^1\| < K_0$;
  (ii)$\|\hat{M}_i^j\| < b$ for $j = 2, \cdots, m$.

(H2) Let $u(s)$ and $v(s) \in \mathbb{R}^m$ be linearly independent, and let $S(s) = S(u(s), v(s))$ be the 2D subspace spanned by u and v. Let $M_i(s) \in$ L$(m, \mathbb{R})$. We assume the maps $s \mapsto u(s), v(s), M_i(s)$ are C^2 with
  (i) $\|u\|_{C^2}, \|v\|_{C^2} < K_0$;
  (ii) $\|\hat{M}_i^1\|_{C^2} < K_0^i$;
  (iii) $\|\hat{M}_i^j\|_{C^2} < K_0^i b$ for $j = 2, \cdots, m$.

Lemma 3.1 *(a) Let M_i be as in (H1), let $S \subset \mathbb{R}^m$ be an arbitrary 2D subspace, and let κ be such that $b^{\frac{1}{3}} < \kappa \leq 1$. If $\|M^{(i)}|_S\| > K_0^{-1}\kappa^{i-1}$ for all $1 \leq i \leq n$, then*

$$\begin{aligned} |e_{i+1}(S) - e_i(S)| &< (Kb\ \kappa^{-2})^i \quad \text{for } i < n; \\ |M^{(i)} e_n(S)| &< (Kb\ \kappa^{-2})^i \quad \text{for } i \leq n. \end{aligned}$$

(b) Let $M_i(s)$ and $S(s)$ be as in (H2), and $b^{\frac{1}{5}} \leq \kappa \leq 1$. If for $1 \leq i \leq n$, $\|M^{(i)}|_S\| > K_0^{-1}\kappa^{i-1}$ for all s, then for $k = 1, 2$,

$$\begin{aligned} |\partial^k e_1(S)| &< K; \\ |\partial^k (e_{i+1}(S) - e_i(S))| &< \Big(Kb\ \kappa^{-(2+k)}\Big)^i \quad \text{for } i < n; \\ |\partial^k M^{(i)} e_n(S)| &< \Big(Kb\ \kappa^{-(2+k)}\Big)^i \quad \text{for } i \leq n. \end{aligned}$$

A proof of Lemma 3.1 is given in Appendix A.5, after some preliminary material in Appendix A.4.

Assumptions for the rest of Section 3 *We consider $T : X \to X$ with the following properties: Let $T = (\hat{T}^1, \cdots, \hat{T}^m)$ be the coordinate maps of T. Then*
  (i) $\|\hat{T}^1\|_{C^3} < K_0$;
  (ii) $\|\hat{T}^j\|_{C^3} < b$ for $j = 2, \cdots, m$.

3.2 A perturbation lemma

The next lemma compares $w_n = DT^n_{z_0}(w_0)$ and $w'_n = DT^n_{z'_0}(w'_0)$ where z_i is near z'_i for $0 \le i < n$ and $w_0 \in X_{z_0}$ and $w'_0 \in X_{z'_0}$ are unit vectors such that $w_0 \approx w'_0$.

Lemma 3.2 *There exists K_1 depending on K_0 such that for κ and η satisfying $\kappa \le 1$ and $b^{\frac{1}{2}} < \eta < K_1^{-1}\kappa^8$, the following hold: Let (z_0, w_0) and (z'_0, w'_0) be such that $\angle(w_0, w'_0) < \eta^{\frac{1}{4}}$, $|w_i| > K_0^{-1}\kappa^{i-1}$ and $|z_i - z'_i| < \eta^{i+1}$ for $1 \le i < n$. Then*
(a) $|w'_n| > \frac{1}{2}K_0^{-1}\kappa^{n-1}$;
(b) $\angle(w_n, w'_n) < \eta^{\frac{n+1}{4}}$.

Lemma 3.2 is proved in Appendix A.6.

3.3 Temporary stable curves and manifolds

One dimensional strong stable curves – temporary or infinite-time – can be obtained by integrating vector fields of most contracted directions. In the proposition below, a neighborhood of 0 in X_{z_0} is identified with a neighborhood of z_0 in X, which in turn is identified with an open set of $\mathbb{R}^m$.

Proposition 3.1 *Let κ and η be as in Lemma 3.2, and let $z_0 \in X$ and $w_0 \in X_{z_0}$ be such that $|w_i| \ge K_0^{-1}\kappa^{i-1}|w_0|$ for $i = 1, \cdots, n$. Let S be a 2D plane in X containing z_0 and $z_0 + w_0$. For any $n \ge 1$, we view $e_n(S)$ as a vector field on S, defined where it makes sense, and let $\gamma_n = \gamma_n(z_0, S)$ be the integral curve to $e_n(S)$ with $\gamma_n(0) = z_0$. Then*
(a) γ_n is defined on $[-\eta, \eta]$ or until it runs out of X;
(b) for all $z \in \gamma_n$, $|T^i z_0 - T^i z| < (\frac{Kb}{\kappa^2})^i \eta$ for all $i \le n$.

Proposition 3.1 is proved in Appendix A.7.

We call γ_n a *temporary stable curve* or *stable curve of order n* through z_0. To obtain the full temporary stable manifold through z_0, we let S vary over all 2D planes containing z_0 and $z_0 + w_0$, obtaining

$$W^s_n(z_0) := \cup_S \ \gamma_n(z_0, S),$$

which we call a *temporary stable manifold of order n* through z_0. Observe that $W^s_n(z_0)$ is a C^1-embedded disk of co-dimension one. (The fact that $W^s_n(z_0)$ is C^1 away from z_0 follows from Lemma 3.1; at z_0 it has continuous partial derivatives.)

3.4 A curvature estimate

Let $\gamma_0 : [c_1, c_2] \to X$ 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 $k_i(s)$. Here $\gamma'_i(s)$ is the tangent vector to $\gamma_i(s)$.

Lemma 3.3 *Let $\kappa > b^{\frac{1}{3}}$, and let γ_0 be such that $k_0(s) \le 1$ for all s. Then the following hold for every $n > 0$: If*

$$|DT^j_{\gamma_{n-j}(s)}(\gamma'_{n-j}(s))| \ge \kappa^j |\gamma'_{n-j}(s)|$$

for every $j < n$, then

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

Lemma 3.3 is proved in Appendix A.8.

Additional assumptions for Sects. 3.5–3.8 *Let $\delta > 0$ be a small number.*

(1) The following is assumed about $\hat{T}^1 : X \to I$ and $f := \hat{T}^1|_{I\times\{0\}}$. Let $C = \{f' = 0\}$. Then
(i) outside of C_δ, f satisfies (P1) in Sect. 2.2;
(ii) inside C_δ, $|f''| > K_0^{-1}$;
(iii) for all $\hat{x} \in C$, there exists i such that $|\partial_{y^i}\hat{T}^1(x,0)| > K_0^{-1}$ for all $x \in C_\delta(\hat{x})$.

(2) From here on we restrict T to $R_1 := I \times \{|y| \le (m-1)^{\frac{1}{2}} b\}$. Note that $T(R_1) \subset R_1$ (see assumption (ii) at the end of Sect. 3.1).

From here on the generic constant K depends on the map $\hat{T}^1$ as well as K_0 and m. We introduce the following notation used in the rest of the paper:

- The first *critical region* $\mathcal{C}^{(1)}$ is defined to be

$$\mathcal{C}^{(1)} = \{(x,y) \in R_1 : |x - \hat{x}| < \delta,\ \hat{x} \in C(f)\}.$$

- $\mathbf{v} \in \mathbb{R}^m$ (identified with X_z, any z) is a fixed unit vector with zero x-component such that $|D\hat{T}^1_{(x,0)}\mathbf{v}| > K_0^{-1}$ for all $x \in C_\delta$. The existence of $\mathbf{v}$ is guaranteed by assumption (1)(iii) above. (We may take it to be orthogonal to the kernel of $D\hat{T}^1_{(\hat{x},0)}$ for $\hat{x} \in C$ but that is not necessary.) In general, $\mathbf{v}$ will be thought of as a reference vector in the "vertical" direction.

3.5 Dynamics outside of $\mathcal{C}^{(1)}$

For $u \in \mathbb{R}^m$, let (u_x, u_y) denote its x and y (or first and last $m-1$) components, and let $s(u) = \frac{|u_y|}{|u_x|}$. Curvature continues to be denoted by k.

Definition 3.1 *Assuming $|f'| > K_0^{-1}\delta$ outside of $\mathcal{C}^{(1)}$, we say $u \in \mathbb{R}^m$ is b-**horizontal** if $s(u) < \frac{3K_0}{\delta} b$. A curve γ in R_1 is called a $C^2(b)$-**curve** if $\gamma'(s)$ is b-horizontal and $k(s)$ is $< \frac{K_1 b}{\delta^3}$ for all s where K_1 is defined explicitly in the proof of Lemma 3.4.* [6]

Lemma 3.4 *(a) For $z \notin \mathcal{C}^{(1)}$, if $u \in X_z$ is b-horizontal, then so is $DT_z(u)$; in fact, $s(DT_z(u)) < \frac{3K_0}{2\delta} b$. Also, for $z \in \mathcal{C}^{(1)}$, $DT_z(\mathbf{v})$ is b-horizontal.*
(b) If γ is a $C^2(b)$-curve outside of $\mathcal{C}^{(1)}$, then $T(\gamma)$ is again a $C^2(b)$-curve.

Proof: The first assertion in (a) follows from the following *invariant cones condition*: Let u be such that $|u_x| = 1$ and $|u_y| < \frac{3K_0}{\delta} b$. Then

$$s(DT_z(u)) \ < \ \frac{b(1 + \frac{3K_0}{\delta} b)}{K_0^{-1}\delta - K_0 \frac{3K_0}{\delta} b} \ < \ \frac{3K_0}{2\delta} b$$

provided b is sufficiently small. For $z \in \mathcal{C}^{(1)}$, $s(DT_z(\mathbf{v})) < 2K_0 b$. For (b) we apply Lemma 3.3 to one iteration of T: Since T is a small perturbation of f, we have $|DTu| > \frac{1}{2} c_1 \delta$ where c_1 is as in (P1). This together with Lemma 3.3 gives $k < \frac{K_1}{\delta^3} b$ where $K_1 = 8 c_1^{-3} K$ and K is as in Lemma 3.3. $\square$

The next lemma says that outside of $\mathcal{C}^{(1)}$, iterates of b-horizontal vectors behave in a way very similar to that in 1D. Its proof is an easy adaption of the arguments in Sects. 2.1 and 2.2 made possible by part (a) of the last lemma.

---

[6] Quantities such as $\frac{K_1}{\delta^3} b$, $\frac{3K_0}{\delta} b$ appearing in this definition will be denoted as $\mathcal{O}(b)$.

Lemma 3.5 *There exists $c_2 > 0$ independent of δ such that the following hold: Let $z_0 \in R_1$ be such that $z_i \in R_1 \setminus \mathcal{C}^{(1)}$ for $i = 0, 1, \cdots, n-1$, and let $w_0 \in X_{z_0}$ be b-horizontal. Then*

(i) $|w_n| > c_2 \delta e^{\frac{1}{4}\lambda_0 n}|w_0|$;

(ii) if, in addition, $z_n \in \mathcal{C}^{(1)}$, then $|w_n| \geq c_2 e^{\frac{1}{4}\lambda_0 n}|w_0|$.

3.6 Properties of $e_1(S)$ for suitable S

We consider in this subsection e_1 of DT restricted to suitable choices of S.

Lemma 3.6 *For $z_0 \notin \mathcal{C}^{(1)}$, let $w \in X_{z_0}$ be b-horizontal, and let $S \subset X_{z_0}$ be any $2D$ plane containing w. Then $\angle(e_1(S), w) > K^{-1}\delta$.*

Proof: Assuming $|w| = 1$, write $e_1 = a_1 w + a_2 v$ where $v \in S$ is a unit vector $\perp$ w. Then $Kb > |DT(e_1)| = |a_1 DT(w) + a_2 DT(v)|$. Since $|DT(w)| > K^{-1}\delta$, it follows that $|a_2| > K^{-1}\delta$. □

Let γ be a $C^2(b)$ curve in $\mathcal{C}^{(1)}$ parametrized by arclength. At each point $\gamma(s)$, we let $S(s) = S(\gamma'(s), \mathbf{v})$. Let $u(s) = \gamma'(s)$, $v(s) = \frac{\mathbf{v} - \langle u, \mathbf{v}\rangle u}{|\mathbf{v} - \langle u, \mathbf{v}\rangle u|}$, i.e. $v(s)$ is a unit vector in $S(s)$ perpendicular to $u(s)$, and let $\eta(s) = \langle e_1(S(s)), v(s)\rangle$.

Lemma 3.7 *Let $\gamma(s)$, $S(s)$ and $\eta(s)$ be as above. Then $e_1(S(s))$ is well-defined on all of γ, and*

$$\left|\frac{d\eta(s)}{ds}\right| > K_1^{-1} \tag{3}$$

for some K_1 independent of γ.

Lemma 3.7 is a direct consequence of our assumptions that $f''(\hat{x}) \neq 0$ and $\partial_{y^i}\hat{T}^1_{(\hat{x},0)} \neq 0$ for $\hat{x} \in C$. A proof is given in Appendix A.9.

3.7 Critical points on $C^2(b)$ curves in $\mathcal{C}^{(1)}$

We fix $\hat{K}_0 > 10K_0$ where K_0 satisfies $|D\hat{T}^1_{(x,0)}\mathbf{v}| > K_0^{-1}$.

Definition 3.2 *Let γ be a $C^2(b)$-curve in $\mathcal{C}^{(1)}$. We say that z_0 is a* **critical point of order** n **on** γ *if*

(a) $|DT^i_{z_0}(\mathbf{v})| \geq \hat{K}_0^{-1}$ for $i = 1, 2, \cdots, n$;

(b) at z_0, $\angle(e_n(S), \gamma') = 0$ with $S = S(\gamma', \mathbf{v})$.

Corollary 3.1 (Corollary to Lemma 3.7) *On any $C^2(b)$-curve traversing the full length of a component of $\mathcal{C}^{(1)}$, there exists a unique critical point of order 1.*

We now turn to the problem of *inducing* new critical points on nearby curves starting from a known critical point on a $C^2(b)$-curve. We begin with two lemmas the exact form of which will be used.

Lemma 3.8 *Let γ and $\hat{\gamma}$ be $C^2(b)$-curves parametrized by arclength in $\mathcal{C}^{(1)}$. Assume*

(a) $\gamma(0)$ is a critical point of order n on γ with $|DT^i_{\gamma(0)}(\mathbf{v})| \geq 2\hat{K}_0^{-1}$ for $i \leq n$;

(b) $|\gamma(0) - \hat{\gamma}(0)|$, $|\gamma'(0) - \hat{\gamma}'(0)| < b^{\frac{n}{4}}$; and

(c) $\hat{\gamma}(s)$ is defined for all $s \in [-b^{\frac{n}{5}}, b^{\frac{n}{5}}]$.

Then there exists a unique s, $|s| < Kb^{\frac{n}{4}}$, such that $\hat{\gamma}(s)$ is a critical point on $\hat{\gamma}$.

Lemma 3.9 *There exists K_2 for which the following holds: Let γ be a $C^2(b)$-curve parametrized by arclength in $\mathcal{C}^{(1)}$, and let $z = \gamma(0)$ be a critical point of order n. If*

(a) $|DT_z^i(\mathbf{v})| \geq 2\hat{K}_0^{-1}$ for $i = 1, 2, \cdots, n+m$, and

(b) $\gamma(s)$ is defined for $s \in [-K_2(Kb)^n, K_2(Kb)^n]$,

then there exists a unique critical point $\hat{z}$ of order $n+m$ on γ, and $|\hat{z} - z| < K_2(Kb)^n$.

Proofs of Corollary 3.1 and Lemmas 3.8 and 3.9 are given in Appendix A.10.

3.8 Tracking $w_n = DT_{z_0}^n(w_0)$: a splitting algorithm

Let $z_0 \in R_1$, and let $w_0 \in X_{z_0}$ be a b-horizontal unit vector. In the case where $z_i \notin \mathcal{C}^{(1)}$ for all i, the resemblance to 1D dynamics is made clear in Lemmas 3.4 and 3.5. Consider next an orbit $z_0, z_1, \cdots$ that visits $\mathcal{C}^{(1)}$ exactly once, say at time $t > 0$. Assume:

(i) There exists $\ell > 1$ such that $|DT_{z_t}^i(\mathbf{v})| \geq K_0^{-1}$ for all $i \leq \ell$, so that in particular $e_\ell(S)$ is defined at z_t with $S = S(\mathbf{v}, w_t)$.

(ii) $\angle(w_t, e_\ell(S)) \geq b^{\frac{\ell}{2}}$.

Then $DT_{z_0}^i(w_0)$ can be analyzed as follows. We split w_t into $w_t = \hat{w}_t + \hat{E}$ where $\hat{w}_t$ is a scalar multiple of $\mathbf{v}$ and $\hat{E}$ is a scalar multiple of $e_\ell(S)$. For $i \leq t$ and $i \geq t+\ell$, let $w_i^* = w_i$. For i with $t < i < t+\ell$, let $w_i^* = DT_{z_t}^{i-t}(\hat{w}_t)$. We claim that all the w_i^* are b-horizontal vectors, and that $\{|w_{i+1}^*|/|w_i^*|\}_{i=0,1,2,\cdots}$ resembles a sequence of 1D derivatives, with $|w_{t+1}^*|/|w_t^*|$ simulating a drop in the derivative when an orbit comes near a critical point in 1D.

In light of Lemma 3.4, to show that w_i^* is b-horizontal, it suffices to consider $w_{t+\ell}^*$. Observe from assumption (ii) above that $|\hat{w}_t| > b^{\frac{\ell}{2}}|\hat{E}|$. (Note that e_ℓ is close to e_1 from Lemma 3.1, and $s(e_1) < K\delta$ for $z \in \mathcal{C}^{(1)}$.) This together with assumption (i) implies that

$$|DT_{z_t}^\ell(\hat{E})| \leq (Kb)^\ell |\hat{E}| \leq K^\ell b^{\frac{\ell}{2}} |\hat{w}_t| \leq K_0 K^\ell b^{\frac{\ell}{2}} |DT_{z_t}^\ell(\hat{w}_t)|.$$

Since $s(DT_{z_t}^\ell(\hat{w}_t)) < \frac{3K_0}{2\delta} b$ (see Lemma 3.4), $w_{t+\ell}^* = DT_{z_t}^\ell(\hat{w}_t) + DT_{z_t}^\ell(\hat{E})$ is b-horizontal.

The discussion above motivates the following

Splitting algorithm We give this algorithm only for $z_0 \in \mathcal{C}^{(1)}$ and $w_0 = \mathbf{v}$ since this is mostly how it will be used. Let $t_1 < t_2 < \cdots$ be the times > 0 when $z_i \in \mathcal{C}^{(1)}$. For each t_j, fix $\ell_{t_j} \geq 2$ with the property that $|DT_{z_{t_j}}^i(\mathbf{v})| > K_0^{-1}$ for $i = 1, \cdots, \ell_{t_j}$ (such ℓ_{t_j} always exist). The following algorithm generates two sequences of vectors w_i^* and $\hat{w}_i$:

1. For $0 \leq i < t_1$, let $w_i^* = \hat{w}_i = w_i$.

2. At $i = t_1$, set $w_i^* = w_i$, and define $\hat{w}_i$ as follows: If w_i^* is a scalar multiple of $\mathbf{v}$, let $\hat{w}_i = w_i^*$. If not, let $S = S(w_i^*, \mathbf{v})$. Then split w_i^* into

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

where $\hat{w}_i$ is a scalar multiple of $\mathbf{v}$ and $\hat{E}_i$ is a scalar times $e_{\ell_i}(S)$.

3. For $i > t_1$, we let

$$w_i^* = DT_{z_{i-1}}(\hat{w}_{i-1}) \ + \sum_{j:\ t_j + \ell_{t_j} = i} DT_{z_{t_j}}^{\ell_{t_j}}(\hat{E}_{t_j}), \tag{4}$$

and define $\hat{w}_i$ as follows: if $i = t_j$, split w_i^* into $w_i^* = \hat{w}_i + \hat{E}_i$ as in item 2; if $i \neq t_j$ for any j, set $\hat{w}_i = w_i^*$.

This algorithm is of interest when the contributions from the $\hat{E}_i$-terms as they rejoin w_i^* are negligible; the meaning of w_i^* and $\hat{w}_i$ are unclear otherwise. The next lemma contains a set of technical conditions describing a "good" situation:

Lemma 3.10 *Let z_0, ℓ_{t_j}, w_i and w_i^* be as above, and let $I_j := [t_j, t_j + \ell_{t_j})$. Assume*

(a) for each $i = t_j$, $|\hat{w}_i| > b^{\frac{\ell_i}{2}} |\hat{E}_i|$;

(b) the I_j are nested, i.e. for $j < j'$, either $I_j \cap I_{j'} = \emptyset$ or $I_{j'} \subset I_j$.

Then the w_i^* are b-horizontal.

A proof of Lemma 3.10 is given in Appendix A.11.

3.9 Attractors arising from interval maps

We explain how to deal with the endpoints of I in the case where I is an interval.

Let $f \in \mathcal{M}$. By assumption, $f(I) \subset \text{int}(I)$. We let $\Lambda = \Lambda^{(n)}$ be as in Lemma 2.3 where n is large enough that $f(I)$ is well inside $[x_1, x_2]$, the shortest interval containing Λ. It is a standard fact that periodic points are dense in topologically transitive shifts of finite type. From this one deduces easily that pre-periodic points are dense in all shifts of finite type, transitive or not. Let y_1 and y_2 be pre-periodic points so that $f(I)$ is well inside $[y_1, y_2]$. For $i = 1, 2$, let k_i and n_i be such that $f^{k_i+n_i}(y_i) = f^{n_i}(y_i)$. Our plan is to prove the following for T when b is sufficiently small:

(i) Near $(f^{k_i}(y_i), 0)$, $i = 1, 2$, T has a periodic point z_i.

(ii) z_i is hyperbolic; it therefore has a codimension one stable manifold $W^s(z_i)$. We claim that W_i, the connected component of $W^s(z_i)$ containing z_i, spans R_1 in the sense that it is the graph of a function from $\{|y| \le (m-1)^{\frac{1}{2}} b\}$ to I.

(iii) Near $(y_i, 0)$ there is a connected component V_i of $W^s(z_i)$; V_i also spans R_1.

(iv) If $\hat{R}_1$ is the part of R_1 between V_1 and V_2, then $T(\hat{R}_1) \subset \hat{R}_1$.

The existence and hyperbolicity of z_i follows from the fact that $|(f^{k_i})'(f^{n_i} y_i)| > 1$ (Lemma 2.1). That W_i spans the cross-section of R_1 follows from Lemma 3.1 and the construction in Sect. 3.3 with $n \to \infty$. Moving on to (iii), the existence of a component of $T^{-k_i} W_i$ near $(y_i, 0)$ follows by continuity. Repeating the arguments at z_i on a (any) point in V_i, we see that not only does V_i span R_1 but its tangent vectors make angles $> K^{-1}\delta$ with the x-axis. Thus the diameter of V_i is arbitrarily small as $b \to 0$, and (iv) follows from $f(I) \subset (y_1, y_2)$.

In Part II, we restrict the domain of T to $\hat{R}_1$. The two ends of $\hat{R}_1$, namely $V_1 \cup V_2$, are asymptotic to the periodic orbits of z_1 and z_2. In particular, they stay away from $\mathcal{C}^{(1)}$. This part of $\partial \hat{R}_1$ is not visible in local arguments. In Sections 7 and 8, in the treatment of monotone branches, there will be some special branches that end in $T^j(V_i)$. Modifications in the arguments are straightforward.

In Part III, we take $z_i(a)$ to be continuations of the same periodic orbits, so that $\hat{R}_1(a)$ varies continuously with a.

Notation for the rest of the paper

- We assume $T = (\hat{T}^1, \cdots, \hat{T}^m) : X \to X$ is such that $\|\hat{T}^j\|_{C^3} < b$ for $j = 2, \cdots, m$.
- $R_1 := I \times \{y \in \mathbb{R}^{m-1} : \ |y| < (m-1)^{\frac{1}{2}} b\}$; $R_k := T^{k-1} R_1$ for $k = 2, 3, \cdots$.
- For definiteness, we let $\mathcal{F}_1$ be the foliation on R_1 given by $\{y = \text{constant}\}$ (this can be replaced by any foliation whose leaves are $C^2(b)$ curves); for $k > 1$, $\mathcal{F}_k := T_*^{k-1}(\mathcal{F}_1)$, i.e. the leaves of $\mathcal{F}_k$ are the T^{k-1}-images of those of $\mathcal{F}_1$.
- A subset $H \subset R_j$ is called a **section** of R_j if it is the diffeomorphic image of $\Phi : [-1, 1] \times D_{m-1} \to R_j$ with $\Phi^{-1}(\partial R_j) = [-1, 1] \times \partial D_{m-1}$. A section H of R_j is called **horizontal** if each component of $\Phi(\{\pm 1\} \times D_{m-1})$ is contained in a hyperplane $\{x = \text{const}\}$ and

all the leaves of $\mathcal{F}_j|_H$ are $C^2(b)$-curves. The **cross-sectional diameter** of a horizontal section H is defined to be the supremum of diam$(V \cap H)$ as V varies over all hyperplanes perpendicular to S^1.

- The distance from z to z' in R_1 is denoted by $|z-z'|$, and their **horizontal distance**, i.e. difference in x-coordinates, is denoted by $|z-z'|_h$.

PART II PHASE-SPACE DYNAMICS

The goal of Part II is to identify, among all maps $T : X \to X$ that are near small perturbations of 1D maps, a class $\mathcal{G}$ with certain desirable features. To explain what we have in mind, consider the situation in 1D. In Sect. 2.2, we show that for maps sufficiently near $f_0 \in \mathcal{M}$, two relatively simple conditions, (G1) and (G2), imply dynamical properties (P1)–(P3), which in turn lead to other desirable characteristics. Our class $\mathcal{G}$ will be modelled after these maps.

The first major hurdle we encounter as we attempt to formulate higher dimensional analogs of (G1) and (G2) is the absence of a well defined *critical set*. As we will show, the concept of a critical set can be defined, but only inductively and only for certain maps. This implies that our "good maps" can only be identified inductively. The task before us, therefore, is the inductive construction of $\mathcal{G}_n$, $n = 1, 2, \cdots$, consisting of maps that are "good" in their first n iterates, and $\mathcal{G}$ is taken to be $\cap_{n \geq 0} \mathcal{G}_n$.

We do not claim in Part II that $\mathcal{G}$ is nonempty, and we consider one map at a time to determine if it is in $\mathcal{G}$; no parameters are involved. The existence (and abundance) of maps in $\mathcal{G}$ is proved in Part III.

Organization Sections 4–9, which comprise Part II, are organized as follows:

Sect. 4.1 contains five statements describing 5 aspects of dynamical behavior. Together, these statements give a snapshot of the maps in $\mathcal{G}_n$ for certain n. The rest of Section 4 is devoted to the elucidation of the ideas introduced.

Implications of these ideas are developed in Section 5, and a formal inductive construction of $\mathcal{G}_n$ for $n \leq N_0 \sim (\log \frac{1}{b})^2$ is given in Section 6.

After N_0 iterates, a fundamental, qualitative change in geometry occurs. The new complexities that arise are dealt with in Sections 7 and 8.

The existence of SRB measures for $T \in \mathcal{G}$ is proved in Section 9.

The notation is as in Section 1, namely that $f : S^1 \to S^1$, $F : R_1 \to S^1$ and $F^\# : R_1 \to R_1$ are related by $F(x, 0) = f(x)$ and $F^\#(x, y) = (F(x, y), 0)$, and $T : R_1 \to R_1$ is a C^3 embedding.

Standing hypotheses *Throughout Part II, we fix $f_0 \in \mathcal{M}$ and $K_0 > 1$, and consider*

- $f : S^1 \to S^1$ *with* $\|f - f_0\|_{C^2} < a$,
- $F : R_1 \to S^1$ *with* $\|F\|_{C^3} < K_0$ *and* $|DF_{(\hat{x},0)}(\mathbf{v})| > K_0^{-1}$ *for* $\hat{x} \in C(f_0)$*, and*
- $T : R_1 \to R_1$ *with* $\|T - F^\#\|_{C^3} < b$

where $a, b > 0$ are as small as need be. The letter K is used as a generic constant which, in Part II, is allowed to depend only on f_0, K_0 and our choice of λ.

4 Critical Structure and Orbits

4.1 Formal assumptions

We describe in this subsection several aspects of geometric and dynamical behaviors to be viewed as desirable. These assumptions, labelled **(A1)–(A5)**, will eventually be part of the inductive cycle up to a certain time. For the moment they are only formal statements.

For purposes of the present discussion, $\lambda > 0$ can be any number $< \frac{1}{5}\lambda_0$ (see Sect. 2.2). We choose α so that $b << \alpha << \min(\lambda, 1)$, and let $\alpha^* = \frac{6}{\lambda}\alpha$. Let $\theta = \frac{K}{\log \frac{1}{b}}$ where K is chosen so that $b^\theta < \|DT\|^{-20}$. Let N be a positive integer $>> 1$. For simplicity of notation, we assume $\theta N, \theta^{-1}, \frac{1}{\alpha^*} \in \mathbb{Z}^+$ (otherwise write $[\theta N], [\theta^{-1}], [\frac{1}{\alpha^*}]$).

(A1) Geometry of critical regions *There are sets* $\mathcal{C}^{(1)} \supset \mathcal{C}^{(2)} \supset \cdots \supset \mathcal{C}^{(\theta N)}$ *called* critical regions *with the following properties:*

(i) $\mathcal{C}^{(1)}$ *is as introduced in Sect. 3.4. For* $1 < k \leq \theta N$, $\mathcal{C}^{(k)}$ *is the union of a finite number of connected components* $\{Q^{(k)}\}$ *each one of which is a horizontal section of* R_k *of length* $\min(2\delta, 2e^{-\lambda k})$ *and cross-sectional diameter* $< b^{\frac{k}{2}}$.

(ii) $\mathcal{C}^{(k)}$ *is related to* $\mathcal{C}^{(k-1)}$ *as follows: For each* $Q^{(k-1)}$*, either* $R_k \cap Q^{(k-1)} = \emptyset$ *or it meets* $Q^{(k-1)}$ *in a finite number of horizontal sections* $\{H\}$ *each one of which extends* $> \frac{1}{2}e^{-\alpha k}$ *beyond the two ends of* $Q^{(k-1)}$*. Each* $H \cap Q^{(k-1)}$ *contains exactly one component of* $\mathcal{C}^{(k)}$ *located roughly in the middle. (See Fig. 1.)*

(iii) Inside each $Q^{(k)}$*, a point* $z_0 = z_0^*(Q^{(k)})$ *whose x-coordinate is exactly half-way between those of the two ends of* $Q^{(k)}$ *is singled out;* z_0 *is a critical point of order* k *in the sense of Definition 3.2 with respect to the leaf of the foliation* $\mathcal{F}_k$ *containing it.*

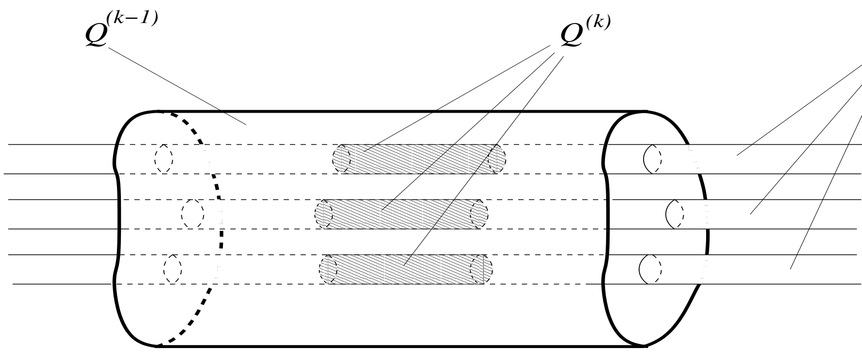


Fig. 1 Structure of critical regions

We call $z_0^*(Q^{(k)})$ a **critical point of generation** k, and let Γ_k denote the set of all critical points of generation $\leq k$. Let $Q^{(k)}(z_0)$ denote the component of $\mathcal{C}^{(k)}$ containing z_0.

The next three assumptions prescribe certain behaviors on the orbits of $z_0 \in \Gamma_{\theta N}$. To state them, we need the following definitions:

First, we define a notion of **distance to critical set** for z_i, denoted $d_{\mathcal{C}}(z_i)$. If $z_i \notin \mathcal{C}^{(1)}$, let $d_{\mathcal{C}}(z_i) = \delta + d(z_i, \mathcal{C}^{(1)})$. If $z_i \in \mathcal{C}^{(1)}$, we let $d_{\mathcal{C}}(z_i) = |z_i - \phi(z_i)|$ where $\phi(z_i)$ is defined as follows. Let j be the largest integer $\leq \alpha^*\theta i$ with the property that $z_i \in \mathcal{C}^{(j)}$. Then $\phi(z_i) := z_0^*(Q^{(j)}(z_i))$ is called the **guiding critical point** for z_i. As the name suggests, the orbit of $\phi(z_i)$ will be thought of as guiding that of z_i through its derivative recovery. Suppose $z_i \in \mathcal{C}^{(1)}$ and $\phi(z_i)$ is of generation j. We say $w \in X_{z_i}$ is **correctly aligned**, or correctly aligned with respect to the leaves of the $\mathcal{F}_j$-foliation, if $\angle(\tau_j(z_i), w) << K_1^{-1} d_{\mathcal{C}}(z_i)$ where K_1^{-1} is a lower bound on $|\frac{d}{ds}e_1|$ along $C^2(b)$-curves in $\mathcal{C}^{(1)}$ in the sense of Lemma 3.7 and $\tau_j(z_i)$ is tangent to the leaf of $\mathcal{F}_j$ through z_i. We say w is correctly aligned with ε-error if $\varepsilon << K_1^{-1}$ and $\angle(\tau_j(z_i), w) < \varepsilon d_{\mathcal{C}}(z_i)$.

For $z_0 \in \Gamma_{\theta N}$, we let $w_0 = \mathbf{v}$, and for a chosen family of ℓ_i corresponding to $z_i \in \mathcal{C}^{(1)}$, let w_i^*, $i = 0, 1, 2, \cdots$, be given by the splitting algorithm in Sect. 3.8. The numbers $\{\ell_i\}$ are called the **splitting periods** for z_0. Let $\varepsilon_0 << K_1^{-1}$ be fixed. We shrink δ if necessary so that it is $<< \varepsilon_0$.

(A2)–(A4) Properties of critical orbits *For $z_0 \in \Gamma_{\theta N}$ of generation k, the following hold for all $i \leq k\theta^{-1}$:*

(A2) $d_{\mathcal{C}}(z_i) > \min(\delta, e^{-\alpha i})$.

(A3) *There exist $\{\ell_j\}$ (to be specified in Sect. 4.4) so that w_i^* is correctly aligned with ε_0-error when $z_i \in \mathcal{C}^{(1)}$.*

(A4) $|w_i^*| > \frac{1}{2}c_2 e^{\lambda i}$ *where c_2 is as in Lemma 3.5.*

Our next assumption gives the relation between z_i and $\phi(z_i)$. Let $\hat{\beta}$ be such that $\alpha << \hat{\beta} << 1$. For $z_0, \xi_0 \in R_1$, let $\hat{p}(z_0, \xi_0)$ be the smallest $j > 0$ such that $|z_j - \xi_j| \geq e^{-\hat{\beta} j}$. For reasons to be explained in Sect. 4.3B, we will be interested in a range of p near $\hat{p}(z_0, \xi_0)$. Inside each $Q^{(k)}$, let

$$B^{(k)} = \{z \in Q^{(k)} : |z - z_0^*(Q^{(k)})|_h < b^{\frac{1}{5}k}\}.$$

(A5) How critical orbits influence nearby orbits *For $z_0 = z_0^*(Q^{(k)})$ and $\xi_0 \in Q^{(k)} \setminus B^{(k)}$, $k \leq \theta N$, the following hold for all $p \in [\hat{p}(z_0, \xi_0), (1 + \frac{9}{\lambda}\alpha)\hat{p}(z_0, \xi_0)]$:*

(i) (Length of bound period) *Suppose $|z_0 - \xi_0| = e^{-h}$. Then*

$$\frac{1}{3 \ln \|DT\|}\ h \ \leq \ p \ \leq \ \frac{3}{\lambda} h$$

the first inequality being valid if $\frac{1}{3\ln\|DT\|}h \leq k\theta^{-1}$ and the second if $\frac{3}{\lambda}h \leq k\theta^{-1}$.

(ii) (Partial derivative recovery) *If $p \leq k\theta^{-1}$, then $|w_p(z_0)||\xi_0 - z_0| \geq e^{\frac{1}{3}\lambda p}$.*

(iii) (Quadratic nature of turns) *Let γ be the $\mathcal{F}_k$-leaf segment joining ξ_0 to $B^{(k)}$. Then for all $\eta_0 \in \gamma$ and $\ell(\eta_0) < i \leq \min\{p, k\theta^{-1}\}$,*

$$|\eta_i - z_i| = \frac{1}{2}\left(|\frac{de_1}{ds}(z_0)| \pm \mathcal{O}(b)\right) \cdot \left(|w_i(z_0)| \pm \mathcal{O}(|\eta_0 - z_0|^{\frac{1}{2}})\right) \cdot |\eta_0 - z_0|^2.$$

Here $\ell(\eta_0)$ is defined by $b^{\frac{\ell(\eta_0)}{2}} = |\eta_0 - z_0|$, and $e_1 = e_1(S)$ where $S = S(\mathbf{v}, \tau_k)$, τ_k being the tangent to the $\mathcal{F}_k$-leaf through z_0.

This completes the formulation of the five statements (A1)–(A5). We also write (A1)(N)–(A5)(N) when more than one time frame is involved. The rest of this section contains some immediate clarifications.

Three important time scales We point out that in the dynamical picture described by (A1)–(A5), there are three distinct time scales: $\theta N << \alpha N << N$. The fastest time scale, N, gives the number of times the map is iterated. The slowest, θN, is the number of generations of critical regions and critical points constructed. The middle time scale, which is on the order of αN ($\alpha^* N$ to be precise), is an upper bound for the lengths of the bound periods initiated by critical orbits returning to $\mathcal{C}^{(1)}$ at times $\leq N$ (this follows from (A2) and (A5)(i) combined).

We assume (A1)–(A5) *for the rest of Section 4.*

4.2 Clustering of critical orbits

In Sect. 4.1, we presented a viewpoint – convenient for some practical purposes – in which a critical point $z_0^*(Q^{(k)})$ in each component $Q^{(k)}$ of $\mathcal{C}^{(k)}$ is singled out for special consideration. To understand the relation among the points in $\Gamma_{\theta N}$, it is more fruitful to group them into clusters. We propose here to view these clusters as represented by $B^{(k)}$. To justify this view, we prove

Lemma 4.1 *For all $k < \hat{k} < \theta N$, if $Q^{(\hat{k})} \subset Q^{(k)}$, then*

$$|z_0^*(Q^{(k)}) - z_0^*(Q^{(\hat{k})})| < Kb^{\frac{k}{4}}$$

and $B^{(\hat{k})} \subset B^{(k)}$.

The proof of this lemma uses the technical estimate below. Both results rely on the geometric information on $Q^{(k)}$ in (A1). Proofs are given in Appendix A.12.

Lemma 4.2 *Let $k < \hat{k}$, $Q^{(\hat{k})} \subset Q^{(k)}$, $z \in Q^{(k)}$, $\hat{z} \in Q^{(\hat{k})}$, and let γ and $\hat{\gamma}$ be the $\mathcal{F}_k$- and $\mathcal{F}_{\hat{k}}$-leaves containing z and $\hat{z}$ respectively. Let τ and $\hat{\tau}$ be the tangent vectors to γ and $\hat{\gamma}$ at z and $\hat{z}$. Then*

$$\angle(\tau, \hat{\tau}) \ \le \ b^{\frac{k}{4}} \ + \ K\delta^{-3}b \cdot |z - \hat{z}|_h.$$

Evolution of critical blobs A theme that runs through our discussion is that orbits emanating from the same $B^{(k)}$ are viewed as essentially indistinguishable for $k\theta^{-1}$ iterates. Informally, we call these finite orbits of $B^{(k)}$ *critical blobs*.

Recall that θ is assumed so that $b^\theta < \|DT\|^{-20}$. This implies that for all $i \le k\theta^{-1}$, $\text{diam}(T^i B^{(k)}) < b^{\frac{1}{5}k}\|DT\|^i < (b^\theta)^{\frac{1}{5}i}\|DT\|^i$. This is $<< e^{-\alpha i}$, the minimum allowed distance to the critical set (see (A2)).

Obviously, we cannot iterate indefinitely and hope that $T^i B^{(k)}$ remains small; that is why we regard $z_0^*(Q^{(k)})$ as **active** for only $k\theta^{-1}$ iterates. The word "active" here refers to both (i) prescribed behavior for z_i (as in (A2)–(A4)) and (ii) the use of z_i as guiding critical orbit or in the sense of (A5).

It is useful to keep in mind the following dynamical picture:

At time $i = 0$, T has a set $B^{(1)}$ corresponding to each critical point of f. For $i \le \theta^{-1}$, the T^i-images of $B^{(1)}$ are relatively small, so that $\{T^i B^{(1)}\}_{i=0,1,\cdots,\theta^{-1}}$ for each $B^{(1)}$ can be treated as a single orbit.

As i increases, the sizes of $T^i B^{(1)}$ become larger, eventually becoming too large for $\{T^i B^{(1)}\}_{i=0,1,\cdots}$ to be treated as a single orbit. We stop considering these critical blobs long before that time, however. At time $i = \theta^{-1}$, we replace each $T^{\theta^{-1}} B^{(1)}$ by the collection of $T^{\theta^{-1}} B^{(2)}$ contained in it. For $\theta^{-1} < i \le 2\theta^{-1}$, $T^i B^{(2)}$ are again relatively small, and so can be viewed as a finite collection of orbits. At time $i = 2\theta^{-1}$, each $T^{2\theta^{-1}} B^{(2)}$ is replaced by the collection of $T^{2\theta^{-1}} B^{(3)}$ inside it, and so on.

As i increases, the number of relevant critical blobs increases, each becoming smaller in size. Blobs that have separated move about "independently". By virtue of (A2), they are allowed to come closer to the critical set with the passage of time.

We finish by recording a technical fact that will be used in conjunction with Lemma 3.8.

Lemma 4.3 *For any $C^2(b)$-curve $s \mapsto l(s)$ traversing a given $B^{(k)} \subset Q^{(k)}$, there exists a point in l, denoted by $l(0)$, such that*

$$\angle(l'(0), \tau(z_0)) < b^{\frac{k}{4}}$$

where $z_0 = z_0^*(Q^{(k)})$ and $\tau(z_0)$ is tangent to the leaf of $\mathcal{F}_k$ at z_0.

As with Lemma 4.2, Lemma 4.3 is proved by a straightforward application of Sublemma A.12.1 in Appendix A.12. We leave it as an exercise.

4.3 Bound periods

Let $z_0 \in \Gamma_{\theta N}$ be of generation k, and let $z_i \in \mathcal{C}^{(1)}$, $i \leq k\theta^{-1}$. In Sect. 4.1, we assigned to z_i a guiding critical point $\phi(z_i) \in \Gamma_{\theta N}$. (A5)(i)–(iii) hold for all $p \in [\hat{p}, (1+\frac{9}{\lambda}\alpha)\hat{p}]$ where $\hat{p} = \hat{p}(z_i, \phi(z_i))$. We now choose a specific number $p = p(z_i)$ in this range with certain desirable properties. This number will be called the **bound period** of z_i.

A. Remarks on $\phi(\cdot)$ and $d_{\mathcal{C}}(\cdot)$

In general, when $z_i \in \mathcal{C}^{(1)}$, it is in many $Q^{(j)}$. Since $\mathcal{C}^{(j)}$ for larger j give better approximations of the eventual critical set, it is natural to want to define $d_{\mathcal{C}}(z_i)$ using the largest j possible. We do not do exactly that; instead, we take $\phi(z_i)$ to be $z_0^*(Q^{(\hat{j})}(z_i))$ where $\hat{j}$ is the largest $j \leq \alpha^*\theta i$ such that $z_i \in Q^{(j)}$. The significance of this upper bound on j will become clear in Section 6. For now we observe

Lemma 4.4 *(i) $|z_i - \phi(z_i)| >> b^{\frac{\hat{j}}{5}}$; in particular, $z_i \in Q^{(\hat{j})} \setminus B^{(\hat{j})}$, so (A5) applies.*
(ii) Let $p \in [\hat{p}, (1+\frac{9}{\lambda}\alpha)\hat{p}]$ be as in (A5). Then $p \leq \hat{j}\theta^{-1}$.

Proof: *Case 1.* $\hat{j}+1 \leq \alpha^*\theta i$. This implies $z_i \in Q^{(\hat{j})} \cap R_{\hat{j}+1} \setminus Q^{(\hat{j}+1)}$, i.e. $d_{\mathcal{C}}(z_i) > e^{-\lambda(\hat{j}+1)}$. Hence $b^{\frac{\hat{j}}{5}} << d_{\mathcal{C}}(z_i)$ and $p << \hat{j}\theta^{-1}$ by (A5)(i).

Case 2. $\hat{j}+1 > \alpha^*\theta i$. Using this relation between i and $\hat{j}$, we see that $d_{\mathcal{C}}(z_i) > e^{-\alpha i} > e^{-\frac{\alpha}{\alpha^*}\theta^{-1}(\hat{j}+1)}$, which we check is $>> b^{\frac{\hat{j}}{5}}$ by the definition of b^θ and the facts that $\frac{\alpha}{\alpha^*} = \frac{\lambda}{6}$ and $e^\lambda < \|DT\|$. Also, $p \leq \frac{3}{\lambda}\alpha i$ by (A2) and (A5)(i). This upper bound is $= \frac{1}{2}\alpha^* i \leq \frac{1}{2}(\hat{j}+1)\theta^{-1} \leq \hat{j}\theta^{-1}$. □

We use $\phi(z_i)$ to define $d_{\mathcal{C}}(z_i)$. One may ask if it makes a significant difference if some other critical point is used. The answer is that when $d_{\mathcal{C}}(z_i)$ is relatively large, for example when $d_{\mathcal{C}}(z_i) > b^{\frac{1}{5}}$, it does not matter much, but when $d_{\mathcal{C}}(z_i)$ is small, the values of $|\hat{z} - z_i|$ or even $|\hat{z} - z_i|_h$ can vary nontrivially as $\hat{z}$ varies over $\Gamma_{\theta N}$. For the same reason, for $z_i, z'_j \in \mathcal{C}^{(1)}$, we cannot conclude – without further information – that $|d_{\mathcal{C}}(z_i) - d_{\mathcal{C}}(z'_j)| \approx |z_i - z'_j|$, for z_i and z'_j can be in very different "layers" of the critical structure, resulting in $\phi(z_i)$ and $\phi(z'_j)$ being relatively far apart.

We do have the following:

Lemma 4.5 *(i) Let $z \in Q^{(k)} \setminus B^{(k)}$. Then for all $\hat{z}, \tilde{z} \in \Gamma_{\theta N} \cap B^{(k)}$ (meaning the $B^{(k)}$ inside $Q^{(k)}(z_i)$), we have $|z - \hat{z}| = (1 \pm \mathcal{O}(b^{\frac{k}{20}}))|z - \tilde{z}|$.*
(ii) Suppose $\hat{z}_0 = \phi(z_i)$, and $\hat{z}_j \in \mathcal{C}^{(1)}$ for some $0 < j < \hat{p}(\hat{z}_0, z_i)$. Then $d_{\mathcal{C}}(z_{i+j}) = (1 \pm \mathcal{O}(e^{-\frac{1}{2}\hat{\beta}j}))d_{\mathcal{C}}(\hat{z}_j)$.

Proof: (i) By Lemma 4.1, $|\hat{z} - \tilde{z}| < Kb^{\frac{k}{4}}$, and by assumption, z is $> b^{\frac{k}{5}}$ from the center of $Q^{(k)}$. This proves $|z - \hat{z}| = (1 \pm \mathcal{O}(b^{\frac{k}{20}}))\ |z - \tilde{z}|$.

(ii) By definition, $|z_{i+j} - \hat{z}_j| < e^{-\hat{\beta}j} << e^{-\alpha j}$, which is $< d_{\mathcal{C}}(\hat{z}_j)$ by (A2). As explained above, this in itself is insufficient for guaranteeing the asserted relationship between $d_{\mathcal{C}}(z_{i+j})$ and $d_{\mathcal{C}}(\hat{z}_j)$. We have, however, the following additional information: By (A5)(iii), there is a curve ω joining z_i to $\hat{z}_0$ such that $\text{diam}(T^j(\omega)) << e^{-\alpha j}$. Now suppose $\hat{z}_j \in \mathcal{C}^{(1)}$ is such that $\phi(\hat{z}_j) = z_0^*(Q^{(\hat{k})})$. Since $\hat{k} << j$, $T^j(\omega)$ is contained, or nearly contained, in $Q^{(\hat{k})}(\hat{z}_j)$. Part (i) now enables us to make the desired comparison. □

B. Definition of bound periods

Consider $z_0 \in \Gamma_{\theta N}$. For each i such that $z_i \in \mathcal{C}^{(1)}$, let $p(z_i)$ be the bound period of z_i to be defined. We say $\{p(z_i)\}$ has a *nested structure* if whenever $i < j$ are such that $z_i, z_j \in \mathcal{C}^{(1)}$ and $j < i + p(z_i)$, we have $j + p(z_j) \leq i + p(z_i)$.

To define $p(z_i)$, we start with $\hat{p}_i := \hat{p}(\phi(z_i), z_i)$ where $\hat{p}(\phi(z_i), z_i)$ is as defined in Sect. 4.1. There is no reason why $\{\hat{p}_i\}$ should have a nested structure. We call $j_0 < j_1 < \cdots < j_n$ a *chain of overlapping bound intervals* if $z_{j_k} \in \mathcal{C}^{(1)}$ and $j_k \in (j_{k-1}, j_{k-1} + \hat{p}_{j_{k-1}})$ for every $k \leq n$. Let Λ_i be the set of all integers $k > i$ such that there is a chain of overlapping bound intervals $j_0 < j_1 < \cdots < j_n$ with $j_0 = i$ and $j_n + \hat{p}_n \geq k$. We define $p(z_i) := i' - i$ where i' is the supremum of the set Λ_i. *A priori*, $p(z_i)$ can be $>> \hat{p}_i$; it can even be infinite. We prove in Lemma 4.6 below that this is not the case.

Lemma 4.6 *For all $z_0 \in \Gamma_{\theta N}$ and all $z_i \in \mathcal{C}^{(1)}$,*
(a) $p(z_i) < (1 + \frac{6}{\lambda}\alpha)\hat{p}_i$.
(b) $\{p(z_i)\}$ has a nested structure.

Proof: (a) For $z_i \in \mathcal{C}^{(1)}$, let j be such that $i < j < i + \hat{p}_i$. Then $d_{\mathcal{C}}(z_j) \approx d_{\mathcal{C}}((\phi(z_i))_{j-i}))$ by Lemma 4.5(ii). Applying (A2) to $\phi(z_i)$ and then (A5)(i) to z_j, we obtain $\hat{p}_j \leq \frac{3}{\lambda}\alpha(j-i) \leq \frac{3}{\lambda}\alpha\hat{p}_i$. If $j_0 < j_1 < \cdots < j_n$ is a chain of overlapping bound intervals with $j_0 = i$, then similar reasoning gives $\hat{p}_{j_k} \leq \frac{3}{\lambda}\alpha\hat{p}_{j_{k-1}}$, so that

$$\hat{p}_{j_0} + \hat{p}_{j_1} + \cdots + \hat{p}_{j_n} < (1 + \frac{3}{\lambda}\alpha + (\frac{3}{\lambda}\alpha)^2 + \cdots)\hat{p}_i < (1 + \frac{6}{\lambda}\alpha)\hat{p}_i.$$

Since this bound is valid for all chains, we have $p(z_i) < (1 + \frac{6}{\lambda}\alpha)\hat{p}_i$.

(b) We need to show that if $j \in (i, i+p(z_i))$, then $j + p(z_j) \leq i + p(z_i)$. Note that since $p(\cdot)$ is finite, there exists a chain of overlapping intervals $i = j_0 < \cdots < j_n$ such that $j_n + \hat{p}_{j_n} = i + p(z_i)$. If $j + p(z_j) > i + p(z_i)$, then the chain that goes from i to $i + p(z_i)$ combined with the one that goes from j to $j + p(z_j)$ forms a new chain starting from i and extending beyond $i + p(z_i)$. This contradicts the definition of $p(z_i)$. □

Let $\beta = \hat{\beta} - \frac{9}{\lambda}\ln\|DT\|\alpha$, and let $p(z_0, \xi_0)$ be the smallest j such that $|z_j - \xi_j| \geq e^{-\beta j}$. An easy calculation gives $\hat{p}(z_0, \xi_0)(1 + \frac{9}{\lambda}\alpha) \leq p(z_0, \xi_0)$.

Clarification: Relation between $\hat{p}(\cdot,\cdot), p(\cdot,\cdot)$ and $p(z_i)$ for $z_0 \in \Gamma_{\theta N}$

1. These definitions are brought about by the tension between our desire to define "bound periods" in terms of the distances separating two orbits, and the advantages of having a nested structure for bound periods along individual orbits. We showed in Lemma 4.6 that a nested structure can be arranged if we allow some flexibility in scale when measuring distances, so that for $z_0 \in \Gamma_{\theta N}$, there exist $\{p(z_i)\}$ with a nested structure *and* satisfying $\hat{p}(z_i, \phi(z_i)) \leq p(z_i) \leq p(z_i, \phi(z_i))$.

2. In general, in results pertaining to a single bound period (e.g. Propositions 5.1), we use $p(\cdot,\cdot)$, so that the result is valid for as long a duration as possible. In situations in which we follow the long range evolution of single orbits (e.g. Sect. 5.2), a nested structure arranged as above is used.

C. Bound and free states

For $z_0 \in \Gamma_{\theta N}$ of generation k, we now have a decomposition of the orbit $z_0, z_1, \cdots, z_{k\theta^{-1}}$ into intervals of **bound** and **free periods**, i.e. we say z_i is free if and only if it is not in a bound period. Calling the maximal bound intervals **primary bound periods**, the nested structure above allows us to speak of **secondary bound periods**, tertiary bound periods, and so on. Returns to $\mathcal{C}^{(1)}$ at the beginning of primary bound periods are called **free returns**, while returns at the start of seconding or higher order bound periods are called **bound returns**.

4.4 The splitting algorithm applied to $DT^i_{z_0}(\mathbf{v})$, $z_0 \in \Gamma_{\theta N}$

The considerations below are motivated by the discussion in Sect. 3.8 and by Lemma 3.10 in particular. We continue to use the notation there.

A. Splitting periods

Fix $z_0 \in \Gamma_{\theta N}$. We explain how the ℓ_i at return times i in Sect. 4.1 are chosen. From Sect. 3.8, we see that the following properties are desirable:
(i) $\ell_i \geq 2$;
(ii) $|DT^j_{z_i}(\mathbf{v})| > K^{-1}$ for $j = 1, 2, \cdots, \ell_i$;
(iii) the intervals $I_i = [i, i + \ell_i)$ have the nested property.

We explain why these properties can, in principle, be arranged. Let i be fixed for now. To obtain property (ii), we use Lemma 3.2 and the fact that $\phi(z_i)$ is a critical point. Observe that (ii) always holds for $\ell \leq 2$, so (i) is not a problem. In general, as a first approximation, let $\hat{\ell}$ be such that $b^{\frac{\hat{\ell}}{3}} = d_{\mathcal{C}}(z_i)$. We claim that (ii) holds for all $\ell \leq \frac{5}{3}\hat{\ell}$. To justify this claim, we need to check that $\ell \leq$ the order of $\phi(z_i)$ as a critical point (this follows from Lemma 4.4(i)), and that the expanding property $|DT^j_{\phi(z_i)}(\mathbf{v})| > K_0^{-1}$ passes to a disk of radius $> d_{\mathcal{C}}(z_i)$ (Lemma 3.2). To achieve (iii), we need to show that if z_j is a return for $i < j < i + \hat{\ell}_i$, then $\hat{\ell}_j < K\alpha(\log \frac{1}{b})^{-1}\hat{\ell}_i$ (for which we follow the proof of Lemma 4.6).

Algorithm for choosing ℓ_i in (A3): Let $\hat{\ell}_i$ be as above. First we set $\ell'_i = \max\{2, \hat{\ell}_i\}$, then increase ℓ'_i to ℓ^*_i if necessary so that the intervals $I_i = [i, i + \ell^*_i)$ are nested, and finally, for convenience, let $\ell_i = \ell^*_i + 1$ or 2 to ensure that no splitting period ends at a return or at the step immediately after a return.

B. Correct alignment implies correct splitting

For $z_0 \in \Gamma_{\theta N}$, we let $w^*_i, i = 1, 2, \cdots$, be generated by the splitting algorithm in Sect. 3.8 using the ℓ_i above. Our next lemma connects the "correct alignment" assumption in (A3) to hypothesis (a) in Lemma 3.10. Suppose $z_i \in \mathcal{C}^{(1)}$ and write $w^*_i = A_i e_{\ell_i}(S) + B_i\mathbf{v}$ where $S = S(\mathbf{v}, w^*_i)$.

Lemma 4.7 *If w^*_i is correctly aligned with ε-error where $\varepsilon << K_1^{-1}$, then*

$$\frac{|B_i|}{|A_i|} > \frac{1}{2}K_1^{-1}d_{\mathcal{C}}(z_i)$$

where K_1^{-1} is the lower bound of $|\frac{d}{ds}e_1|$ in Lemma 3.7.

When the conclusion of Lemma 4.7 holds, we say w^*_i **splits correctly**. We caution that when $d_{\mathcal{C}}(\cdot)$ is very small, b-horizontal vectors do not necessarily split correctly.

Corollary 4.1 *If at all returns, w^*_i is correctly aligned with ε-error where $\varepsilon << K_1^{-1}$, then w^*_i is b-horizontal and splits correctly.*

A proof of Lemma 4.7 is given in Appendix A.13. Corollary 4.1 follows from a direct application of Lemma 3.10 once we note that $\frac{1}{2}K_1^{-1}d_{\mathcal{C}}(z_i) >> b^{\frac{\ell_i}{2}}$.

5 Properties of Orbits Controlled by Critical Set

We continue to assume (A1)–(A5). This section contains a general discussion of the extent to which the orbits of $z_0 \in \Gamma_{\theta N}$ can be used to guide other (noncritical) orbits, or, put differently, the extent to which (ξ_0, w_0) for arbitrary $\xi_0 \in R_1$ and $w_0 \in X_{\xi_0}$ can be *controlled* by $\Gamma_{\theta N}$. The word *control* is given a formal definition in Sect. 5.2.

5.1 Copying segments of critical orbits

For z, ξ in the same component of $\mathcal{C}^{(1)}$, let $p(z,\xi)$ be as defined in Sect. 4.3B, i.e. it is the smallest j such that $|T^j z - T^j \xi| > e^{-\beta j}$. For $z_0 = z_0^*(Q^{(k)}) \in \Gamma_{\theta N}$ and $\xi_0, \xi_0' \in Q^{(1)}(z_0)$, we let $p(z_0;\xi_0,\xi_0') := \min\{p(z_0,\xi_0), p(z_0,\xi_0'), k\theta^{-1}\}$. Unlike (A5), we do not presuppose here any geometric relationship between ξ_0, ξ_0' and z_0. In particular, $p(z_0;\xi_0,\xi_0')$ may not be in the time range for which (A5) is applicable.

Let $w_0(\xi_0) = w_0(\xi_0') = w_0(z_0) = \mathbf{v}$. We apply the splitting algorithm to z_0, ξ_0 and ξ_0' for $i \le p(z_0;\xi_0,\xi_0')$ using for all three points the splitting periods for z_0 as specified in Sect. 4.4. Our next proposition compares $w_i^*(\xi_0)$ and $w_i^*(\xi_0')$. Let

$$\Delta_n(\xi_0,\xi_0') := \sum_{s=0}^{n} b^{\frac{s}{4}} 2^{\ell_{n-s}} |\xi_{n-s} - \xi_{n-s}'| \tag{5}$$

where ℓ_{n-s} is the length of the longest splitting period z_{n-s} find itself in, 0 if z_{n-s} is out of all splitting periods.

Proposition 5.1 *There is a constant K_1 such that for all ξ_0, ξ_0' and z_0 as above and $i < p(z_0;\xi_0,\xi_0')$,*

$$\frac{|w_i^*(\xi_0)|}{|w_i^*(\xi_0')|}, \quad \frac{|w_i^*(\xi_0')|}{|w_i^*(\xi_0)|} \;\le\; \exp\left\{ K_1 \sum_{n=1}^{i-1} \frac{\Delta_n(\xi_0,\xi_0')}{d_{\mathcal{C}}(z_n)} \right\} \tag{6}$$

and

$$\angle(w_i^*(\xi_0), w_i^*(\xi_0'))) \le b^{\frac{1}{2}} \Delta_{i-1}(\xi_0,\xi_0'). \tag{7}$$

This proposition would not be very useful without *a priori* bounds for the quantities involved. We explain how a bound for the right side of equations (6) and (7) can be arranged.

Lemma 5.1 *Assume that (i) β is sufficiently large compared to α, (ii) δ is sufficiently small depending on α and β, and (iii) b is small enough. Then for all z_0, ξ_0, ξ_0', i and n as above,*

$$\Delta_n < 2e^{-\frac{1}{2}\beta n} << \varepsilon_0 d_{\mathcal{C}}(z_n)$$

and

$$K_1 \sum_{n=1}^{i-1} \frac{\Delta_n(\xi_0,\xi_0')}{d_{\mathcal{C}}(z_n)} << 1.$$

Proposition 5.1 and Lemma 5.1 are proved in Appendix A.14. Our first application of Proposition 5.1 is to the case where $\xi_0' = z_0$. We assume α, β, δ and b are chosen so that the following is an immediate corollary of Proposition 5.1 and Lemma 5.1.

Corollary 5.1 *Let $z_0 \in \Gamma_{\theta N}$ be of generation k. Then for $\xi_0 \in Q^{(1)}(z_0)$ and $i < \min\{k\theta^{-1}, p(z_0,\xi_0)\}$,*

(i) $|w_i^*(\xi_0)| > \frac{1}{4} c_2 e^{\lambda i}$;

(ii) at return to $\mathcal{C}^{(1)}$, $w_i^*(\xi_0)$ is correctly aligned with $2\varepsilon_0$-error.

5.2 A formal notion of "control"

Very roughly, a controlled orbit is one obtained by splicing together a finite number of orbit segments each one of which is either free or bound to a critical orbit. The goal of this subsection is to identify sufficient conditions at the joints that will guarantee that the resulting orbit has desirable properties.

Let $\xi_0 \in R_1$ be an arbitrary point.

Definition 5.1 *We say ξ_0 is* **controlled** *by $\Gamma_{\theta N}$ for M iterates, or equivalently, the orbit segment $\xi_0, \xi_1, \cdots, \xi_{M-1}$ is controlled by $\Gamma_{\theta N}$, if the following hold: whenever $\xi_i \in \mathcal{C}^{(1)}$, $0 \le i < M$, there exists $Q^{(k)}$, $k \le \theta N$, such that*

(i) $\xi_i \in Q^{(k)} \setminus B^{(k)}$, and

(ii) $\min(\hat{p}(z_0, \xi_i), M - i) \le k\theta^{-1}$ where $z_0 = z_0^*(Q^{(k)})$.

Condition (i) guarantees that (A5) applies to ξ_i. Condition (ii) guarantees that the guiding orbit z_0 remains active until either the bound period or the period of control expires.

Orbits controlled by $\Gamma_{\theta N}$ can be seen as follows:

Let $n_1 \ge 0$ be the first time $\xi_i \in \mathcal{C}^{(1)}$. For $i \le n_1$, we regard ξ_i as free. At time n_1, we assume there exists $z_0 \in \Gamma_{\theta N}$ satisfying the conditions in Definition 5.1. Such a critical point is usually not unique. We make an arbitrary choice, call it $\tilde{\phi}(\xi_{n_1})$, and defined $\tilde{d}_{\mathcal{C}}(\xi_{n_1}) := |\tilde{\phi}(\xi_{n_1}) - \xi_{n_1}|$. From Lemma 4.1 we see that among the admissible choices of $\tilde{\phi}(\xi_{n_1})$, $\tilde{d}_{\mathcal{C}}(\xi_{n_1})$ do not differ substantially. Instead of $\tilde{\phi}(\cdot)$ and $\tilde{d}_{\mathcal{C}}(\cdot)$, we write $\phi(\cdot)$ and $d_{\mathcal{C}}(\cdot)$ for notational simplicity.

For the next $\hat{p}(\xi_{n_1}, \phi(\xi_{n_1}))$ iterates, we think of ξ_{n_1} as bound to $\phi(\xi_{n_1})$ as in Sect. 5.1, inheriting from the orbit of $\phi(\xi_{n_1})$ bound and splitting periods. At the end of the $\hat{p}(\xi_{n_1}, \phi(\xi_{n_1}))$ iterates, there may be some bound periods that have not expired. In the interest of a nested structure for bound periods, we extend $\hat{p}(\xi_{n_1}, \phi(\xi_{n_1}))$ to p_1, so that $n_1 + p_1$ is the first moment when all bound periods initiated before $n_1 + p_1$ have expired. For the same reason as in the proof of Lemma 4.6, we have $p_1 < (1 + \frac{6}{\lambda}\alpha)\hat{p}(\xi_{n_1}, \phi(\xi_{n_1}))$. (This uses condition (i) in Definition 5.1.)

We regard $\xi_{n_1+p_1}$ as "free", and think of its orbit as remaining free until n_2, the first time $\ge n_1 + p_1$ when $\xi_{n_2} \in \mathcal{C}^{(1)}$. For a controlled orbit, we are guaranteed the existence of at least one critical point satisfying the conditions of Definition 5.1 with respect to ξ_{n_2}. We think of ξ_{n_2} as bound to $\phi(\xi_{n_2})$ for p_2 iterates, and so on.

The process continues until the period of control expires. Splitting periods with a nested structure are defined similarly.

Next we discuss what it means for a (ξ_0, w_0)-pair to be controlled. Let ε_1 be such that $4\varepsilon_0 < \varepsilon_1 << K_1^{-1}$ where ε_0 and K_1 are as in Sect. 4.1. Let ξ_0 be a controlled orbit, and let $w_0 \in X_{\xi_0}$ be an arbitrary unit vector. The vectors $w_i^*(\xi_0)$ are obtained by using the splitting periods defined above.

Definition 5.2 *We say (ξ_0, w_0) is* **controlled** *by $\Gamma_{\theta N}$ for M iterates, or equivalently, the sequence $(\xi_0, w_0), \cdots, (\xi_{M-1}, w_{M-1})$ is controlled by $\Gamma_{\theta N}$, if ξ_0 is controlled for M iterates and the following holds: whenever $\xi_i \in \mathcal{C}^{(1)}$, $0 \le i < M$, w_i^* is correctly aligned with ε_1-error, i.e. if $\phi(\xi_i)$ is of generation j and $d_{\mathcal{C}}(\xi_i)$ is as above, then $\angle(w_i^*(\xi_0), \tau) < \varepsilon_1 d_{\mathcal{C}}(\xi_i)$ where τ is the tangent to the leaf of $\mathcal{F}_j$ through ξ_i.*

A slightly expanded definition: It is convenient to expand the definition of control to allow the following initial condition: If $\xi_0 \in \mathcal{C}^{(1)}$ and $w_0 = \mathbf{v}$, then the conditions in Definitions 5.1 and 5.2 are waived at time 0. (The rationale for this inclusion is that since no splitting occurs at time 0, derivative recovery is automatic.)

The properties of a controlled (ξ_0, w_0)-pair can be summarized as follows:

Proposition 5.2 *Assume that (ξ_0, w_0) is controlled by $\Gamma_{\theta N}$ for M iterates. Then*

(1) there exist $0 \le n_1 < n_1 + p_1 \le n_2 < n_2 + p_2 \le n_3 \cdots < M$ such that for each i,

(i) there is $\phi(\xi_{n_i}) \in \Gamma_{\theta N}$ to which ξ_{n_i} is bound for p_i iterates, $p_i \sim log\frac{1}{d_{\mathcal{C}}(\xi_{n_i})}$;

(ii) $\xi_j \notin \mathcal{C}^{(1)}$ for $n_i + p_i \le j < n_{i+1}$;

(2) w_i^* has the following growth properties:

$$\frac{|w^*_{n_i+p_i}|}{|w^*_{n_i}|} > K^{-1}e^{\frac{1}{3}\lambda p_i}; \quad \frac{|w^*_{n_{i+1}}|}{|w^*_{n_i+p_i}|} > \frac{1}{2}c_2 e^{\frac{1}{4}\lambda_0(n_{i+1}-(n_i+p_i))}.$$

Proof: (1) is a summary of the discussion following Definition 5.1; the estimate for p_i uses (A5)(i). The second inequality in (2) follows immediately from Lemma 3.5. The first is proved as follows: By Proposition 5.1, we have $|DT^{p_i}_{\xi_{n_i}}(\mathbf{v})| > \frac{1}{2}|DT^{p_i}_{\phi(\xi_{n_i})}(\mathbf{v})|$. For purposes of this proof, it is simplest to split off a vector from $w^*_{n_i}$ that is known to contract for p_i iterates. Let $e = e_{p_i}$ be the most contracted direction for $DT^{p_i}_{\xi_{n_i}}$ on $S = S(w_{n_i}, \mathbf{v})$. We claim that if $w^*_{n_i} = Ae + B\mathbf{v}$, then $|B| > K^{-1}d_{\mathcal{C}}(\xi_{n_i})$. (Reason: correct splitting is assumed at time n_i; the (normal) splitting period, ℓ, is $<< p_i$; and so $\angle(e, e_\ell) < (Kb)^\ell$, which is $<< b^{\frac{\ell}{3}} \approx d_{\mathcal{C}}(\xi_{n_1})$.) (A5)(ii) then gives $|DT^{p_i}_{\xi_{n_i}}(B\mathbf{v})| > K^{-1}|DT^{p_i}_{\phi(\xi_{n_i})}(\mathbf{v})|d_{\mathcal{C}}(\xi_{n_i}) > K^{-1}e^{\frac{1}{3}\lambda p_i}$. The addition of $DT^{p_i}_{\xi_{n_i}}(Ae)$ has negligible effect. □

In Sect. 2.2, we proved that for a class of "good" 1D maps, *every* orbit not passing through the critical set has the properties in Proposition 5.2. A consequence of the definition of control, therefore, is that (ξ_0, w_0)-pairs have 1D behavior.

5.3 A collection of useful facts

We record in this subsection a miscellaneous collection of facts related to controlled orbits that are used in the future. Lemmas 5.2–5.6 are proved in Appendices A.15–A.17. Proposition 5.3 is proved in Appendix A.18.

A. Relation between $|w_i|$ and $|w_i^*|$

Lemma 5.2 *Assume that (ξ_0, w_0) is controlled by $\Gamma_{\theta N}$ for M iterates. Under the additional assumption that $d_{\mathcal{C}}(\xi_i) > \min(\delta, e^{-\alpha i})$ for all $i < M$, we have*

$$K^{-\varepsilon i}|w_i^*(\xi_0)| \le w_i(\xi_0) \le K^{\varepsilon i}e^{2\alpha i}|w_i^*(\xi_0)|, \quad \varepsilon = K\alpha\theta. \tag{8}$$

B. Angles at bound returns

Lemma 5.3 *Let ξ_0 be controlled by $\Gamma_{\theta N}$ for M iterates, and assume that at all* free *returns, w_i^* is correctly aligned with $< \varepsilon_1$-error. Then at all* bound *returns, w_i^* is correctly aligned with $< 3\varepsilon_0$-error.*

Since ε_1, the error in alignment of w_i^* at free returns, can be $>> 3\varepsilon_0$, Lemma 5.3 implies that the magnitudes of the errors at free returns are not reflected in the angles at returns during ensuing bound periods provided they are within an acceptable range.

C. Growth of $|w_i|$, $|w_i^*|$ and $\|DT^i\|$

The next three results provide more detailed information on derivative growth than Proposition 5.2.

Lemma 5.4 *There exists λ' with $\lambda' > \frac{1}{3}\lambda - \mathcal{O}(\sqrt{b})$ such that if (ξ_0, w_0) is controlled by $\Gamma_{\theta N}$ for M iterates , then for every $0 \le k < n < M$,*

$$|w_n^*| \ge K^{-1}d_{\mathcal{C}}(\xi_j)e^{\lambda'(n-k)}|w_k^*|$$

where j is the first time $\ge k$ when a bound period extending beyond time n is initiated. If no such j exists, set $d_{\mathcal{C}}(\xi_j) = \delta$.

Lemma 5.5 *The setting and notation are as in Lemma 5.4. If in addition ξ_n is free, then*

$$|w_n| > \delta K^{-K\theta(n-k)} e^{\lambda'(n-k)} |w_k|.$$

If ξ_n is a free return, then δ on the right side can be omitted.

We finish by recording a technical lemma that will be used in Part III.

Lemma 5.6 *Suppose (ξ_0, w_0) is controlled for M iterates by $\Gamma_{\theta N}$, and that $d_{\mathcal{C}}(\xi_i) > e^{-\alpha i}$ for all $i \le M$. Then there exist constants K and $\hat{\lambda} > 0$ slightly smaller than $\frac{1}{3}\lambda$ such that for every $0 \le s < i < M$,*

$$\|DT^{i-s}_{\xi_s}\| \le K e^{-\hat{\lambda} s} |w_i|.$$

D. Quadratic properties of turns

We consider in this paragraph the special situation where the critical point on a $C^2(b)$-curve is controlled. The quadratic distance formula in Proposition 5.3 is used to prove estimates of the kind in (A5).

The precise setting is as follows: Let $\gamma \subset \mathcal{C}^{(1)}$ be a $C^2(b)$-curve, and let $z_0 \in \gamma$ be a critical point of order M on γ in the sense of Definition 3.2. (There is no restriction on the size of M; it can be $> N$.) We assume that

(1) $(z_0, \mathbf{v})$ is controlled by $\Gamma_{\theta N}$ for M iterates; and
(2) $d_{\mathcal{C}}(z_i) > \min(\delta, e^{-\alpha i})$ for all $0 < i \le M$.

Let $s \mapsto \xi_0(s)$ be the parametrization of γ by arclength with $\xi_0(0) = z_0$.

Proposition 5.3 *For given $s_1 > 0$, let $p(s_1) = \min\{p(\xi_0(s_1), z_0), M\}$. Then for all $0 < s \le s_1$ and $i \in [\ell(s), p(s_1)]$ with $\ell(s) = 2\frac{\log s}{\log b}$, we have*

$$|\xi_i(s) - z_i| \approx \frac{1}{2}\left|\frac{d}{ds} e_1(0)\right| |w_i(0)| \; s^2$$

where $e_1 = e_1(S)$ and $S = S(\gamma', \mathbf{v})$.

6 Identification of Hyperbolic Behavior: Formal Inductive Procedure

6.1 Global constants (mostly review)

For $N = 1, 2, \cdots$, we define below a set of "good" maps $T : X \to X$ denoted by

$$\mathcal{G}_N = \mathcal{G}_N(f_0, K_0, a, b; \; \lambda, \alpha; \; \delta, \beta, \varepsilon_0, \theta).$$

The arguments on the right side can be understood conceptually as follows:

1. The first group consists of $f_0 \in \mathcal{M}$ and three constants, K_0, a and b. These items appear in the *Standing Hypotheses* at the beginning Part II; they define an open set in the space of C^3 embeddings of X into itself.

2. In the next group are λ and α, two constants that appear in (A2) and (A4). As we will see, (A2) and (A4) play a special role in determining if T in the open set above is in $\mathcal{G}_N$; they are analogous to (G1) and (G2) for 1D maps (see Sect. 2.2).

3. Unlike the situation in 1D, auxiliary constructions are needed before we are able to properly formulate (A2) and (A4). The constants in the last group, namely $\delta, \beta, \varepsilon_0$ and θ,

appear in these auxiliary constructions. They do not directly impact whether a map is in $\mathcal{G}_N$, but help maintain uniform estimates in the constructions.

In the definition of $\mathcal{G}_N$, f_0 is chosen first; it can be any element of $\mathcal{M}$. We then fix K_0, which can be any number $> \|f_0\|_{C^3}$. Precise conditions imposed on the rest of the constants are given in the text. We review below their (rough) meanings and give the order in which they are chosen. To ensure consistency in our choices, it is important that (i) only upper bounds are imposed on each constant, and (ii) these bounds are allowed to depend only on the constants higher up on the list (in addition to f_0, K_0, and m, the dimension of X). Except for λ, all the constants listed below are $<< 1$ and must be taken to be as small as necessary.

Important constants: their meanings, and the order in which they are chosen

– λ is our targeted Lyapunov exponent; it can be anything $< \frac{1}{5}\lambda_0$ where λ_0 is a growth rate of $|f_0'|$ (see Definition 1.1). Once chosen, it is fixed throughout.

– Next we fix α and β and think of $e^{-\alpha n}$ and $e^{-\beta n}$ as representing two small scales. The requirements are that $0 < \alpha, \beta << \min\{\lambda, 1\}$ and $\beta > K\alpha$ for some K depending on f_0 and K_0. The meaning of α is that critical orbits are not allowed to approach the critical set at speeds faster than $e^{-\alpha n}$. Two orbits $\{z_i\}$ and $\{z_i'\}$ with $|z_i - z_i'| < e^{-\beta i}$ are to be thought of as "bound together".

– ε_0, which depends only on f_0, K_0 and m, has the following meaning: For $z \in \mathcal{C}^{(1)}$, vectors $v \in X_z$ that make angles $< \varepsilon_0 d_{\mathcal{C}}(z)$ with certain $\mathcal{F}_k$-leaves are viewed as "correctly aligned".

– The size of δ is limited by many factors. Examples of which include $\delta < \delta_0$ where δ_0 is as in Definition 1.1, a bound used in distortion (Lemma 5.1), the Taylor formula estimate at "turns" (Proposition 5.3), $\delta << \varepsilon_0$, and some purely numerical inequalities (e.g. if $\delta = e^{-\mu}$, then $\frac{1}{\mu^2} << e^{-\mu}$).

– Chosen last are a and b. It is best to think of a and b as very small numbers that we may need to decrease a finite number of times as we go along.
  - The smaller a is, the longer $f^n(\hat{x}), \hat{x} \in C$, can be kept out of C_{δ_0}.
  - The smaller b is, the more closely T mimics $F^{\#}$.

– There is an important constant defined the same time as b, namely $\theta := \frac{K}{\log \frac{1}{b}}$ where K is chosen so that $b^\theta = \|DT\|^{-20}$. With this choice of θ, critical orbits emanating from the same $B^{(k)}$ can be viewed as a single orbit for $k\theta^{-1}$ iterates. We may, therefore, regard the number of critical orbits (or "critical blobs") present at time N as $\leq K^{\theta N}$ for some K depending on f_0.

When referring to $\mathcal{G}_N$ in the future, it will be understood that the arguments above are implicit. In particular, $\mathcal{G}_0$ is the set of maps T satisfying the conditions at the beginning of Part II with regard to some fixed f_0, K_0, a and b. Constants (such as K_1) not on this list are regarded as local in context; they must be specified each time they are used. Finally, we emphasize that the generic constant K that appears in many of our results is allowed only to depend on f_0, K_0 and m provided that the other constants are appropriately small.

6.2 Three stages of evolution

Our construction of $\mathcal{G}_N$ comes in three distinct stages: For $N \leq \theta^{-1}$, the situation is, in many ways, not far from that in 1D. This part is simple and is disposed of immediately in the next paragraph. At time $N = \theta^{-1}$, certain *local* complexities of higher dimensional maps begin to develop, "turns" play a more prominent role, and the definition of $\mathcal{G}_N$ becomes necessarily inductive. We have been building up the dynamical picture for this part in Sections 4 and 5 and will complete its construction in Sects. 6.3 and 6.4. At $N = \theta^{-2}$, the *global* structure of T begins to depart from those of 1D maps. New ideas are needed; they are discussed in Sections 7 and 8.

Getting started: the first θ^{-1} steps

Let $T \in \mathcal{G}_0$, and assume the leaves of $\mathcal{F}_1$ are parallel to the x-axis. Let $C = \{\hat{x}_1, \cdots, \hat{x}_q\}$ be the critical set of f. Then near each $\hat{x}_i$, there exists $\tilde{x}_i$ such that $e_1(S) = \partial_x$ at $(\tilde{x}_i, 0)$ where $S = S(\partial_x, \mathbf{v})$. A simple computation gives $|\tilde{x}_i - \hat{x}_i| < Kb$. Let $\Gamma_1 = \{(\tilde{x}_1, 0), \cdots, (\tilde{x}_q, 0)\}$. These are the only critical points for the first θ^{-1} iterates. Components of $\mathcal{C}^{(1)}$ centered at these points are constructed as required in (A1).

Before proceeding further, we observe that if γ_0 is a $C^2(b)$-segment with the property that $\gamma_i := T^i(\gamma_0)$ does not meet $B^{(1)}$ for all $i < n$, then the curves γ_i are roughly horizontal for all $i \leq n$. This follows immediately from the fact that for z with $d_{\mathcal{C}}(z) > b^{\frac{1}{5}}$ and $u \in X_z$ with $s(u) < b^{\frac{3}{4}}$, $s(DT_z(u)) < b^{\frac{3}{4}}$ (see Sect. 3.5).

For $N = 1, 2, \cdots, \theta^{-1}$, let

$$\mathcal{G}_N \ = \ \{T \in \mathcal{G}_0 \mid \ \text{(A2) and (A4) hold for all } z_0 \in \Gamma_1 \text{ and } i \leq N\}.$$

(A2) and (A4) are, as noted earlier, analogs of (G1) and (G2) in Sect. 2.2.

We claim that for $T \in \mathcal{G}_N$, (A3) and (A5) are satisfied automatically. (A3) is easily verified since all b-horizontal vectors are correctly aligned at $d_{\mathcal{C}} > e^{-\alpha\theta^{-1}}$ ($b^\theta < e^{-\alpha}$ by definition, so $b << e^{-\alpha\theta^{-1}}$). (A5) follows from 1D estimates: Let γ_0 be the curve joining $\xi_0 \in Q^{(1)} \setminus B^{(1)}$ to $B^{(1)}$ in (A5). Then during its bound period, all tangent vectors to γ_i are roughly horizontal as explained above. An argument entirely parallel to that in Appendix A.1 proves (A5)(i)-(iii).

Inductive scheme for going from $N = \theta^{-1}$ to $N = \theta^{-2}$

Beyond $N = \theta^{-1}$, more critical points are needed as orbits emanating from $B^{(1)}$ begin to diverge. To help describe the structures needed for the identification of new critical points, we have introduced a set of assumptions, namely (A1)–(A5). In Sect. 6.3, we will add another one, (A6), which is also trivially satisfied up to time θ^{-1}. Let

$$\mathcal{G}_N \ := \ \{T \in \mathcal{G}_0 \mid \text{(A1)(N)} - \text{(A6)(N) hold } \}, \quad N \leq \theta^{-2}.$$

Observe that this definition is consistent with the one defined earlier for $N \leq \theta^{-1}$. The goal of Sects. 6.3 and 6.4 is to prove the following:

(*) Let $\theta^{-1} < N < \frac{1}{\alpha^*}N \leq \theta^{-2}$. We assume $T \in \mathcal{G}_N$, and prove that if T satisfies (A2) and (A4) up to time $\frac{1}{\alpha^*}N$, then it is in $\mathcal{G}_{\frac{1}{\alpha^*}N}$.

The time step of the construction above is determined by the fact that at times $\leq \frac{1}{\alpha^*}N$, the lengths of the bound periods are $\leq N$. This ratio is noted in the paragraph on "three important time scales" in Sect. 4.1.

Why stop at $N = \theta^{-2}$?

We emphasize that the material in this section is for iterates $N \leq \theta^{-2}$. The reason for this time restriction is that as mentioned above, the sets $T^k B^{(1)}$ begin to get "large" at $k = \theta^{-1}$, affecting the geometry of the critical regions. (A1) and (A6), which we will introduce shortly, cannot be sustained as formulated.

Notation: In this section and the next, we will be working with the foliations $\mathcal{F}_k$. Given that we have defined $\mathcal{F}_1$ to be the initial foliation on R_1, it is advantageous in discussions involving $\mathcal{F}_k$ to let ξ_1 denote arbitrary points in R_1 and τ_1 unit tangent vectors to the leaves of $\mathcal{F}_1$. This convention (instead of the usual (ξ_0, τ_0)) leads to more pleasing notation such as $\xi_k \in R_k$ and τ_k as tangent vectors to the leaves of $\mathcal{F}_k$.

6.3 Controlling $\mathcal{F}_k$, and pushing forward (A1) and (A6)

One way to gain a better grip on the geometry of R_k is to control (ξ_i, τ_i) for $\xi_1 \in R_1$.

Rules for setting control

(1) We stop controlling (ξ_i, τ_i) once ξ_i enters $B^{(i)}$; this is compatible with the idea that $T^k B^{(i)}, k = 1, 2, \cdots, i\theta^{-1}$, is to be seen as the orbit of a single point.

(2) In our inductive scheme to be detailed shortly, the control of (ξ_i, τ_i) proceeds in parallel with the construction of Γ_i. For this reason, we will take $\phi(\xi_i) \in \Gamma_i$.

(3) As explained in Sect. 5.2, it suffices to set control at free returns. Let i be a free return. Then $\phi(\xi_i)$ is chosen as follows: If there exists $j < i$ such that $\xi_i \in \mathcal{C}^{(j)} \setminus \mathcal{C}^{(j+1)}$, then we let $\phi(\xi_i) = z_0^*(Q^{(j)})$ where $Q^{(j)} = Q^{(j)}(\xi_i)$. If $\xi_i \in Q^{(i)}$, we have no choice (in view of (2)) but to let $\phi(\xi_i) = z_0^*(Q^{(i)})$.

To the five assumptions (A1)–(A5) in Sect. 4.1, we now add another one. We say the **foliation** $\mathcal{F}_{k+1}$ is **controlled** on R_{k+1} by Γ_k if for all $\xi_1 \in R_1$ and $i \le k$, (ξ_i, τ_i) is controlled by Γ_k provided $\xi_i \notin B^{(i)}$ for all $i \le k$. (The indices in the last sentence are intended as written: we say $\mathcal{F}_{k+1}$ is controlled because control of (ξ_i, τ_i) for $i \le k$ leads to geometric knowledge of the leaves of $\mathcal{F}_{k+1}$.)

(A6)(N) *For all $k \le \theta N$, $\mathcal{F}_{k+1}$ is controlled on R_{k+1} by Γ_k.*

At this point we would like to assert that (A1)($\frac{1}{\alpha^*}N$) and (A6)($\frac{1}{\alpha^*}N$) hold for $T \in \mathcal{G}_N$. A proof would involve simultaneously constructing $\mathcal{C}^{(k)}$ and Γ_k, and using Γ_k to control $\mathcal{F}_{k+1}$. What prevents us from making a clean statement to this effect at this time is that without having first assumed or proved (A2)($k\theta^{-1}$)–(A5)($k\theta^{-1}$) for $k > \theta N$, we cannot, in principle, conclude that orbits controlled by Γ_k have the properties in Sect. 5.3.

We examine the situation more closely:

Assume T satisfies (A1)(N)–(A6)(N). Fix $\theta N < i \le \frac{1}{\alpha^*}\theta N$, and assume (A1)($i\theta^{-1}$) holds. Let $\xi_i \in \mathcal{C}^{(1)}$ be an arbitrary point. We define $\phi(\xi_i)$ as in (3) above and *assume* that τ_i is correctly aligned (with respect to $\mathcal{F}_j$ where j is the generation of $\phi(\xi_i)$). The discussion below pertains only to time $\le \frac{1}{\alpha^*}\theta N$.

Case 1. $j \le \theta N$. In this case, (ξ_i, τ_i) is controlled by $\Gamma_{\theta N}$ for the next $\min(p, \frac{1}{\alpha^*}\theta N - i)$ iterates where p is the bound period between ξ_i and $\phi(\xi_i)$.

Case 2. $j > \theta N$, and $\xi_i \notin B^{(\theta N)}$. The conclusion is as in Case 1. The orbit of $\hat{z}_0 := \phi(\xi_i)$ and that of $z_0 = z_0^*(B^{(\theta N)}(\phi(\xi_i)))$ remain extremely close during the period in question (more precisely, $|\hat{z}_k - z_k| < \|DT\|^k b^{\frac{\theta N}{5}} << e^{-\beta k}$), and it makes no difference whether we view ξ_i as bound to $\hat{z}_0$ or to z_0.

Case 3. $j > \theta N$ and $\xi_i \in B^{(\theta N)}$. The estimate in the last paragraph shows that ξ_i is bound to $\phi(\xi_i)$ – and to $z_0^*(B^{(\theta N)}(\phi(\xi_i))$ – through time $\frac{1}{\alpha^*}\theta N$. From Proposition 5.1, we know that e_ℓ is well defined on all of $B^{(\theta N)}$ for all $\ell \le N$, and by our correct alignment assumption together with Lemma 4.7, τ_i splits correctly. The evolution of the $\mathbf{v}$-component at ξ_i can then be compared to that at $z_0^*(B^{(\theta N)})$ using Proposition 5.1 and Lemma 5.3.

The discussion above tells us that in the control $(\xi_1, \tau_1), (\xi_2, \tau_2), \cdots$ up to time $\frac{1}{\alpha^*}\theta N$, the only role played by $\Gamma_k \setminus \Gamma_{\theta N}$ and $\mathcal{F}_k$ for $k > \theta N$ is to determine the correctness of alignment and subsequent splitting period at free returns. The rest of the control is really provided by $\Gamma_{\theta N}$. We have argued that Lemmas 5.2–5.6 apply up to time $\frac{1}{\alpha^*}\theta N$. Nevertheless, to distinguish between the present situation and that after we have conferred (A2)–(A5) upon Γ_k, we will say, if correct alignment holds for all $i \leq k$, that the sequence $(\xi_1, \tau_1), \cdots, (\xi_k, \tau_k)$ is *provisionally controlled* by Γ_k.

We now state the main result of this subsection.

Proposition 6.1 *Let $\theta^{-1} \leq N < \frac{1}{\alpha^*} N \leq \theta^{-2}$, and assume T satisfies (A1)(N)–(A6)(N). Then for $\theta N < k \leq \frac{1}{\alpha^*}\theta N$:*

(a)$_k$ $\mathcal{C}^{(k)}$ and Γ_k with the properties in (A1) can be constructed;

(b)$_k$ if $\xi_1 \in R_1$ is such that $\xi_i \notin B^{(i)}$ for all $i \leq k$, the sequence $(\xi_1, \tau_1), \cdots (\xi_k, \tau_k)$ is provisionally controlled by Γ_k.

Proof: We assume (a)$_i$ and (b)$_i$ for all $i < k$.

Proof of (a)$_k$: Noting that it makes sense to speak about those segments of $\mathcal{F}_k$-leaves that are provisionally controlled as being in a bound or free state, we begin with the following result of independent interest:

Lemma 6.1 *Let γ be a leaf of $\mathcal{F}_k$. If every $\xi_k \in \gamma$ is free, then γ is a $C^2(b)$-curve.*

Proof: That τ_k is b-horizontal follows from Corollary 4.1. As for curvature, we appeal to Lemma 3.3 after using Lemma 5.5 to establish that $|\tau_k| \geq \delta K^{-K\theta(k-i)}|\tau_i|$ for all $i < k$. $\square$

Let γ be a leaf segment of $\mathcal{F}_k$ meeting $\mathcal{C}^{(k-1)}$. We claim that it is contained in a maximal free segment that traverses the entire length of $Q^{(k-1)}$, extending as a $C^2(b)$-curve by $> \frac{1}{2}e^{-\alpha k}$ on both sides. To see this, let $\xi_k \in R_k$ be such that $d_{\mathcal{C}}(\xi_k) < \frac{1}{2}e^{-\alpha k}$, and suppose it is *not* free. Then there are only two possibilities: (1) For some $i < k$, $\xi_i \in B^{(i)}$ and we stopped controlling its orbit, or (2) ξ_i is controlled for all $i < k$, and ξ_k is in a bound period initiated at some time $i < k$. (1) is not feasible, for if we let $z_0 = z_0^*(B^{(i)})$, then $d_{\mathcal{C}}(z_{k-i}) > e^{-\alpha(k-i)}$, and $\text{diam}(T^{k-i}B^{(i)}) << e^{-\alpha(k-i)}$, contradicting $d_{\mathcal{C}}(\xi_k) < \frac{1}{2}e^{-\alpha k}$. (2) is also impossible, for if we let $z_0 = \phi(\xi_i)$, then $d_{\mathcal{C}}(z_{k-i}) > e^{-\alpha(k-i)}$ while $|\xi_k - z_{k-i}| < e^{-\beta(k-i)} << e^{-\alpha(k-i)}$.

We have proved that $R_k \cap Q^{(k-1)}$, if non-empty, is the union of a collection of horizontal sections $\{H\}$. In each H, we arbitrarily pick an $\mathcal{F}_k$-leaf γ. The critical point $z_0^*(Q^{(k-1)})$ constructed in step (a)$_{k-1}$ induces a critical point of order $k-1$ on γ (Lemma 4.3 and 3.8). By Lemma 3.9, this critical point can be upgraded to one of order k. We make it an element of Γ_k, and construct a $Q^{(k)}$ of length $\min(2\delta, e^{-\lambda k})$ centered at it. Doing this for every horizontal section H that passes through every $Q^{(k-1)}$ completes the construction of $\mathcal{C}^{(k)}$ and Γ_k.

It follows directly from the next lemma that the sectional diameter of $Q^{(k)}$ is $< b^{\frac{k}{2}}$.

Lemma 6.2 *Every $\xi_k \in Q^{(k)}$ is contained in a codimension one manifold W with the property that*

(i) W meets every connected component of $\mathcal{F}_k$-leaf in $Q^{(k)}$ in exactly one point;

(ii) for all $\xi_1, \xi_1' \in T^{-(k-1)}W$, $|\xi_i - \xi_i'| < b^{\frac{i}{2}}$ for all $i \leq k$.

Lemma 6.2 is proved in Appendix A.19.

Proof of $(b)_k$*:* As noted earlier, it suffices to consider the case where ξ_k is a free return, and it suffices to show correct alignment of τ_k at ξ_k. Let γ be the maximal free segment of $\mathcal{F}_k$-leaf containing ξ_k. Then the endpoints of γ are in bound state, and so are outside of $\mathcal{C}^{(k)}$. This leaves two possibilities for the relation between γ and $\mathcal{C}^{(k)}$.

Case 1. γ passes through the entire length of some $Q^{(k)}$. We consider $Q^{(k)} \subset Q^{(k-1)} \subset Q^{(k-2)} \subset \cdots$ until we reach the first $Q^{(j)}$ that contains ξ_k. Since $\xi_k \in (Q^{(j)} \setminus \mathcal{C}^{(j+1)})$, $d_{\mathcal{C}}(\xi_k) > e^{-\lambda(j+1)}$. We let $\hat{\gamma}$ be the $\mathcal{F}_j$ leaf through ξ_k, and apply Lemma 4.3 to obtain two points $\gamma(0)$ and $\hat{\gamma}(0)$ in γ and $\hat{\gamma}$ respectively with

(i) $|\xi_k - \gamma(0)| \approx |\xi_k - \hat{\gamma}(0)| \approx d_{\mathcal{C}}(\xi_k)$, and
(ii) $\angle(\hat{\gamma}'(0), \gamma'(0)) < Kb^{\frac{j}{4}}$.

Letting $\hat{\tau}_j$ and τ_k denote the tangents to $\hat{\gamma}$ and γ respectively at ξ_k and using the fact that γ and $\hat{\gamma}$ are $C^2(b)$-curves between the points in question, we have

$$\begin{aligned}\angle(\hat{\tau}_j, \tau_k) &\leq \angle(\hat{\tau}_j, \hat{\gamma}'(0)) + \angle(\hat{\gamma}'(0), \gamma'(0)) + \angle(\gamma'(0), \tau_k) \\ &< \frac{Kb}{\delta^3} d_{\mathcal{C}}(\xi_k) + Kb^{\frac{j}{4}} + \frac{Kb}{\delta^3} d_{\mathcal{C}}(\xi_k) \\ &< b^{\frac{1}{5}} d_{\mathcal{C}}(\xi_k).\end{aligned}$$

Case 2. γ does not meet $\mathcal{C}^{(k)}$. We first formally treat the geometry before making the required angle estimates.

Geometry: (i) Let j be the largest integer such that $\xi_k \in Q^{(j)}$, so that $\xi_k \in (H \setminus Q^{(j+1)})$ where H is a component of $R_{j+1} \cap Q^{(j)}$. Suppose for definiteness that ξ_k lies in the right chamber of $H \setminus Q^{(j+1)}$. We move along γ to the left until we reach either ξ, the left end point of γ, or the right boundary of $Q^{(j+1)}$, whichever happens first. Once ξ is reached, we stop. Otherwise we continue moving through $Q^{(j+1)}$ until we reach either ξ or the right boundary of $Q^{(j+2)}$. (We have used implicitly the fact that γ, which is a leaf of $\mathcal{F}_k$, does not meet ∂R_i for $i < k$.) By assumption, ξ is reached before we arrive at $Q^{(k)}$, so that $\xi \in Q^{(j')} \setminus \mathcal{C}^{(j'+1)}$ for some j' with $k > j' \geq j$.

(ii) We note that ξ can also be regarded as in bound state, and argue now that $\phi(\xi)$ is to the left of ξ. More precisely, we write $\xi = \eta_k$, let $\eta_i, i < k$, be the last free return, and let $\phi(\eta_i) = \hat{z}_0$. Recalling the definition of $\phi(\cdot)$ for critical orbits (Sects. 4.1 and 4.3A), we deduce that $\phi(\xi) = \phi(\hat{z}_{k-i})$ is of generation j'' for some $j'' \leq j'$ (it can be considerably smaller), and that both ξ and $\hat{z}_{k-i}$ are in $Q^{(j'')} \setminus B^{(j'')}$. To see that ξ is in the right chamber of $Q^{(j'')} \setminus B^{(j'')}$, we interpolate between $Q^{(j')} \subset \cdots \subset Q^{(j'')}$, noting that the right chamber of each $Q^{(i)}$ does not meet the left chamber of $Q^{(i-1)}$.

Angles: Let $\hat{\gamma}$ be the leaf of $\mathcal{F}_j$ through ξ_k and $\tilde{\gamma}$ the leaf of $\mathcal{F}_{j''}$ through ξ. We will use the following notation: τ_{k,ξ_k} and $\tau_{k,\xi}$ are tangents to γ at ξ_k and ξ respectively; $\hat{\tau}_{j,\xi_k}$ is tangent to $\hat{\gamma}$ at ξ_k, and $\tilde{\tau}_{j'',\xi}$ is tangent to $\tilde{\gamma}$ at ξ. Then

$$\angle(\tau_{k,\xi_k}, \hat{\tau}_{j,\xi_k}) \leq \angle(\tau_{k,\xi_k}, \tau_{k,\xi}) + \angle(\tau_{k,\xi}, \tilde{\tau}_{j'',\xi}) + \angle(\tilde{\tau}_{j'',\xi}, \hat{\tau}_{j,\xi_k}).$$

The terms above are estimated by

(i) $\angle(\tau_{k,\xi_k}, \tau_{k,\xi}) < \frac{Kb}{\delta^3}|\xi - \xi_k|$ since γ is free and hence $C^2(b)$;
(ii) $\angle(\tau_{k,\xi}, \tilde{\tau}_{j'',\xi}) < 3\varepsilon_0|\xi - \phi(\hat{z}_{k-i})|$ since $\xi = \eta_k$ is a bound return (Lemma 5.3);
(iii) $\angle(\tilde{\tau}_{j'',\xi}, \hat{\tau}_{j,\xi_k}) < b^{\frac{1}{4}\min(j,j")} + \frac{Kb}{\delta^3}|\xi - \xi_k|$ from Lemma 4.2.

Also, we have argued that $\phi(\hat{z}_{k-i})$ is to the left of ξ, so $|\xi - \xi_k|$, $|\xi - \phi(\hat{z}_{k-i})| < d_{\mathcal{C}}(\xi_k)$. These inequalities together with $d_{\mathcal{C}}(\xi_k) >> b^{\frac{j}{4}}$ and $d_{\mathcal{C}}(\xi_k) > d_{\mathcal{C}}(\xi) > b^{\frac{j''}{5}}$ give $\angle(\tau_{k,\xi_k}, \hat{\tau}_{j,\xi_k}) < 4\varepsilon_0 d_{\mathcal{C}}(\xi_k)$. □

6.4 What it takes to go from $\mathcal{G}_N$ to $\mathcal{G}_{\frac{1}{\alpha^*}N}$

In this subsection we fix N with $\theta^{-1} \le N < \frac{1}{\alpha^*}N \le \theta^{-2}$, and assume that $T \in \mathcal{G}_N$, i.e. (A1)(N)–(A6)(N) hold for T. It is proved in Proposition 6.1 that without further assumptions, (A1)($\frac{1}{\alpha^*}N$) holds automatically. The purpose of this subsection is to determine what constraints we need to impose on T to put it in $\mathcal{G}_{\frac{1}{\alpha^*}N}$.

(A2): Rate of approach to critical set

By assumption, all $z_0 \in \Gamma_{\theta N}$ obey (A2). At issue is whether or not $z_0 \in \Gamma_{\frac{1}{\alpha^*}\theta N} \setminus \Gamma_{\theta N}$ obeys (A2) as stated in Sect. 4.1. We distinguish between the two time intervals $[1, N]$ and $[N+1, \frac{1}{\alpha^*}N]$: On $[1, N]$, each $z_0 \in \Gamma_{\frac{1}{\alpha^*}\theta N} \setminus \Gamma_{\theta N}$ follows closely a critical point of generation θN (Corollary 5.1). The two orbits do differ by a little, however, so it is possible for z_i to violate slightly the condition in (A2). On $[N+1, \frac{1}{\alpha^*}N]$, there is no reason why (A2) is respected by $z_0 \in \Gamma_{\frac{1}{\alpha^*}\theta N} \setminus \Gamma_{\theta N}$. We conclude that (A2)($\frac{1}{\alpha^*}N$) is a new condition that must be imposed on T if it is to belong in $\mathcal{G}_{\frac{1}{\alpha^*}N}$.

(A3): Correct alignment

We will prove that (A3), in fact, comes for free. The mechanisms for ensuring correct alignment of w_i^* at free returns and at bound returns are entirely different. At free returns, this comes from *geometry*, from the "rank one" character of T in particular. At bound returns, it comes from *copying*. We emphasize that we do not deduce (A3)($\frac{1}{\alpha^*}N$) directly from (A3)(N). We prove it from scratch, in a sense, keeping track of the increase in error each time the picture is copied.

Proposition 6.2 *Let $T \in \mathcal{G}_N$. We assume $\Gamma_{\frac{1}{\alpha^*}\theta N}$ is constructed, and fix $z_0 \in \Gamma_{\frac{1}{\alpha^*}\theta N}$. We assume the condition in (A2) is imposed on z_0 up to time $\frac{1}{\alpha^*}N$.*[7] *Then $(z_0, \mathbf{v})$ is controlled by $\Gamma_{\theta N}$ up to time $\frac{1}{\alpha^*}N$. In fact, we have the following stronger results:*
(i) If z_i is a free *return, then w_i is aligned correctly with error $< d_{\mathcal{C}}(z_i) << \varepsilon_0$;*
(ii) If z_i is a bound *return, then w_i^* is aligned correctly with $< \varepsilon_0$-error.*

Proof: To establish control for the orbit of (z_0, w_0) with $w_0 = \mathbf{v}$, we define $\phi(z_i)$ for $i \le \frac{1}{\alpha^*}N$ according to the rule in Sect. 4.1, i.e. $\phi(z_i)$ is the critical point of highest generation $j \le \alpha^*\theta i \le \theta N$ with the property that $z_i \in \mathcal{C}^{(j)}$. Note that this implies $\phi(z_i) \in \Gamma_{\theta N}$.[8] With (A2) imposed on z_0, Lemma 4.4 shows that the conditions in Definition 5.1 are met. Correct splitting is proved inductively as follows. Assume that (z_0, w_0) is controlled for $k-1$ iterates by $\Gamma_{\theta N}$ for some $k \le \frac{1}{\alpha^*}N$.

(i) z_k is a free return Let $j+1$ be the generation of $\phi(z_k)$, and let $\tau_{z_k,j+1}$ be tangent to $\mathcal{F}_{j+1}$ at z_k. We need to show $\angle(w_k, \tau_{z_k,j+1}) << \varepsilon_0 d_{\mathcal{C}}(z_k)$. Let $\xi_1 = T^{-j}z_k$, i.e. $z_k = \xi_{j+1}$. From (A6)(N), we know that (ξ_1, τ_1) is controlled for j iterates unless $\xi_n \in B^{(n)}$ for some $n \le j$, which is impossible because that would contradict our assumption that z_k is free. Observe that

$$\begin{aligned}\angle(w_k, \tau_{z_k,j+1}) &= \angle(DT^j_{\xi_1} w_{k-j}, DT^j_{\xi_1}\tau_1) \le \frac{|DT^j_{\xi_1} w_{k-j} \wedge DT^j_{\xi_1}\tau_1|}{|w_k||\tau_{j+1}|} \\ &\le (Kb)^j \frac{|w_{k-j} \wedge \tau_1|}{|w_k||\tau_{j+1}|} \le (Kb)^j \frac{|w_{k-j}|}{|w_k||\tau_{j+1}|}.\end{aligned}$$

[7] By this, we mean the condition $d_{\mathcal{C}}(z_i) > \min(\delta, e^{-\alpha i})$ is assumed but only for the orbit of z_0, i.e. no assumptions are made on the behavior of other critical orbits beyond time N.

[8] This is the reason why we require $j \le \alpha^*\theta i$ in our definition of $\phi(z_i)$ in Sect. 4.1.

The second inequality above comes from Sublemma A.4.2 in Appendix A.4; the rest are straightforward. To estimate the quantity in the final bound, we claim that $\frac{|w_{k-j}|}{|w_k|} < Ke^{-\lambda'' j}$ where λ'' is slightly smaller than $\frac{1}{3}\lambda$. This is because (z_0, w_0) is, by inductive assumption, controlled by $\Gamma_{\theta N}$ for k iterates, and z_k being a free return, Lemma 5.5 applies. Next we claim that $\frac{1}{|\tau_{j+1}|} < Ke^{-\lambda'' j}$. This is again a consequence of Lemma 5.5, after we establish the following: With $j \leq \theta N$, (ξ_1, τ_1) is a controlled pair by (A6)(N), and ξ_{j+1} is a free return because the bound-free structure on the orbit segment $\xi_1, \xi_2, \cdots, \xi_{j+1}$ can be taken to be identical to that of $z_{k-j}, z_{k-j+1}, \cdots, z_k$ except that bound periods for z_i initiated before time $k-j$ do not count for ξ_1. Thus $\angle(w_k, \tau_{z_k, j+1}) < (Kb)^j$.

Correct alignment at z_k is now straightforward: If $j+1 < \theta N$, then $\angle(w_k, \tau_{z_k, j+1}) < (Kb)^j << e^{-2\lambda j} \approx (d_{\mathcal{C}}(z_k))^2$. In general, $d_{\mathcal{C}}(z_k) > e^{-\alpha k} \geq e^{-\alpha(\frac{1}{\alpha^*}N)}$, so that if $j+1 = \theta N$, then

$$\angle(w_k, \tau_{z_k, j+1}) < (Kb)^j = (Kb)^{\theta N - 1} < e^{-\frac{\lambda}{3}N} = e^{-\alpha(\frac{2}{\alpha^*}N)} \leq (d_{\mathcal{C}}(z_k))^2.$$

This completes the proof of the free return case.

(ii) z_k is a bound return Consider first the following scenario:

Suppose $z_j, j < k$, is a free return, and the bound period initiated at that time extends beyond time k. Let $\phi(z_j) = \hat{z}_0$, and assume $\hat{z}_{k-j}$ is a free return.

We estimate the error in alignment at z_k as follows: Let g and $\hat{g}$ be the generations of $\phi(z_k)$ and $\phi(\hat{z}_{k-j})$. We claim that $g \geq \hat{g} - 1$. This is because if $\hat{z}_{k-j} \in Q^{(n)}$, then z_k is also in $Q^{(n)}$ – or it is just outside, in which case it is in the $Q^{(n-1)}$ containing $Q^{(n)}$ (see Lemma 4.5). Now in the definition of $\phi(\cdot)$, there is an upper bound on the generation of the guiding critical orbit. Since $k > k-j$, a more stringent upper bound is imposed on $\hat{z}_{k-j}$ than on z_k. Hence the assertion.

In the discussion to follow, we let $\tau_{z,n}$ denote the tangent to the $\mathcal{F}_n$-leaf at z. The angle to be estimated, $\angle(w_k^*(z_0), \tau_{z_k, g})$, is bounded above by

$$\angle(w_k^*(z_0), \tau_{z_k, g}) \leq \angle(w_k^*(z_0), w_{k-j}(\hat{z}_0)) + \angle(w_{k-j}(\hat{z}_0), \tau_{\hat{z}_{k-j}, \hat{g}}) + \angle(\tau_{\hat{z}_{k-j}, \hat{g}}, \tau_{z_k, g}).$$

From part (i), we have that the second term on the right is $< (d_{\mathcal{C}}(z_k))^2$: $\hat{z}_{k-j}$ is free, and $d_{\mathcal{C}}(z_k) \approx d_{\mathcal{C}}(\hat{z}_{k-j})$ from Lemma 4.5. The third term is $< Kb^{\frac{1}{4}(\hat{g}-1)} + Kb\delta^{-3}e^{-\beta(k-j)}$ from Lemma 4.2. To estimate the first term we write $w_j(z_0) = Ae_{k-j} + B\mathbf{v}$ to obtain

$$\angle(w_k^*(z_0), w_{k-j}(\hat{z}_0)) < \angle(w_{k-j}(z_j), w_{k-j}(\hat{z}_0)) + (Kb)^{k-j} < \frac{1}{2}e^{-\frac{1}{2}\beta(k-j)} + (Kb)^{k-j};$$

the second inequality is from Proposition 5.1 and Lemma 5.1. Plugging $d_{\mathcal{C}}(z_k) > b^{\frac{1}{5}(\hat{g}-1)}$ (Lemma 4.4) and $d_{\mathcal{C}}(z_k) > e^{-\alpha(k-j)}$ into the three estimates above, we obtain

$$\begin{aligned} \angle(w_k^*(z_0), \tau_{z_k, g}) &< d_{\mathcal{C}}(z_k) \, \{d_{\mathcal{C}}(z_k) + Ke^{(\frac{1}{2}\beta - \alpha)(k-j)} + (Kb)^{k-j} + Kb^{\frac{1}{20}(\hat{g}-1)}\} \\ &< \varepsilon_0 d_{\mathcal{C}}(z_k). \end{aligned}$$

In general, there exist $j_1 < j_2 < \cdots < j_n < k$ and $\hat{z}_0^{(1)}, \cdots, \hat{z}_0^{(n)} \in \Gamma_{\theta N}$ such that

– z_{j_1} is the last free return before time k, with $\phi(z_{j_1}) = \hat{z}_0^{(1)}$;
– $\hat{z}_{k-j_1}^{(1)}$ is not a free return; its last free return is at time j_2, with $\phi(\hat{z}_{j_2-j_1}^{(1)}) = \hat{z}_0^{(2)}$;
– $\hat{z}_{k-j_2}^{(2)}$ is not a free return; its last free return is at time j_3, and so on, until
– finally, $\hat{z}_{k-j_n}^{(n)}$ is a free return.

Considerations similar to those above show that as we go through the different layers of bindings, the errors in alignment form a geometric series which add up to $\leq e^{-\frac{1}{2}\beta(k-j_n)} + (Kb)^{k-j_n}$. □

(A4): Growth of $|w_i^*|$

Corollary 6.1 *Under the hypothesis of Proposition 6.2, we have, for the critical orbit z_0 in question,*

$$\frac{|w_i^*(z_0)|}{|w_N^*(z_0)|} > K^{-1}e^{(\frac{1}{3}\lambda-2\alpha)(i-N)}, \qquad N < i \leq \frac{1}{\alpha^*}N.$$

This corollary, which follows from the control of $(z_0, \mathbf{v})$ proved in Proposition 6.2, the condition in (A2), and Lemma 5.4, is in the direction of maintaining (A4)($\frac{1}{\alpha^*}N$) but the gain in Lyapunov exponent is only about $\frac{1}{3}$ of what is needed for that purpose. Indeed, (A4) is not a self-perpetuating condition: the exponent in Corollary 6.1, when copied again in future inductive steps, may lead to a downward spiral in the Lyapunov exponent along critical orbits.

We conclude that (A4)($\frac{1}{\alpha^*}N$) must be imposed (by external means) to ensure that $T \in \mathcal{G}_{\frac{1}{\alpha^*}N}$.

(A5): Quadratic turns, lengths of bound periods and derivative recovery

Proposition 6.3 *Let $T \in \mathcal{G}_N$ be such that (A2)($\frac{1}{\alpha^*}N$) and (A4)($\frac{1}{\alpha^*}N$) hold. Then (A5)($\frac{1}{\alpha^*}N$) holds automatically.*

Proof: Let $z_0 \in \Gamma_{\frac{1}{\alpha^*}\theta N}$ be fixed. Suppose $z_0 = z_0^*(Q^{(k)})$, and let $\xi_0 \in Q^{(k)} \setminus B^{(k)}$. For the definitions of and relation between $\hat{p}(z_0, \xi_0)$ and $p(z_0, \xi_0)$, see Sect. 4.3B.

Proof of (A5)(iii): Let γ be the $\mathcal{F}_k$-leaf containing ξ_0 in $Q^{(k)}$. Then there exists $\tilde{z}_0 \in \gamma \cap B^{(k)}$ such that $\tilde{z}_0$ is a critical point of order $p = \min\{p(z_0, \xi_0), k\theta^{-1}\}$ on γ. For all practical purposes, z_0 and $\tilde{z}_0$ are indistinguishable for p iterates, so we may regard $\tilde{z}_0$ as satisfying the hypotheses of Proposition 5.3 (which z_0 has been shown to satisfy). Proposition 5.3 then gives the desired result (with $\tilde{z}_0$ instead of z_0).

Proof of (A5)(i): For the lower bound, we have, for all $j \leq \frac{h}{3\log\|DT\|}$,

$$|\xi_j - z_j| < \|DT\|^j |\xi_0 - z_0| < e^{-\frac{2h}{3}} << e^{-\hat{\beta}\frac{h}{3\log\|DT\|}},$$

proving $\hat{p}(z_0, \xi_0) \geq \frac{h}{3\log\|DT\|}$.

For the upper bound, by Proposition 5.3, $p = p(z_0, \xi_0)$ is the smallest integer i such that $|w_i(0)| \cdot |z_0 - \xi_0|^2 > K_1^{-1}e^{-\beta i}$. We claim that $p \leq \frac{3h}{\lambda}$, for

$$|w_{\frac{3h}{\lambda}}(z_0)| \cdot |z_0 - \xi_0|^2 > K^{-\varepsilon\frac{3h}{\lambda}} |w^*_{\frac{3h}{\lambda}}(z_0)| \cdot |z_0 - \xi_0|^2 > K^{-\varepsilon\frac{3h}{\lambda}} e^{\lambda\cdot\frac{3h}{\lambda}} e^{-2h} > 1.$$

Lemma 5.2 is used in the first inequality above.

Proof of (A5)(ii): First we consider $\hat{p} = \hat{p}(z_0, \xi_0)$, for which we have $|w_{\hat{p}}(\tilde{z}_0)| \approx |w_{\hat{p}}(z_0)|$ (Corollary 5.1), and

$$\begin{aligned} |z_0 - \xi_0||w_{\hat{p}}(z_0)| &= (|z_0 - \xi_0||w_{\hat{p}}(z_0)|^{\frac{1}{2}}) \cdot |w_{\hat{p}}(z_0)|^{\frac{1}{2}} \\ &> K^{-1}|z_{\hat{p}} - \xi_{\hat{p}}| \cdot |w_{\hat{p}}(z_0)|^{\frac{1}{2}} \quad \text{by (A5)(iii)} \\ &> K^{-1}e^{-\frac{\hat{\beta}}{2}\hat{p}} \cdot e^{\frac{\lambda\hat{p}}{2}}. \end{aligned}$$

Now let $p \in [\hat{p}, \hat{p}(1 + \frac{9}{\lambda}\alpha)]$. From Lemmas 5.2 and 5.4, $|w_p(z_0)| \geq e^{-4\alpha p}|w_{\hat{p}}(z_0)|$, and so

$$|z_0 - \xi_0||w_p(z_0)| > K^{-1}e^{-4\alpha p}e^{-\frac{\hat{\beta}}{2}\hat{p}} \cdot e^{\frac{\lambda\hat{p}}{2}} > e^{\frac{1}{3}\lambda p}.$$

□

(A6): Control of foliations

With (A2) and (A4) assumed and (A3) and (A5) proved up to time $\frac{1}{\alpha^*}N$, the provisional control proved in Proposition 6.1 is, by definition, upgraded to control in the usual sense.

Summary: (*) in Sect. 6.2 is proved.

7 Global Geometry via Monotone Branches

The purpose of this section is to introduce the main geometric ideas needed to construct $T \in \mathcal{G}_n$ beyond $N = \theta^{-2}$, and to reformulate (A1) and (A6) to accommodate these new geometric structures.

7.1 Introduction

The idea of studying piecewise monotonic 1D maps via their monotone branches has been used many times. We attempt in this section to introduce a corresponding notion for T. For small n, it is easy to see that R_n is the union of sets that are tubular neighborhoods of 1D monotone curve segments. These should be, by any definition, monotone branches of T. For $n \leq \theta^{-1}$, we have seen that R_n is punctuated by (tiny) sections that are T^i-images of $B^{(n-i)}$, i.e. the "critical blobs" of Sect. 4.2. Intuitively, these sets are located at sharp "turns"; they divide R_n into connected components that are comparatively "straight". Leaving precise definitions for later, we think of these components as monotone branches of generation n.

The picture in the last paragraph cannot be maintained indefinitely, however, for it relies on the fact that critical blobs are very small compared to their distances to the critical set. As n increases, it is inevitable that the images of $B^{(k)}$ will grow large, making it impossible to keep them away from the critical regions. See Fig. 2. As explained in Sect. 6.2, the significance of time $N = \theta^{-2}$ is that at time θ^{-2}, the geometry of $R_{\theta^{-1}}$ becomes relevant, and θ^{-1} is the time beyond which we cannot guarantee the smallness of the images of $B^{(1)}$.

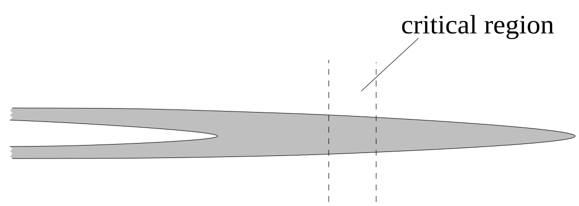


Fig. 2 The images of $B^{(k)}$ cannot avoid critical regions forever

To avoid dealing with the situation depicted in Fig. 2, we declared in Sect. 4.2 that $B^{(k)}$ ceases to be active after $k\theta^{-1}$ iterates. Once $B^{(k)}$ ceases to be active, we must "discontinue", i.e. stop considering, the monotone branches that end in $T^{k\theta^{-1}}B^{(k)}$. We do not wish to relinquish control completely of the region occupied by a discontinued branch, however, for it is likely to contain part of the attractor.

Central to our scheme is the idea of *branch replacement*. We will prove that all monotone branches that are discontinued can be systematically replaced by branches of higher generations, so that at every step n, there is a collection of "good" branches of generations $\sim n$ that together account for all parts of the attractor. This replacement procedure is discussed at the end of Section 8, after we make precise the notion of a monotone branch and integrate these new geometric ideas into the dynamical picture described in sections 4, 5 and 6.

A 2D version of monotone branches was introduced in [WY1]. They were not used, however, in the inductive construction of the dynamical picture.

7.2 Formal definitions and assumptions

The idea of monotone branches is inseparable from that of critical regions. Any definition necessarily assumes that certain relevant critical structures have been identified. Likewise, the identification of these critical structures relies on the idea of monotone branches. Definition 7.1 below is how we have elected to enter this inductive cycle.

We say $A_1 \cup A_2 \cup A_3$ are contiguous sections of R_k if (i) each A_i is a section of R_k, and (ii) if $\Phi_i : [-1,1] \times D_{m-1} \to R_k$ are the defining maps for A_i (see the definitions immediately preceding Part II), then for $i = 1, 2$, $\Phi_i(\{1\} \times D_{m-1}) = \Phi_{i+1}(\{-1\} \times D_{m-1})$.

Definition 7.1 *Let $T \in \mathcal{G}_0$ and $n \geq 2$. Suppose that for each $k < n$, a collection of sections $\{B^{(k)}\}$ of R_k has been identified. Then a* **monotone branch** *M* **of generation** *n for T is a section of R_n that is the union of three contiguous sections $E \cup M^\circ \cup E'$ with the following properties:*

(a) There exist $i, i' < n$ and $\hat{B}^{(n-i)}, \hat{B}^{(n-i')}$ such that

(i) $T^{-i}(E) = \hat{B}^{(n-i)}$ and $i \leq (n-i)\theta^{-1}$;

(ii) $T^{-i'}(E') = \hat{B}^{(n-i')}$ and $i' \leq (n-i')\theta^{-1}$;

(b) for all $i < n$, $T^{-i}(M^\circ) \cap B^{(n-i)} = \emptyset$ for all $B^{(n-i)}$.

E and E' are called the *ends* of the monotone branch M, and M° is its *main body*. We say the end E has reached the end of its period of activity if $i = (n-i)\theta^{-1}$. For $T \in \mathcal{G}_{\theta^{-2}}$ and $k \leq \theta^{-1}$, $\{B^{(k)}\}$ is as in Sect. 4.1. Thus we know from Sections 4–6 that monotone branches of generation $n \leq \theta^{-1} + 1$ are well defined. In this time range, every R_n is the union of a finite number of monotone branches, with adjacent branches overlapping in $T^i B^{(n-i)}$ for some $B^{(n-i)}$.

Remark. In the case that I is an interval, there are two special branches of generation n for all $n \geq 2$ with the property that one of their two ends is $T^{n-1}(V_i)$. Here V_1 and V_2 are the two "vertical" components in the boundary of $\hat{R}_1$; see Sect. 3.9.

Tree of monotone branches

Associated with $T \in \mathcal{G}_{\theta^{-2}}$ is a combinatorial object $\cup_{1 \leq k \leq \theta^{-1}} \mathcal{T}_k$ defined as follows: We declare R_1 to be the unique monotone branch of generation 1 (even though it has no ends), and let $\mathcal{T}_1 = \{R_1\}$. In general, $\mathcal{T}_k$ consists of a collection of monotone branches of generation k. Let $M \in \mathcal{T}_k$ for some $k < \theta^{-1}$. Here is how M *reproduces*: By construction, M either does not intersect any of the $B^{(k)}$, or it contains in its main body a finite number of them, say $B_1, B_2, \cdots, B_s$, in that order. In the first case, $T(M) \in \mathcal{T}_{k+1}$. In the second, $T(M)$ is the union of $s+1$ elements of $\mathcal{T}_{k+1}$, the main bodies of which connect $T(E)$ to $T(B_1)$, $T(B_i)$ to $T(B_{i+1})$, and $T(B_s)$ to $T(E')$. This construction defines a finite tree with θ^{-1} levels.

Suppose for $n \geq \theta^{-1}$ the tree $\cup_{1 \leq k \leq n} \mathcal{T}_k$ is defined, i.e. each $\mathcal{T}_k$ consists of a collection of monotone branches of generation k and $\mathcal{T}_k$ and $\mathcal{T}_{k+1}$ are related as above. We assume further that the critical regions $Q^{(n)}$ and $B^{(n)}$ have been identified. Then we may extend the tree to level $n+1$: Consider one $M \in \mathcal{T}_n$ at a time. First we check to see if either one of the two ends E and E' of M has reached the end of its period of activity. If so, the branch M is "discontinued", i.e. we do not iterate it further. If not, then M reproduces as in the last paragraph, and $\mathcal{T}_{n+1}$ consists of all the offsprings so obtained as M ranges over $\mathcal{T}_n$. (Note that no branch is discontinued for $n < \theta^{-1}$.)

Provided that the relevant monotone branches and $\{B^{(k)}\}$ are well defined and are related in the manner described above, one can extend $\mathcal{T}_n$ indefinitely and obtain, as $n \to \infty$, an infinite tree $\mathcal{T} := \cup_{k \geq 1} \mathcal{T}_k$.

We turn next to the inductive construction of $\mathcal{T}_k$ and $\mathcal{C}^{(k)}$. From the previous discussion, it is clear that the construction of these two objects must proceed hand in hand. Moreover, after time $N = \theta^{-2}$, due to the discontinuation of certain branches, the structure of critical regions becomes more complex, and (A1) and (A6) have to be modified accordingly. In (A1') and (A6') below, we try to give as complete a geometric description of these structures as possible, without seeking to present a minimal set of conditions.

(A1')(N) **Critical regions** *For $1 \leq k \leq \theta N$, there are sets $\mathcal{C}^{(k)}$ called* critical regions *with the following properties:*

(I) **Geometric structure** *$\mathcal{C}^{(1)}$ is as defined after Sect. 3.4. For $k \leq \theta N$, $\mathcal{C}^{(k)}$ has a finite number of connected components $\{Q^{(k)}\}$ each one of which is a horizontal section of R_k.*

(a) Relation among different $Q^{(k)}$

(i) *For $Q^{(k)}$ and $Q^{(k')}$ with $k < k'$, either $Q^{(k')} \subset Q^{(k)}$ or $Q^{(k)} \cap Q^{(k')} = \emptyset$. This defines a partial ordering on the set $\{Q^{(k)}, 1 \leq k \leq \theta N\}$ with $Q < Q'$ if and only if $Q' \subset Q$.*

(ii) *If $Q^{(k_1)} < \cdots < Q^{(k_n)}$ is a maximal chain,*[9] *then $k_{i+1} \leq k_i(1 + 2\theta)$.*

(b) Properties of individual $Q^{(k)}$

(i) *$Q^{(k)}$ has length $\min(2\delta, 2e^{-\lambda k})$ and cross-sectional diameter $< b^{\frac{k}{2}}$.*

(ii) *Exactly halfway between the two ends of $Q^{(k)}$ a point $z_0 = z_0^*(Q^{(k)})$ is singled out; z_0 is a critical point of order k with respect to the leaf of the foliation $\mathcal{F}_k$ containing it, and $B^{(k)} := \{z \in Q^{(k)} : |z - z_0|_h < b^{\frac{k}{5}}\}$.*

(II) **Construction and relation to $\mathcal{T}_k$** *Let $M \in \mathcal{T}_k$, $k \leq \theta N$, and let Q be a component of $\mathcal{C}^{(j)}, (1 + 2\theta)^{-1} k \leq j < k$. Then either $M \cap Q = \emptyset$ or it is the union of a finite number of horizontal sections each one of which extends $> \frac{1}{2}e^{-\alpha k}$ on both sides of Q. If $M \cap Q \neq \emptyset$, then each connected component H of $M \cap Q$ contains a unique $Q^{(k)}$, which is located in roughly the middle of H in terms of x-coordinate. All $Q^{(k)}$ are constructed this way.*

(A6')(N) **Monotone branches** *$\cup_{0 \leq k \leq \theta N} \mathcal{T}_{k+1}$ is defined with the following properties:*

(I) **Construction and relation to $\mathcal{C}^{(k)}$** *For each $M \in \mathcal{T}_k$, $1 \leq k \leq \theta N$, either M does not meet any $B^{(k)}$, or it contains in its main body a finite number of them, and it reproduces as described in the paragraph on "Tree of monotone branches".*

(II) **Dynamical control** *For $M \in \cup_{0 \leq k \leq \theta N} \mathcal{T}_{k+1}$, $\mathcal{F}_{k+1}$ is controlled on M° by Γ_k.*

(III) **Relation to R_k** *For $k \leq \theta N(1 + 2\theta)^{-1}$,*

$$R_{k(1+2\theta)} \subset \cup\{M, M \in \cup_{k<j<k(1+2\theta)} \mathcal{T}_j\}.$$

As before, we call $z_0^*(Q^{(k)})$ a critical point of generation k, and let Γ_k denote the set of critical points of generation $\leq k$.

We are finally in a position to give the definition of $\mathcal{G}_N$ that is valid for all $N \in \mathbb{Z}^+$:

$$\mathcal{G}_N := \{T \in \mathcal{G}_0 \mid \text{(A1')(N), (A2)(N)} - \text{(A5)(N), and (A6')(N) hold}\}.$$

The goal for the remainder of Part II, then, is to prove the following, which is a more general statement than (*) in Sect. 6.2:

(**) For all $N \geq \theta^{-1}$, if $T \in \mathcal{G}_N$ satisfies (A2) and (A4) up to time $\frac{1}{\alpha^*}N$, then it is in $\mathcal{G}_{\frac{1}{\alpha^*}N}$.

[9] By "maximal" we mean no other $Q^{(k)}$ can be squeezed between between $Q^{(k_i)}$ and $Q^{(k_{i+1})}$.

7.3 Clarification and implications

We have tried to capture in (A1') and (A6') a relatively concise summary of the geometric structures that appear after generation θ^{-1}. Before embarking on a formal proof of (**), we would like to take this subsection to elaborate on the implications of these statements and to highlight those features that are new.

In the discussion below, T is assumed to be in $\mathcal{G}_N$.

(1) **Neighborhoods of attractor** Our attractor Ω is defined to be $\Omega = \cap_{k\geq 1} R_k$. For $k \leq \theta^{-1}$, it is natural to see R_k as an approximation of Ω with "finite geometry". Beyond $k = \theta^{-1}$, we are forced to choose between R_k, the geometry of which becomes increasingly complex, and something with simpler geometry. We opted for the latter. The set $\cup_{M\in\mathcal{T}_k} M$, is in general not a good approximation of R_k; in particular, $\cup_{M\in\mathcal{T}_k} M \not\supset \Omega$. On the other hand, (A6')(III) tells us that for all j, $\cup\{M, M \in \cup_{j<k\leq j(1+2\theta)}\} \supset \Omega$. That is to say, while any one level of the tree $\mathcal{T}$ may not be adequate, sets that are unions of branches from $\sim j\theta$ levels starting with level j are *bona fide* neighborhoods of Ω with finite geometry.

(2) **Structure of critical regions** (a) The structure of $\mathcal{C}^{(k)}$ in (A1') is not as complete as that in (A1). First, these regions are no longer nested, i.e. it is not necessarily the case that $\mathcal{C}^{(k+1)} \subset \mathcal{C}^{(k)}$. From its construction in (A1')(II), however, it follows that

$$\mathcal{C}^{(k)} \subset \bigcup_{(1+2\theta)^{-1}k\leq j<k} \mathcal{C}^{(j)}.$$

(b) The following structure inside each component Q of $\mathcal{C}^{(k)}$ is used many times in the analysis to follow: [10]

- With regard to the partial order in (A1')(I)(a), there exist components of $\mathcal{C}^{(j)}$, $j > k$, which lie immediately below Q, i.e. if we call these components Q_i, then $Q_i > Q$ and there is no other Q' with $Q_i > Q' > Q$.
- By (A1')(I)(a)(ii), Q_i is of generation k_i with $k < k_i \leq k(1+2\theta)$. We remark that for $k \leq \theta^{-1}$, $k_i = k+1$. For $k > \theta^{-1}$, this is not necessarily the case. The phenomenon described here will be referred to as the "**skipping of generations**"; it introduces a number of technical problems (that will be addressed in Sect. 8.1).
- By (A1')(II), each Q_i is contained in a horizontal section H_i that stretches across Q, extending considerably beyond, and
- if $k \leq \theta N(1+2\theta)^{-1}$, then by (A6')(III), $(Q \cap R_{k(1+2\theta)}) \subset \cup_i H_i$.

We caution that there may be points $z \in Q^{(k)} \cap R_{k+1}$ that are not in $\cup_i H_i$.

(3) **Eventual set of critical points** At the end of our inductive construction, there is a set $\mathcal{C}$ defined by

$$\mathcal{C} \;=\; \lim_{k\to\infty} \Gamma_k \qquad \text{or, equivalently,} \qquad \mathcal{C} \;=\; \cap_{n>0} \cup_{k\geq n} \mathcal{C}^{(k)}.$$

$\mathcal{C}$ is the set of *critical points* for $T \in \mathcal{G}$. Note that all orbits of $z_0 \in \mathcal{C}$ satisfy (A2) and (A4).

(4) **Critical blobs and geometry of monotone branches**

(a) In the notation of Definition 7.1, E and E' are precisely what we called critical blobs in Sect. 4.2. The requirements that $n-i \leq i\theta^{-1}$ and $n-i' \leq i'\theta^{-1}$ are equivalent to discontinuing M

[10] If the degree of T is zero, then it can happen that $Q \cap \left(\cup_{j>k}\mathcal{C}^{(j)}\right) = \emptyset$.

as soon as one of $z_0^*(B^{(i)})$ or $z_0^*(B^{(i')})$ ceases to be active.

(b) We state a result which together with Lemma 6.1 reinforces our mental picture of what a monotone branch should be: either M is relatively small (such as when all or most of it is in bound state), or it consists of a relatively long horizontal section, namely the part that is free, connecting two relatively small pieces at the ends consisting of points that are in bound state.

Lemma 7.1 *Let $T \in \mathcal{G}_N$, $N \geq \theta^{-1}$. Then for $M \in \mathcal{T}_k$, $1 \leq k \leq \theta N$, the set*

$$\{\xi_{k+1} \in M : \xi_{k+1} \text{ is free}\},$$

if nonempty, is a connected set omitting a neighborhood of $E \cup E'$ in M.

Corollary 7.1 below is a direct consequence of Lemma 7.1. A bound of this type is needed in the treatment of parameter issues in Part III.

Corollary 7.1 *There exist K_1, K_2 depending only on f_0 such that for all $T \in \mathcal{G}_N$, the following hold for $k \leq \theta N$:*
(i) $M \in \mathcal{T}_k$ has at most K_1 children;
(ii) $\mathcal{C}^{(k)}$ has at most K_2^k connected components.

Lemma 7.1 and Corollary 7.1 are proved in Appendix A.20.

We note again that (A1') and (A6') are consistent with (A1) and (A6) for $N \leq \theta^{-2}$. This is because no monotone branches are discontinued before time $N = \theta^{-2}$.

8 Completion of Induction

8.1 Preparation: Sections 4, 5 and 6 revisited

Assume $T \in \mathcal{G}_N$. We begin by bringing to the foreground how the new geometry introduced in Section 7 affects the statements and/or proofs in Sections 4, 5 and 6.

1. *Angles between leaves of $\mathcal{F}_k$ for different k* (Lemma 4.2): The statement of Lemma 4.2 is unchanged. Its proof, which as stated compares the leaves of $\mathcal{F}_k$ and $\mathcal{F}_{k'}$ by going through the leaves of $\mathcal{F}_j$ for all intermediate j. Since not all of the $Q^{(j)}$ are present, the argument needs to be modified slightly: Replace $k+i$ by k_i, $i = 0, 1, \cdots, n$, where $Q^{(k)} = Q^{(k_0)} \supset Q^{(k_1)} \supset \cdots \supset Q^{(k_n)} = Q^{(\hat{k})}$ are the critical regions present. Using the fact that $k_{i+1} \leq k_i(1+2\theta)$, the proof goes through as is.

2. *Distances between critical points* (Lemma 4.1): The statement of Lemma 4.1 is unchanged. In the proof, which estimates $|z_0^*(Q^{(k)}) - z_0^*(Q^{(k+1)})|$, replace $k+1$ by k' where k' is the generation of the next $Q^{(i)}$ inside $Q^{(k)}$. To use Lemma 3.8 to induce a new critical point it suffices to have $\hat{\gamma}$ traverse $B^{(k)}$. This holds easily because $k' \leq k(1+2\theta)$. The order of the newly induced critical point is then updated as before.

3. *"Reproduction" of critical blobs* (last paragraph of Sect. 4.2): Let Q and Q_i be as in paragraph (2)(b) in Sect. 7.3, and let $B^{(k)}$ and $B^{(k_i)}$ be associated with Q and Q_i respectively. Then at time $k\theta^{-1}$, $T^{k\theta^{-1}}(B^{(k)})$ is replaced by $\{T^{k\theta^{-1}}(B^{(k_1)}), \cdots, T^{k\theta^{-1}}(B^{(k_s)})\}$, i.e. in the absence of in-between generations, some of the critical blobs $T^{k\theta^{-1}}(B^{(k_j)})$ may be born a little earlier than before.

4. *Existence of suitable* $\phi(\cdot)$ *for critical orbits* (Lemma 4.4): This is a genuine concern, since fewer critical regions and therefore fewer critical points are available. Both the definition of $\phi(z_i)$ and the statement of Lemma 4.4 are unchanged. Its proof is modified as follows:

Case 1. $\hat{j}(1+2\theta) \leq \alpha^*\theta i$. This implies $z_i \in Q^{(\hat{j})} \cap (H \setminus Q^{(j)})$ for some horizontal section H of generation j crossing $Q^{(\hat{j})}$, $\hat{j} < j \leq \hat{j}(1+2\theta)$.

Case 2. $\hat{j} \leq \alpha^*\theta i < \hat{j}(1+2\theta)$. Here all $\hat{j}+1$ are replaced by $\hat{j}(1+2\theta)$, and $d_{\mathcal{C}}(z_i) > e^{-\alpha i}$ is used as before.

In both cases, there is enough room for the new estimates to go through.

With regard to Section 5, we remark that unlike the situation in the last paragraph, the results in this section *assume* the existence of guiding critical orbits and so are unaffected. We go directly to Sect. 6.3.

5. *Setting control for* (ξ_k, τ_k) (Sect. 6.3): Here we need to set control for all $\xi_i \in M^\circ \cap \mathcal{C}^{(1)}$, $M \in \mathcal{T}_i$. The rules on the selection of $\phi(\xi_i)$ at free returns are essentially the same as (1)–(3) in the beginning of Sect. 6.3, with (3) modified to read as follows: If there exists $j < i$ such that $\xi_i \in \mathcal{C}^{(j)} \setminus \cup_{k>j}\mathcal{C}^{(k)}$, then we let $\phi(\xi_i) = z_0^*(Q^{(j)}(\xi_i))$; otherwise $\phi(\xi_i)$ is of generation i. At issue is the suitability of this choice of $\phi(\xi_i)$. Given that $\phi(\xi_i)$ may be of a lower generation than it would have been (due to the skipping of generations) and that certain monotone branches are discontinued, two questions are: (i) Do we know that $\phi(\xi_i)$ will remain active for as long as it is needed? (ii) If $\phi(\xi_i)$ is of generation $j < i$, is $\xi_i \notin B^{(j)}$?

Let j be the generation of $\phi(z_i)$. There are two cases to consider.

Case 1. $j \leq i(1+2\theta)^{-1}$. Here ξ_i, which is in R_i, lies in a horizontal section of generation j' crossing $Q^{(j)}(\xi_i)$. We may assume $j < j' \leq j(1+2\theta)$; see (2)(b) in Sect. 7.3. Observe that since $j' > j$, ξ_i is not in $\mathcal{C}^{(j')}$ by assumption. It follows that $d_{\mathcal{C}}(\xi_i) > e^{-\lambda(j+1)(1+2\theta)}$, so $\xi_i \notin B^{(j)}$ and $p(\xi_i) << \theta^{-1}j$.

Case 2. $j > i(1+2\theta)^{-1}$. By assumption, $\xi_i \in M_i \cap Q^{(j)}$ for some $M_i \in \mathcal{T}_i$. It follows, by (A1')(II), that M_i extends all the way across $Q^{(j)}$ and $M_i \cap Q^{(j)}$ contains a component $Q^{(i)}$ of $\mathcal{C}^{(i)}$. If $\xi_i \notin Q^{(i)}$, the situation is as in Case 1. If $\xi_i \in Q^{(i)}$, then $\phi(\xi_i) = z_0^*(Q^{(i)}(\xi_i))$. It is easy to show that for $k = 1, 2, \cdots$, $T^k\xi_i$ and $T^k\phi(\xi_i)$ lie in the same element of $\mathcal{T}_{i+k}$ until either this branch is discontinued or the bound period of ξ_i expires. Recall that when the orbit of $\phi(\xi_i)$ ceases to be active, the monotone branch containing it is automatically discontinued.

6. *Proposition 6.1* (Sect. 6.3): With the idea of provisional control as before, we reformulate Proposition 6.1 as follows:

Proposition 8.1 *For* $T \in \mathcal{G}_N$*, and* $\theta N < k \leq \frac{1}{\alpha^*}\theta N$*:*

(a)$_k$ $\mathcal{C}^{(k)}$ *and* Γ_k *with the properties in (A1') can be constructed;*

(b)$_k$ $\mathcal{T}_{k+1}$ *with the properties in (A6') – except for the provisional nature of the control in (A6')(II) – can be constructed.*

The proof of this proposition is postponed to Sect. 8.2. Assuming it for now, we continue with our list of modifications.

7. *Alignment of vectors at free returns* (Sect. 6.4): In the proof of Proposition 6.2(i), we view z_{k-j} as a point $\xi_1 \in R_1$ and show that (ξ_1, τ_1) is controlled for j iterates. In this type of arguments, one needs to verify that there exists $M \in \mathcal{T}_{j+1}$ such that $\xi_{j+1} \in M$, i.e. the ancestors of this branch were not discontinued. Here we know M exists because our choice of $\phi(z_k)$ implies the existence of $Q^{(j+1)}$ with $\phi(z_k) = z_0^*(Q^{(j+1)})$ and $\xi_{j+1} = z_k \in Q^{(j+1)}$, and, by definition, every $Q^{(j+1)}$ is contained in some $M \in \mathcal{T}_{j+1}$. The rest of the proof of Proposition 6.2 is not affected.

The quadratic estimate in (A5) relies on the behavior of the guiding critical orbits and not on global geometry; it is therefore not affected.

This completes our list of modifications for Sections 4, 5 and 6 and hence the proof of **(**)** stated at the end of Sect. 7.2 – modulo the proof of Proposition 8.1.

8.2 Construction of critical regions and monotone branches

Proof of Proposition 8.1: We follow in outline the proof of Proposition 6.1, focusing on those aspects of the situation that are new.

Assume (a)$_i$ and (b)$_i$ for all $i < k$. For $M \in \mathcal{T}_k$, if $\xi_1 \in T^{-(k-1)}M^\circ$, then the sequence $(\xi_1, \tau_1), \cdots (\xi_{k-1}, \tau_{k-1})$ is provisionally controlled by Γ_{k-1}. Thus it makes sense to speak about the leaf segments of $\mathcal{F}_k$ on M° as being in bound or free states. We divide the proof of (a)$_k$ and (b)$_k$ into the following steps:

1. *Construction of* $\mathcal{C}^{(k)}$ *and* Γ_k

Let $M \in \mathcal{T}_k$, and let $Q = Q^{(j)}$, $k(1+2\theta)^{-1} \le j < k$, be such that $M \cap Q \neq \emptyset$. We observe that $M \cap Q$ is the union of horizontal sections each extending $> \frac{1}{2}e^{-\alpha k}$ on both sides of Q. This is true because by (A2) for a much earlier time, $(E \cup E') \cap Q = \emptyset$ where E and E' are the ends of M. We then consider one $\mathcal{F}_k$-leaf segment in M at a time and argue as in the proof of Proposition 6.1.

Next we explain *where* $Q^{(k)}$ is constructed (postponing *how* it is done to the next paragraph). For each $M \in \mathcal{T}_k$, let $\mathcal{I}(M)$ be the set of all connected components of $M \cap Q^{(j)}$, $k(1+2\theta)^{-1} \le j < k$. We define a partial order on $\mathcal{I}(M)$ by set inclusion, i.e. $S < S'$ if $S \supset S'$. Let $\mathcal{H}(M)$ be the set of maximal elements. A critical region $Q^{(k)}$ is constructed in each $H \in \mathcal{H}(M)$.

As to *how* to construct $Q^{(k)}$, let $H \in \mathcal{H}(M)$ be a component of $M \cap Q^{(j)}$. We fix an arbitrary $\mathcal{F}_k$-leaf γ in H, and use Lemma 3.8 and $z_0^*(Q^{(j)})$ to induce a (unique) critical point z_0' of order j on γ. In x-coordinate, we know from Lemma 3.8 that z_0' is $< Kb^{\frac{j}{4}}$ away from the center of $Q^{(j)}$. Lemma 3.9 then tells us that near z_0' there is a unique critical point z_0 of order k on γ. We put $z_0 \in \Gamma_k$, and construct a $Q^{(k)}$, i.e. a section of length $2\min(\delta, e^{-\lambda k})$ centered at it.

$\mathcal{C}^{(k)}$ is defined to be the union of all the $Q^{(k)}$ constructed as we let M vary over $\mathcal{T}_k$.

2. *Verification of (A1'):*

The procedure above gives immediately the following: (1) The $Q^{(k)}$ constructed are disjoint sections of R_k and hence are genuine connected components of $\mathcal{C}^{(k)}$. (2) The partial order in (A1')(I)(a) is extended to $\{Q^{(k')}, k' \le k\}$; this is because each $Q^{(k)}$ constructed lies immediately below a unique $Q^{(j)}$, $k(1+2\theta)^{-1} \le j < k$, in this partial order. (2) implies (3), namely that the jumps in generation in (A1')(I)(ii) are as claimed. Observe also that (A1')(II) is fulfilled. As for (A1')(I)(b), all statements are true by construction except the one regarding sectional diameter, which follows directly from the next lemma.

Lemma 8.1 *Let* $M \in \mathcal{T}_k$. *Then for all* $\xi_k \in M^\circ$, *there exists a codimension one manifold* W *with* $\xi_k \in W$ *such that*

- *W meets every connected component of $\mathcal{F}_k$-leaf in M in exactly one point;*
- *for all $\xi_1, \xi_1' \in T^{-k+1}W$, $|\xi_i - \xi_i'| < (Kb)^{\frac{i}{2}}$ for $i \le k$.*

The proof of this lemma is a small modification of that of Lemma 6.2. Details are left to the reader.

3. *Construction of* $\mathcal{T}_{k+1}$ *and verification of (A6')(I) and (II)*

The relation between $M \in \mathcal{T}_k$ and $B^{(k)}$ follows immediately from our construction in Step 1. We construct $\mathcal{T}_{k+1}$ as described in the paragraph on "Tree of monotone branches" in Sect.

7.2, proving (A6')(I). To prove (A6')(II), it suffices, as in the proof of Proposition 6.1, to prove correct alignment of the τ-vectors at free returns, and we consider the two cases as before. To make transparent the effect of the "missing generations", we describe in some detail the geometry in Case 2, the case where γ, our maximal free $\mathcal{F}_k$-segment, does not meet $\mathcal{C}^{(k)}$.

Let ξ_k be fixed. We let $j = j_0$ be the largest integer such that $\xi_k \in \mathcal{C}^{(j)}$, and let $Q^{(j_0)} = Q^{(j_0)}(\xi_k)$. Then $j_0(1+2\theta) < k$, otherwise γ would lie in a monotone branch crossing $Q^{(j_0)}$, contradicting our assumption that it does not meet $\mathcal{C}^{(k)}$. By (A6')(III), $\xi_k \in R_{j_0(1+2\theta)} \subset \cup\{M, M \in \cup_{j_0<\ell<j_0(1+2\theta)}\mathcal{T}_\ell\}$. Thus $\xi_k \in H$ where H is a horizontal section of generation j_1, $j_0 < j_1 \le j_0(1+2\theta)$, that crosses the entire length of $Q^{(j_0)}(\xi_k)$. We may assume $Q^{(j_1)}$ is immediately below $Q^{(j_0)}$ in the partial order. Suppose for definiteness that ξ_k lies in the right chamber of $H \setminus Q^{(j_1)}$. We move left along γ until we reach either ξ, the left endpoint of γ, or the right boundary of $Q^{(j_1)}$, noting that since $j_1 < k$, γ cannot meet ∂R_{j_1}. If we reach the boundary of $Q^{(j_1)}$ before reaching ξ, then we continue to move left along γ, going into $Q^{(j_1)}$. For the same reason as above, $j_1(1+2\theta) < k$, so that as we enter $Q^{(j_1)}$, we have entered a horizontal section of generation j_2, $j_1(1+2\theta)^{-1} \le j_2 < j_1$, that crosses the entire length of $Q^{(j_1)}$, and so on. After going through a finite number of $Q^{(j_i)}$, we must arrive at ξ, for the j_i are strictly increasing with i and $< k(1+2\theta)^{-1}$.

The argument showing $\phi(\xi)$ is to the left of ξ is essentially the same, except for the fact that the interpolating chain $Q^{(j')} \subset \cdots \subset Q^{(j'')}$ also involves skips in generation. One way to see that such a chain exists is to start from $Q^{(j'')}$.

After these preparations, the angle estimates are unchanged. This completes the verification of (A6')(II).

4. *Proof of (A6')(III)* This step involves a very different set of ideas. We formulate the result as Proposition 8.2 and give the proof in the next subsection.

Proposition 8.2 *We consider $T \in \mathcal{G}_N$, and let $n \le \frac{1}{\alpha^*}\theta N$. Assume (a)$_k$ and (b)$_k$ in Proposition 8.1 for all $k \le n-1$. Then for all $k \le n(1+2\theta)^{-1}$,*

$$R_{k(1+2\theta)} \subset \cup\{M, M \in \cup_{k\le j<k(1+2\theta)}\mathcal{T}_j\}. \tag{9}$$

Modulo this result, the proof of Proposition 8.1 is now complete. □

8.3 Branch replacement: Proof of Proposition 8.2

As explained in Sect. 8.1, a monotone branch is discontinued before either one of its ends becomes too large. The problem of "branch replacement", roughly speaking, is one of finding a collection of branches of higher generations that together cover the part of the attractor "exposed" by the removal of the discontinued branch. Proposition 8.2 tells us explicitly what neighborhoods are covered by which collections of branches.

We begin with some preliminary definitions. Let M_1 be a monotone branch of generation $k_1 > k$. We say M_1 is *subordinate* to M if (i) $M_1 \subset M$, (ii) the ends of M and M_1 are related as follows: Let E and E' be the ends of M, and E_1 and E_1' the ends of M_1. Suppose $T^{-i}E = \hat{B}^{(k-i)}$, then $T^{-i}E_1 = \hat{B}_1^{(k_1-i)}$ with $\hat{B}_1^{(k_1-i)} \subset \hat{B}^{(k-i)}$; and E' and E_1' are related the same way. A collection of monotone branches $\{M_j\}$ subordinate to M is called a *viable replacement* for M if $(M^\circ \cap \Omega) \subset \cup_j M_j$.

Lemma 8.2 *There exists $K > 0$ for which the following holds: Suppose $M \in \mathcal{T}_k$, $k \le n$, has an end E with the property that $T^{-i}E = \hat{B}^{(k-i)}$, $i \ge K\alpha(k-i)$. Then $T^{-i}M$ is contained in a horizontal section H of R_{k-i} of length $< e^{-2\alpha(k-i)}$ centered at $\hat{B}^{(k-i)}$.*

Lemma 8.2 is proved in Appendix A.21. Assume, for definiteness, that $T^{-i}M$ lies in the right half of H (it contains, needless to say, $\hat{B}^{(k-i)}$). To look for a viable replacement for M, we examine the structures inside H more closely.

For $j = 1, 2, \cdots, i-1$, let $S_j \in \mathcal{T}_{k-i+j}$ be the ancestors of M, and let $\hat{S}_0 = T^{-1}S_1$, so that $\hat{S}_0$ is a section of R_{k-i} containing the right half of H. From Lemma 8.2, it follows that there exists ℓ with $\ell < K\alpha(k-i)$ such that $T^{-\ell}S_\ell \subset H$.

Consider now $P \in \mathcal{T}_p$, $(k-i) < p < (1+2\theta)(k-i+1)$, such that $P \cap H \neq \emptyset$. Then $P \cap H$ is the union of horizontal sections that run the entire length of H. (See Fig. 3.) We fix one component of $H \cap P$, call it $\hat{H}$, and let $\hat{B}^{(p)}$ denote the $B^{(p)}$ in $\hat{H}$. Whenever possible, we define $P_j \in \mathcal{T}_{p+j}$, $j = 1, 2, \cdots, i$, as follows: P_1 is the child of P such that $T^{-1}P_1$ contains the right half of $\hat{H}$; for $j > 1$, P_j is the child of P_{j-1} one of whose ends is $T^j\hat{B}^{(p)}$. If well defined, P_j depends on M, P and $\hat{H}$; we write $P_j(M, P, \hat{H})$.

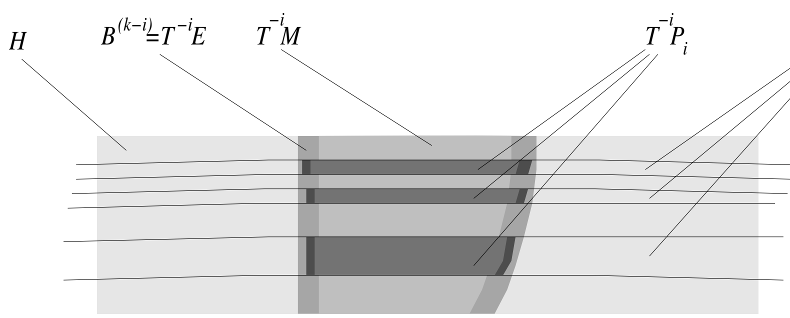


Fig. 3 Replacement branches: $\cup P$ is used to replaced M

We observe that P_j may not be defined: First, P may be discontinued, in which case it has no children. If that is not the case, then P_1 is defined. Let E_1' denote the end of P_1 not equal to $T(\hat{B}^{(p)})$. This is clearly the older of the two ends of P_1. It may cause P_1 to be discontinued. We may also have $P_2 = T(P_1)$, with $T(E_1')$ causing P_2 to be discontinued, and so on.

Let n be as in the statement of Proposition 8.2, and let M, P and H be as above.

Lemma 8.3 *We assume $k \leq (1+2\theta)^{-1}n$. Then the following hold for every component $\hat{H}$ of $P \cap H$: If P_1 is well defined and E_1' remains active for ℓ generations for some ℓ with $T^{-\ell}S_\ell \subset H$, then*

(i) P_j is well defined for all $j \leq i$, and

(ii) P_i is subordinate to M.

Lemma 8.3 is proved in Appendix A.21.

Proof of Proposition 8.2: Our strategy is to construct, for $m = 1, 2, \cdots, n(1+2\theta)^{-1}$, a collection of monotone branches $\mathcal{S}_m$ with the following properties:

(i) For every $M \in \mathcal{S}_m$, if $M \in \mathcal{T}_k$ and E is an end of M with $T^{-i}E = B^{(k-i)}$, then $i \leq \frac{2}{3}(k-i)\theta^{-1}$;

(ii) $\mathcal{S}_m \subset \cup_{m \leq k < m(1+2\theta)} \mathcal{T}_k$, and

(iii) $\cup\{M, M \in \mathcal{S}_m\} \supset R_{m(1+2\theta)}$.

Proposition 8.2 follows immediately from (ii) and (iii). We say $M \in \mathcal{S}_m$ is at *replacement time* if equality holds in (i), i.e. $i = \frac{2}{3}(k-i)\theta^{-1}$, for one of its ends. For reasons to become clear momentarily, we have elected to define replacement time to occur somewhat before the branch is discontinued.

Let $\mathcal{S}_1 = \{R_1\}$. We assume for all $k \le m$, $\mathcal{S}_k$ has been constructed and has properties (i)–(iii).

Construction of $\mathcal{S}_{m+1}$ with property (i): We consider $M \in \mathcal{S}_m$ one at a time. If M has not reached its replacement time, then we put all the children of M into $\mathcal{S}_{m+1}$. If it has, then we put into $\mathcal{S}_{m+1}$ the children of $\{P'\}$ where $\{P'\}$ is defined as follows: Suppose $M \in \mathcal{T}_k$ and E is an end with $T^{-i}E = \hat{B}^{(k-i)}$ and $i = \frac{2}{3}(k-i)\theta^{-1}$. Then $\{P'\} = \{P_i(M, P, \hat{H}) : P \in \mathcal{S}_{k-i+1}$ and $\hat{H}$ is a component of $H \cap P\}$.[11]

First we show that P' is well defined as an element of $\mathcal{T}_{p+i}$ where p is the generation of P. To do that, it suffices to verify the hypotheses of Lemma 8.3. To begin with, P_1 is well defined as an element of $\mathcal{T}_{p+1}$ because $P \in \mathcal{S}_{k-i+1}$ and by property (i) for $\mathcal{S}_{k-i+1}$, both ends of P will remain active for some period of time. Let E_1' be as above, i.e. the "other" end of P_1. We claim that E_1' will last $>> K\alpha(k-i)$ generations: Suppose it was created ℓ generations prior to $p+1$. Then $\ell \le \frac{2}{3}(p+1-\ell)\theta^{-1}$, so that

$$\begin{aligned} K\alpha(k-i) + \ell &\le K\alpha(k-i-\ell) + (K\alpha+1)\ell \\ &\le K\alpha(k-i-\ell) + (K\alpha+1)\frac{2}{3}(p+1-\ell)\theta^{-1} \\ &< (p+1-\ell)\theta^{-1}. \end{aligned}$$

The hypothesis of Lemma 8.3 is verified and P' is defined.

To prove that the children of P' meet the condition in property (i) for $\mathcal{S}_{m+1}$, we let $\hat{E}$ and $\hat{E}'$ be its two ends, $\hat{E}$ being the one contained in E. This end is created the same time E is created. Clearly, $i < \frac{2}{3}p\theta^{-1}$ since $p > k-i$. As for $\hat{E}'$, it follows from the analysis in Lemma 8.3 that this is the younger of the two ends. Thus it cannot have reached replacement time if $\hat{E}$ has not. This completes the construction of $\mathcal{S}_{m+1}$ with property (i).

Proof of property (ii) for $\mathcal{S}_{m+1}$: Let $M \in \mathcal{S}_m$. If M is not replaced, then the children of M are obviously of acceptable generation. If replacement occurs, then the generation of P' is estimated as follows: Let all notation be as above. Since $P \in \mathcal{S}_{k-i+1}$, we have, by inductive assumption, $p < (1+2\theta)(k-i+1)$. Combining this with $\frac{3}{2}\theta i = k-i$, we have

$$p + i \le i + (1+2\theta)(k-i+1) = i + (1+2\theta)(\frac{3}{2}\theta i + 1) < (i+1)(1+2\theta).$$

To complete the proof, we show that $i \le m$. First, it is true for $m = 2$. In general, we claim that if E is an end of M and $T^{-i}E = B^{(k-i)}$, then $i \le m$. This is obviously so if no replacement occurs. In a replacement procedure, observe that even though the generations of the new branches are higher, their ends are created exactly the same number of generations earlier as the branch replaced. (See the proof of Lemma 8.3.)

Proof of property (iii) for $\mathcal{S}_{m+1}$: As before, it suffices to consider the case where $M \in \mathcal{S}_m$ is at replacement time. We claim that $\{P'\}$ is a viable replacement for M. Let H be as above. By induction hypothesis, more specifically, by (iii) for $k-i+1$,

$$H \cap R_{(k-i+1)(1+2\theta)} \subset H \cap (\cup\{M', M' \in \mathcal{S}_{k-i+1}\}) \subset H \cap (\cup P) ,$$

that is to say, $H \setminus (\cup P)$ does not meet $R_{(k-i+1)(1+2\theta)}$. Thus the part of phase space deleted as we replace M by $\{P'\}$ does not meet R_g where

$$g := (k-i+1)(1+2\theta) + i .$$

[11]If $H \cap P = \emptyset$ for all $P \in \mathcal{S}_{k-i+1}$, then there is no need for replacement. Also, where I is an interval (see Sect. 3.9), the two special branches in $\mathcal{T}_m$ having $T^{m-1}V_i$ as one of their ends are always in $\mathcal{S}_m$.

The same computation as in the proof of property (ii) gives $g < (m+1)(1+2\theta)$.

This completes the proof of Proposition 8.2. □

Remarks 1. As the proof shows, there is a natural time step for branch replacements. They occur very infrequently, roughly once every $\sim \theta^{-1}$ iterates. Thus when working with the replacement of a branch of generation k, it is mostly structures of generation up to $\sim \theta k$ – including assumption (A6')(III) for these times – that count, although certain properties of critical orbits and the well-definedness of branches up to time $k-1$ are also needed.

2. In spite of the qualitative flavor of the statement, the existence of viable replacements in the setting above reflects the fact that monotone branches reproduce at rates much faster than the speed with which critical blobs are allowed to approach the critical set.

9 Construction of SRB Measures

The definition of SRB measures is given in Section 1. For more information on the subject, see [Y1] and [Y2]. The goal of this section is to prove

Proposition 9.1 *Let $T \in \mathcal{G}$. Then T has an SRB measure.*

9.1 Generic part of construction

Our construction can be thought of as having a "generic" part, i.e. a part that can be used for many dynamical systems, and a "situation-dependent" part, i.e. a part that relies (seriously) on the properties of the map in question. The goal of this subsection is to treat the generic part. We pinpoint what specific information is needed and then *assume* it to complete the construction. For notational convenience, we give the proof in the setting of Part II of this paper, remarking that aside from $\dim W^u = 1$, other properties of T used below are inessential.

Step 1. Pushing forward Lebesgue measure on a W^u-leaf

Let l_0 be a piece of local unstable manifold through $\hat{z}$ where $\hat{z}$ is a hyperbolic periodic point or belongs in a uniformly hyperbolic invariant set. We let m_0 be the Riemannian measure on l_0, and for $n = 1, 2, \cdots$, define

$$\nu_n \;=\; \frac{1}{n}\sum_{i=0}^{n-1} T_*^i(m_0)$$

where $T_*^i(m_0)$ is the measure with $T_*^i(m_0)(E) = m_0(T^{-i}(E))$ for all Borel sets E. Let ν be a limit point of ν_n in the weak* topology. It is easy to see that ν is T-invariant.

To prove that ν is an SRB measure, it is necessary to show that it has absolutely continuous conditional measures (*accm*) on unstable manifolds. By design, this property is enjoyed by ν_n for all n. Whether it is passed from ν_n to ν, however, depends on a number of factors that are situation-dependent. We describe in Step 2 a construction to facilitate this passage if the conditions are right. Our construction is based on the idea (also used in [BY]) that it suffices to control a small fraction of ν_n.

Step 2. "Catching" a fraction of ν with accm *on unstable curves*

First we introduce some language convenient for our purposes. We call a curve γ an *unstable curve* if there exist $\kappa < 1$ and $K > 1$ such that for all $z \in \gamma$ and $\tau \in X_z$ tangent to γ, $|DT_z^{-n}\tau| \le K\kappa^n|\tau|$ for all $n \ge 0$. Next we introduce the objects used to "catch" a part of ν.

Let L be an interval, and let Σ be a compact set. We say $\Psi : L \times \Sigma \to R_1$ is *a continuous family of unstable curves* if

(a) Ψ maps $L \times \Sigma$ homeomorphically onto its image;

(b) for each $\alpha \in \Sigma$, $\Psi|_{L\times\{\alpha\}}$ is a C^1 embedding, and $D_\alpha := \Psi(L \times \{\alpha\})$ is an unstable curve;

(c) $\alpha \mapsto \Psi|_{L\times\{\alpha\}}$ is continuous as a map from Σ to $C^1(L, R_1)$.

We will use the notation $\mathcal{N} = \Psi(L \times \Sigma) = \cup_\alpha D_\alpha$.

The following condition, the validity of which depends on the specifics of the map in question, is assumed for the rest of this subsection:

(S) There exist $c > 0, K \geq 1$, a continuous family of unstable curves $\mathcal{N} = \cup_\alpha \mathcal{D}_\alpha$, and a sequence of integers $n_1 < n_2 < \cdots$ for which the following hold. For each $i \geq 0$, there is a collection $\{\omega_j^{(i)}\}$ of subsegments of l_0 such that

(i) for each i, j, $T^i(\omega_j^{(i)}) = D_\alpha$ for some α;

(ii) letting $\tau(z)$ denote a unit vector tangent to l_0 at $z \in l_0$, we have, for all $z, z' \in \omega_j^{(i)}$,

$$\frac{|DT_z^i \tau(z)|}{|DT_{z'}^i \tau(z')|} < K;$$

(iii) $\frac{1}{n_k} \sum_{i=0}^{n_k-1} m_0(\cup_j \omega_j^{(i)}) > c\, m_0(l_0)$ for all n_k.

Let $\hat{\nu}_{n_k} = \frac{1}{n_k} \sum_{i=0}^{n_k-1} T_*^i(m_0|_{\cup_j \omega_j^{(i)}})$, and let $\hat{\nu}$ be a limit point of $\hat{\nu}_{n_k}$. It follows from (iii) in Condition **(S)** that $\hat{\nu}(\mathcal{N}) > c m_0(l_0) > 0$. From (i) and (ii), we see that for each n_k, $\hat{\nu}_{n_k}$ is supported on a finite number of D_α, and its densities with respect to arclength measure on D_α are bounded between K and $\frac{1}{K}$. The absolute continuity of conditional measures of $\hat{\nu}$ on $\{D_\alpha\}$ is now a simple exercise: Let $\eta_1 < \eta_2 < \cdots$ be any increasing sequence of finite partitions on Σ such that $\bigvee_{i>0} \eta_i$ partitions Σ into points, and let $\mathcal{E}_i$ be the partition on $\mathcal{N}$ given by $\{\Psi(L \times S) : S \in \eta_i\}$. Then $\mathcal{E}_\infty := \bigvee \mathcal{E}_i$ is the partition of $\mathcal{N}$ into $\{D_\alpha\}$. Let $\ell \subset L$ be an arbitrary interval, and let $A = \Psi(\ell \times \Sigma)$. Then there exists K' depending on the constant K in **(S)**(ii) and on the norms of the embeddings $\Psi|_{L\times\{\alpha\}}$ such that for all n_k and i,

$$\frac{1}{K'} \frac{|\ell|}{|L|} \;\leq\; (\hat{\nu}_{n_k}|\mathcal{E}_i)(A) \;\leq K' \frac{|\ell|}{|L|}. \tag{10}$$

Here $\hat{\nu}_{n_k}|\mathcal{E}_i$ denotes the conditional measure of $\hat{\nu}_{n_k}$ given $\mathcal{E}_i$. The relation in (10) is first passed to $\hat{\nu}|\mathcal{E}_i$ by letting $n_k \to \infty$. It is then passed, by the martingale convergence theorem, to $\hat{\nu}|\mathcal{E}_\infty$. Our assertion on the conditional measures of $\hat{\nu}$ follows as we let ℓ range over a countable basis of the Borel topology on L.

Step 3. Extracting an SRB measure from ν

Let ν be as in Step 1. Then $\nu(\mathcal{N}) \geq \hat{\nu}(\mathcal{N}) > 0$, and $\nu_{\mathcal{N}} := \nu|_{\mathcal{N}}$ is $T_{\mathcal{N}}$-invariant where $T_{\mathcal{N}} : \mathcal{N} \to \mathcal{N}$ is the first return map of T to $\mathcal{N}$. Let $R : \mathcal{N} \to \mathbb{Z}^+$ be the first return time, and assume for the moment that $\nu_{\mathcal{N}}$ has *accm* on $\{D_\alpha\}$. We claim that

$$\mu := \sum_{n\geq 0} \sum_{k=0}^{n-1} T_*^k(\nu_{\mathcal{N}}|_{\{R=n\}})$$

normalized is an SRB measure. To prove this, it suffices to show (i) T has a positive Lyapunov exponent μ-a.e., and (ii) the D_α-curve through a.e. $z \in \mathcal{N}$ is its local unstable manifold. Both are true by the construction in Step 2.

While $\hat{\nu}$ (which is not necessarily $T_{\mathcal{N}}$-invariant) has *accm* on $\{D_\alpha\}$, we do not know in general that $\nu_{\mathcal{N}}$ does as well. The following procedure is used to extract a part of $\nu_{\mathcal{N}}$ with the desired property: Let $\nu_\Sigma = (\pi_\Sigma)_*(\Psi_*^{-1}(\nu_{\mathcal{N}}))$ where $\pi_\Sigma : L \times \Sigma \to \Sigma$ is projection. Let λ denote Lebesgue measure on L, and let $\nu_{Leb} = \Psi_*(\lambda \times \nu_\Sigma)$. We then decompose $\nu_{\mathcal{N}}$ into $\nu_{\mathcal{N}} = \nu_{ac} + \nu_\perp$ where ν_{ac} is absolutely continuous with respect to ν_{Leb} and $\nu_\perp$ is singular with respect to it (written $\nu_{ac} << \nu_{Leb}$ and $\mu_\perp \perp \nu_{Leb}$). Since $(T_{\mathcal{N}})_*(\nu_{ac}) << \nu_{Leb}$ and $(T_{\mathcal{N}})_*(\nu_\perp) \perp \nu_{Leb}$, it follows that both ν_{ac} and $\nu_\perp$ are $T_{\mathcal{N}}$-invariant. By Condition **(S)**, $\nu_{ac}(\mathcal{N}) \geq \hat{\nu}(\mathcal{N}) > 0$. The construction of μ above can now be carried out with ν_{ac} in the place of $\nu_{\mathcal{N}}$.

Summary: We have shown that if for some $l_0 = W^u_{loc}(\hat{z})$ there is a family of unstable curves $\mathcal{N} = \cup_\alpha \mathcal{D}_\alpha$ for which Condition **(S)** is satisfied, then T admits an SRB measure.

For the rest of this section, we assume $T \in \mathcal{G}$.

9.2 Dynamics on unstable manifolds

Let $\Omega_\delta := \{z_0 \in \Omega : d_{\mathcal{C}}(z_i) \geq \delta \ \forall i \in \mathbb{Z}\}$. From Sect. 3.5, we know that $T|_{\Omega_\delta}$ is uniformly hyperbolic. Fix $\hat{z} \in \Omega_\delta$, and let $l_0 = W^u_r(\hat{z})$ denote its local unstable curve of radius r. We assume r is small enough that for all $\xi_0 \in l_0$, $d_{\mathcal{C}}(\xi_{-i}) > \frac{1}{2}\delta$ for all $i \geq 0$. In the rest of this section, we let $\tau_0(\xi_0) \in X_{\xi_0}$ denote the positively oriented unit vector tangent to l_0, and use τ to denote generic unit vectors tangent to $l_i := T^i l_0$.

A. Control of (ξ_0, τ_0) for $\xi_0 \in l_0$

For $z \in \Omega \setminus \mathcal{C}$, a natural choice of $\phi(z)$ is $\phi(z) = z_0^*(Q^{(j)})$ where j is the largest k such that $z \in Q^{(k)}$. Observe that $j = \infty$ corresponds exactly to $z \in \mathcal{C}$. Thus for all $\xi_0 \in \Omega$ such that $\xi_i \notin \mathcal{C}$ for all $0 \leq i < k$, ξ_0 is controlled by $\cup_{j>0}\Gamma_j$ for k iterates. Proof of control for $\tau_0(\xi_0)$ is obtained by leveraging (A6'); details are given in Appendix A.22.

Lemma 9.1 *For all $\xi_0 \in l_0$, the sequence $(\xi_0, \tau_0), \cdots, (\xi_k, \tau_k)$ is controlled by $\cup\Gamma_j$ provided $\xi_i \notin \mathcal{C}$ for all $i < k$.*

Once control is established, the evolution of l_i is very similar to that of $\mathcal{F}_k$-leaves. We record their geometric and dynamical properties in B and C below.

B. Geometry of l_i

The proof of the following is entirely parallel to that of Lemma 7.1:

(i) For each $i > 0$, l_i is partitioned by $\{T^{i-k} z_k, k < i, z_0 \in \bar{\Gamma}\}$ where $\bar{\Gamma}$ is the closure of $\cup_j \Gamma_j$ into a finite disjoint union of *monotone segments* $\{\sigma\}$.

(ii) The free part of σ, if nonempty, is connected, and the function on σ giving the number of iterates before a point becomes free is U-shaped.

(iii) If $\sigma \cap Q^{(1)} \neq \emptyset$, then either (a) $\sigma \cap Q^{(1)}$ meets $\bar{\Gamma}$ in a single point $\hat{z}_0$, and σ contains a $C^2(b)$ curve of length $e^{-\alpha i}$ centered at $\hat{z}_0$, or (b) $\sigma \cap Q^{(1)}$ lies strictly to one side of $\bar{\Gamma}$; in this case we let $\hat{z}_0 = \phi(\xi)$ where ξ is the point in $\sigma \cap Q^{(1)}$ closest to $\mathcal{C}$.

C. 1D behavior

Behavior near the "turns" excepted, the dynamics of $l_0 \to l_1 \to l_2 \to \cdots$ bear a striking resemblance to those of iterated 1D maps. By this, we mean a *qualitative* resemblance rather than the existence of a specific map $f : I \to I$ with the property that $f^k(\pi_x(l_0)) \approx \pi_x(l_k)$ for all k. Here π_x denotes projection onto the x-axis. To make precise this qualitative resemblance, we formulate three properties in analogy with (P1)–(P3) in Sect. 2.2.

Proposition 9.2 *$T^k|_{l_0}, k = 1, 2, \cdots,$ satisfy (P1')–(P3') below.*

(P1') (Outside of $\mathcal{C}^{(1)}$) Let $f_0 \in \mathcal{M}$ be as in Sect. 6.1, and let $\varepsilon = \max(\mathcal{O}(a), \mathcal{O}(b))$. Then for all $z \in l_i \setminus \mathcal{C}^{(1)}$ in a free state, we have

$$|\pi_x(Tz) - f_0(\pi_x(z))|, \quad ||DT_z(\tau)| - f_0'(\pi_x(z))| \ < \varepsilon.$$

From this it follows that results analogous to (P1)(i),(ii) with $|(f^n)'|$ replaced by $|DT^n(\tau)|$ hold with slightly weaker constants for segments of l_i in free state.

(P2') (Bound periods and derivative recovery) Let ω be the maximal free segment in a component of $l_i \cap \mathcal{C}^{(1)}$, let $\hat{z}_0$ be as in B(iii) above, and let $\mathcal{P}^{\hat{z}_0}$ be the partition in Sect. 2.2 centered at $\pi_x(\hat{z}_0)$. Then there exist K_0 and K_1 such that

(i) $K_0^{-1} \log \frac{1}{|z-\hat{z}_0|} \leq p(z) \leq K_0 \log \frac{1}{|z-\hat{z}_0|}$ for all $z \in \omega$;

(ii) $|DT_z^p(\tau)| > e^{\frac{1}{4}\lambda p(z)}$ for all $z \in \omega$;

(iii) if $\pi_x(\omega) \approx I_{\mu j}$ for some $I_{\mu j} \in \mathcal{P}^{\hat{z}_0}$, then $|T^p(\omega)| > e^{-K_1 \alpha |\mu|}$.

Let ω be a segment of l_i. We say all $z \in \omega$ have the *same itinerary* for $n-1$ iterates if there exist $t_1 < t_1 + p_1 \leq t_2 < t_2 + p_2 \leq \cdots \leq n$ such that for every k, $\pi_x \circ T^{t_k}\omega \subset P^+$ for some $P \subset C_\delta$, $p_k = \min_{z\in\omega} p(T^{t_k}z)$, and for all $i \in [0, n) \setminus \cup_k [t_k, t_k + p_k)$, $\pi_x \circ T^{t_k}\omega \subset P^+$ for some $P \cap C_\delta = \emptyset$.

(P3') (Distortion estimate) There exists K_2 such that the following hold for all i, n and $\omega \subset l_i$ satisfying (i) all $z \in \omega$ have the same itinerary for $n-1$ iterates and (ii) both ω and $T^n(\omega)$ are free. Then for all $z, z' \in \omega$,

$$\frac{|DT_z^n(\tau(z))|}{|DT_{z'}^n(\tau(z'))|} \ < K_2.$$

A proof of Proposition 9.2 is given in Appendix A.22.

9.3 Distribution of free segments of length $> \delta$

For definiteness, assume l_0 is such that either (i) $\pi_x(l_0) = I_{\mu_0 j_0}$ where $I_{\mu_0 j_0}$ is one of the outermost $I_{\mu j}$ or (ii) $l_0 \cap \mathcal{C}^{(1)} = \emptyset$ and has length $> K^{-1}\delta$. Following the procedure in Sect. 2.3, we introduce on l_0 an increasing sequence of partitions $\mathcal{Q}_0 < \mathcal{Q}_1 < \mathcal{Q}_2 < \cdots$ with $\mathcal{Q}_i$ representing a *canonical subdivision by itinerary*. This means in particular that $\mathcal{Q}_0 = \{l_0\}$, and each $\omega \in \mathcal{Q}_{i-1}$ has the property that all $z \in \omega$ have the same itinerary through step $i-1$ in the sense of Sect. 9.2C. We are particularly interested in those $\omega \in \mathcal{Q}_{i-1}$ for which $T^i\omega$ is free and $|T^i\omega| > \delta$. These are the segments that will be used in our constructions in Sect. 9.1. Observe that (P3') holds for $T^i|_\omega$.

As in Sect. 2.3, let S be the stopping time on l_0 defined by $S(z) = i$ if and only if $T^i(\mathcal{Q}_{i-1}(z))$ is free and has length $> \delta$. We introduce a sequence of stopping times $S_0 < S_1 < S_2 < \cdots$ on l_0 as follows: Let $S_0 = 0$ and $S_1 = S$. On $\omega \in \mathcal{Q}_{i-1}$ such that $S_k|_\omega = i$, we define $S_{k+1}(z)$ to be the smallest $j > i$ such that $T^j(\mathcal{Q}_{j-1}(z))$ is free and has length $> \delta$.

Lemma 9.2 *There exists K_3 such that the following hold for all $k \geq 0$ and $n > K_3 \log \frac{1}{\delta}$: If $\omega \in \mathcal{Q}_{i-1}$ is such that $S_k|_\omega = i$, then*

$$m_0(\omega \cap \{(S_{k+1} - S_k) > n\}) \ < e^{-K_3^{-1}n} \ m_0(\omega).$$

Proof: We claim that results corresponding to those in Sect. 2.3 are valid in the present setting: the proofs in Sect. 2.3 depend only on (P1)–(P3), and they continue to hold as (P1)–(P3) are replaced by (P1')–(P3'). Partitioning $T^i(\omega)$ into segments of length between δ and 3δ, we appeal to Corollary 2.1 and (P3'). $\square$

The next lemma locates sites suitable for the construction of $\mathcal{N}$.

Lemma 9.3 *There exist an interval $\tilde{L} \subset I$, a number $\tilde{c} > 0$, a sequence of integers $n_1 < n_2 < \cdots$, and a collection of segments $\{\tilde{\omega}_j^{(i)}\}$ of l_0 such that*

(i) $\pi_x(T^i(\tilde{\omega}_j^{(i)}) = \tilde{L}$;

(ii) $\frac{1}{n_k}\sum_{i=0}^{n_k-1} m_0(\cup_j \tilde{\omega}_j^{(i)}) \geq \tilde{c}\; m_0(l_0)$.

Proof: (1) *Estimate from below of the total measure of $\omega \in \mathcal{Q}_{i-1}$ with $|T^i(\omega)| > \delta$.* Let $\mathcal{R}_{ik} = \{\omega \in \mathcal{Q}_{i-1} : S_k|_\omega = i\}$. By Lemma 9.2, there exists K'' such that

$$\int_\omega (S_{k+1} - S_k) dm_0 \;\leq K'' m_0(\omega) \quad \text{for all } \omega \in \mathcal{R}_{ik}.$$

Writing $S_n = \sum_{k=0}^{n-1}(S_{k+1} - S_k)$ and summing over all $\omega \in \cup_i \mathcal{R}_{ik}$ for each k, we obtain

$$\int_{l_0} S_n dm_0 \;\leq\; K'' n \cdot m_0(l_0). \tag{11}$$

Let N be a large integer. Applying Chebychev's Inequality to (11), we obtain

$$m_0\{S_{[\frac{1}{2K''}N]} > N\} \;\leq\; \frac{1}{N}\int S_{[\frac{1}{2K''}N]} dm_0 \;\leq\; \frac{1}{N}\;(K''\frac{1}{2K''}N)\; m_0(l_0) \;=\; \frac{1}{2} m_0(l_0).$$

This implies

$$\sum_{i \leq N} m_0(\cup\{\omega \in \cup_k \mathcal{R}_{ik}\}) \;\geq\; \frac{1}{4K''} N\; m_0(l_0). \tag{12}$$

(2) *Selection of $\tilde{L}$.* We partition I into intervals $L_1, L_2, \cdots, L_{\frac{3}{\delta}}$ of length $\frac{1}{3}\delta$ each. For $\omega \in \mathcal{R}_{ik}$, since $|T^i(\omega)| > \delta$, there exists $q = \psi(\omega)$ such that $\pi_x(T^i(\omega)) \supset L_q$. Let $\hat{\omega} = \omega \cap T^{-i}\pi_x^{-1} L_q$. By (P3'), there exists K''' independent of ω such that $m_0(\hat{\omega}) > K'''^{-1} m_0(\omega)$. Together with (12), this implies that for each N, there exists $q(N)$ such that

$$\frac{1}{N}\sum_{i \leq N} m_0(\cup\{\hat{\omega} : \omega \in \cup_k \mathcal{R}_{ik}, \psi(\omega) = q\}) \;\geq\; \frac{\delta}{12K''K'''}\; m_0(l_0).$$

Let $\tilde{L} = L_q$ where $q = q(N)$ for infinitely many N. For each i, then, the collection $\{\tilde{\omega}_j^{(i)}\}$ is $\{\hat{\omega} : \omega \in \cup_k \mathcal{R}_{ik},\; \psi(\omega) = q\}$. $\square$

9.4 Completing the proof of Proposition 9.1

As explained in Sect. 9.1, it suffices to verify Condition **(S)**. Let l_0 be as in Sect. 9.2. Using the notation in Lemma 9.3, we let L be the middle $\frac{9}{10}$ of $\tilde{L}$, and let

$$\mathcal{N} \;:=\; \text{closure}\;\left\{\left(\cup_{i,j>0}\; T^i(\tilde{\omega}_j^{(i)})\right) \cap (L \times D_{m-1})\right\}.$$

It remains is to show that $\mathcal{N}$ is a continuous family of unstable curves.

Lemma 9.4 *Let $M \in \mathcal{T}_n$ be such that $M \cap \mathcal{N} \neq \emptyset$. Then*
(i) there exists $\tilde{\omega}_j^{(i)}$ such that $T^i(\tilde{\omega}_j^{(i)}) \subset M^\circ$;
(ii) $M \cap (L \times D_{m-1})$ is a horizontal section.

Proof: (i) Since $\mathcal{N} \subset \Omega \subset \text{int}(R_n)$, $M \cap \mathcal{N} \neq \emptyset \implies M \cap \gamma \neq \emptyset$ for some $\gamma = T^i(\tilde{\omega}_j^{(i)})$. It remains to show that γ cannot meet $E \cup E'$, the ends of M. This is because all points in $E \cup E'$ are in bound state, while γ is free. (We need to know $\xi_i \in \gamma$ is free viewing the orbit as starting from ξ_{i-n}. For $n > i$, this is true because for $\xi_0 \in l_0$, $d_{\mathcal{C}}(\xi_j) > \frac{1}{2}\delta$ for all $j < 0$.) (ii) By Lemma 8.1, there is, through each $\xi \in \gamma$, a codimension one stable manifold $V(\xi) := T^n(W_n^s(T^{-n}\xi))$. Each $V(\xi)$ has diameter $< Kb^{\frac{n}{5}}$ and spans the cross-section of M, i.e. $\cup_{\xi \in \gamma} V(\xi)$ is a section of R_n. All points in this section are free, so it is a horizontal section of length $\approx \frac{1}{3}\delta$, of which roughly the middle $\frac{9}{10}$ is occupied by $M \cap (L \times D_{m-1})$. □

Verification of Condition (S)

Construction of $\{\mathcal{E}_i\}$. Let $\mathcal{A}_n = \{M \cap (L \times D_{m-1}) : M \in \mathcal{T}_n,\ M \cap \mathcal{N} \neq \emptyset\}$. By Lemma 9.4, elements of $\mathcal{A}_n$ are horizontal sections. We give an algorithm below that selects, for each n, a cover $\mathcal{E}'_n$ of $\mathcal{N}$ by a finite number of pairwise disjoint elements of $\cup_{j \geq n} \mathcal{A}_j$. We then let $\mathcal{E}_n = \{E \cap \mathcal{N} : E \in \mathcal{E}'_n\}$.

Let $\mathcal{E}'_1 = \mathcal{A}_1$, and assume $\mathcal{E}'_{i-1}$ is constructed. To construct $\mathcal{E}'_i$, we first put all $E \in \mathcal{E}'_{i-1}$ of generation $\geq i$ in $\mathcal{E}'_i$. Each $E \in \mathcal{E}'_{i-1}$ of generation $i-1$ is then replaced systematically by elements of $\{A \in \mathcal{A}_j, j \geq i, A \subset E\}$ as follows: first pick all $F \in \mathcal{A}_i$, then pick all $F \in \mathcal{A}_{i+1}$ that cover some points in $\mathcal{N} \cap E$ not yet covered, then pick all $F \in \mathcal{A}_{i+2}$ covering some points not covered before, and so on. Notice at each stage that the branches chosen are pairwise disjoint. Moreover, the process stops in finite time, for every $z \in \mathcal{N}$ lies in some $M \in \mathcal{T}$.

Properties of $\mathcal{E}_\infty = \bigvee_{n>0} \mathcal{E}_n$. First we show that the elements of $\mathcal{E}_\infty$ form a continuous family of C^1 curves. Since every $E \in \mathcal{E}_\infty$ is the nested intersection of a sequence of horizontal sections whose cross-sectional diameters tend to zero, it is the graph of a function $\varphi : L \to D_{m-1}$. By Lemma 9.4, φ is the pointwise limit of a sequence of functions φ_k the graph of each one of which is contained in $T^i(\tilde{\omega}_j^{(i)})$ for some $\tilde{\omega}_j^{(i)}$. Since $T^i(\tilde{\omega}_j^{(i)})$ is a $C^2(b)$ curve, $|\varphi''_k|$ is uniformly bounded for all k; therefore a subsequence φ_{k_i} converges to φ in the C^1 norm.

To see that the curves in $\mathcal{E}_\infty$ are unstable curves, we use the fact that $T^i(\tilde{\omega}_j^{(i)})$ are unstable curves (Lemma 5.5). The uniform derivative estimates along these curves in backward time are passed to the graph of φ, and the distortion estimate in **(S)**(ii) is verified similarly.

Finally, **(S)**(iii) is given by Lemma 9.3. This completes the verification of Condition **(S)** and the proof of Proposition 9.1. □

PART III PARAMETER ISSUES

Let $\mathcal{G} = \cap_{n \geq 0} \mathcal{G}_n$. The purpose of Part III is to prove the existence and abundance of maps in $\mathcal{G}$. More specifically, we will prove that for 1-parameter families $T_a : X \to X$ satisfying the Standing Hypotheses in Section 1, the set $\{a : T_a \in \mathcal{G}\}$ has positive Lebesgue measure. Our plan is to construct a set $\Delta \subset \{a : T_a \in \mathcal{G}\}$ with a generalized Cantor structure in which the gap ratios tend to zero exponentially fast.

We cannot overemphasize the dependence of Part III on earlier sections. Results from Part II on properties of individual $T \in \mathcal{G}$ are clearly relevant as we now seek to identify such maps from a given 1-parameter family. Since the criteria for belonging in $\mathcal{G}$ reside with the behavior of critical orbits, a major focus of the present study is on the evolution of critical curves, i.e. curves of the form $a \mapsto z_i(a), i = 0, 1, 2, \cdots$, where z_0 is a critical point. We will show that $a \mapsto z_i(a)$

define processes that have a great deal in common with the 1D maps studied in Section 2. Part of the analysis involves adapting the results of Section 2 to the present context.

Each of the first three sections of Part III discusses one important aspect of the problem. These ideas culminate in Section 13, which contains the actual construction of Δ.

Hypotheses for Part III: *We assume*
(1) $\{T_a, a \in [a_0, a_1]\}$ *satisfies the Standing Hypotheses in Section 1;*
(2) $[a_0, a_1]$ *is in a sufficiently small neighborhood of* a^* *that* T_a *satisfies the hypotheses at the beginning of Part II for all* $a \in [a_0, a_1]$.
The generic constant K here depends on the family T_a as well as our choice of λ.

10 Dependence of Dynamical Structures on Parameter

Notation such as $\mathcal{T}_k(a)$ and $\mathcal{C}^{(k)}(a)$ are used to indicate dependence on the map T_a.

10.1 Continuation of critical regions and critical points

Definition 10.1 *Let* $J \subset [a_0, a_1]$ *be an interval, and assume that for some* $\hat{a} \in J$, $T_{\hat{a}} \in \mathcal{G}_n$. *We say* $\{T_a, a \in J\}$ *is a* **continuation of** $T_{\hat{a}}$ **in** $\mathcal{G}_n$ *if the following hold:*

(1) For all $a \in J$, $T_a \in \mathcal{G}_n$, *and there is a choice of* $\Gamma_{\theta n}(a)$ *with the following properties:*

(2) The monotone branches of $T_{\hat{a}}$ *of generation* $\leq \theta n$ *deform continuously on* J, *i.e. for each* $k \leq \theta n$, *there is a map* Φ_k *defined on* $J \times \mathcal{T}_k(\hat{a})$ *such that*

(i) for each fixed a, $M \mapsto \Phi_k(a, M)$ *is a bijection between* $\mathcal{T}_k(\hat{a})$ *and* $\mathcal{T}_k(a)$;

(ii) for each fixed M, $a \mapsto \Phi_k(a, M)$ *is continuous (in the Hausdorff metric).*

(3) The critical regions of $T_{\hat{a}}$ *of generation* $\leq \theta n$ *deform continuously on* J, *i.e. for each* $k \leq \theta n$, *there is a map* Ψ_k *defined on* $J \times \{Q^{(k)}(\hat{a})\}$ *such that*

(i) for each fixed a, $Q \mapsto \Psi_k(a, Q)$ *is a bijection between* $\{Q^{(k)}(\hat{a})\}$ *and* $\{Q^{(k)}(a)\}$;

(ii) for each fixed Q, $a \mapsto \Psi_k(a, Q)$ *is continuous.*

(4) The critical points of $T_{\hat{a}}$ *of generation* $\leq \theta n$ *continue smoothly to all of* J, *i.e. for each* $z_0(\hat{a}) = z_0^*(Q^{(k)}(\hat{a}))$, $k \leq \theta n$, $a \mapsto z_0(a)$ *is a* C^2 *curve satisfying*

(i) $z_0(a) = z_0^*(Q^{(k)}(a))$ *where* $Q^{(k)}(a) = \Psi(a, Q^{(k)}(\hat{a}))$;

(ii) if $\xi_1(\hat{a}) = T_{\hat{a}}^{-(k-1)} z_0(\hat{a})$ *and* l *is the* $\mathcal{F}_1$*-leaf containing* $\xi_1(\hat{a})$,[12] *then there is a* C^2*-function* $\xi_1 : J \to l$ *such that* $z_0(a) = T_a^{k-1}(\xi_1(a))$.

We refer to $a \mapsto \Gamma_{\theta n}(a)$ with property (4) as a *coherent choice* of $\Gamma_{\theta n}(a)$.[13]

Observe that if $\{T_a, a \in J\}$ is a continuation of $T_{\hat{a}}$, then (i) $\{\Phi_k\}$ "commutes" with the actions of T_a, i.e. if $M \in \mathcal{T}_k(\hat{a})$ is such that $T_{\hat{a}} M = \cup_{i=1}^s M_i$, $M_i \in \mathcal{T}_{k+1}(\hat{a})$, then $T_a \Phi_k(a, M) = \cup_{i=1}^s \Phi_k(a, M_i)$; and (ii) the partial order on $\{Q^{(k)}, k \leq \theta n\}$ is respected by $\{\Psi_k\}$. The validity of these statements is easily seen by comparing two nearby a.

Our first goal is to give sufficient conditions for the existence of continuations. To ensure that a nontrivial continuation exists, we choose $T_{\hat{a}}$ in the "interior" of $\mathcal{G}_N$. Let

$$\mathcal{G}_N^{\#} = \{T \in \mathcal{G}_N : T \text{ satisfies (A2)}^{\#} \text{ and (A4)}^{\#}\}$$

[12] We assume $\mathcal{F}_1$ is independent of a.

[13] For $T_a \in \mathcal{G}_n$, $\Gamma_{\theta n}(a)$ is determined only up to a finite precision; the exact location of $\Gamma_{\theta n}(a)$ depends on choices of $\mathcal{F}_k$-leaves on which critical points are constructed.

where (A2)$^\#$ and (A4)$^\#$ below require that for all $z_0 \in \Gamma_{\theta N}$ of generation k, the following hold for all $i \le k\theta^{-1}$:

(A2)$^\#$ $d_{\mathcal{C}}(z_i) > \min(\delta, 2e^{-\alpha i})$;

(A4)$^\#$ $|w_i^*| > c_2 e^{\lambda^* i}$ *where* $\lambda^* = \lambda + \frac{1}{100}\lambda_0$.

Clearly, $\mathcal{G}_N^\# \subset \mathcal{G}_N$.

Proposition 10.1 (Dynamical continuation) *There exists $\rho > 0$ depending only on $\|T_a\|_{C^3}$ and c_2 for which the following holds: Assume $T_{\hat{a}} \in \mathcal{G}_N^\#$. For $n \le N$, let $J_n = [\hat{a} - \rho^n, \hat{a} + \rho^n]$. Then $\{T_a, a \in J_n \cap [a_0, a_1]\}$ is a continuation of $T_{\hat{a}}$ in $\mathcal{G}_n$.*

Proof: We assume the following hold for $n < N$ and prove it for $n+1$:

(i) $\{T_a, a \in J_n\}$ is a continuation of $T_{\hat{a}}$ in $\mathcal{G}_n$;

(ii) (*a priori* estimate on $|\frac{d}{da} z_0(a)|$) there is a constant $K > 0$ independent of ρ such that for all $a \in J_n$, if $z_0(a) \in \Gamma_{[\theta n]}(a)$ is of generation j, then

$$\left| \frac{d}{da} z_0(a) \right| < K^j.$$

There is nothing to do if $k\theta^{-1} < n+1 < (k+1)\theta^{-1}$: no new monotone branches or critical regions are constructed, and all critical points of generation $\le k$ are treated in the previous step. We assume therefore that $n+1 = (k+1)\theta^{-1}$ for some k.

1. Coherent choice of $\Gamma_{k+1}(a)$, construction of $\mathcal{C}^{(k+1)}(a)$ and $\mathcal{T}_{k+1}(a)$, and verification of (2) and (3) in Definition 10.1 for objects of generation $k+1$:

For each individual $a \in J_n$, since $T = T_a \in \mathcal{G}_n$, we know by Proposition 8.1 that $\mathcal{C}^{(k+1)}$ and $\mathcal{T}_{k+1}$ can be constructed. Moreover, for each $M \in \mathcal{T}^{(k)}$ and $Q = Q^{(j)}$, $(k+1)(1+2\theta)^{-1} \le j \le k$, there is the following dichotomy: either $TM \cap Q = \emptyset$, or a horizontal section of TM pierces through the entire length of Q .

Now let a vary over J_n. For $M \in \mathcal{T}_k(\hat{a})$, we know by inductive assumption that $M(a) := \Phi_k(a, M)$ varies continuously with a, as does $Q(a) := \Psi_j(Q)$. Since the dichotomy above holds for *all* a – and there is no way to go from one scenario to the other in a continuous manner – it follows that exactly one of the two scenarios must prevail for all $a \in J_n$. Indeed, the number of times $T_a M(a)$ goes through $Q(a)$ is constant for all a. This proves properties (2)(i) and (3)(i) in Definition 10.1.

Suppose for M and Q as above, $T_{\hat{a}}$ has a critical point $\hat{z}_0(a) \in T_{\hat{a}} M(\hat{a}) \cap Q(\hat{a})$. Let $l_k(\hat{a})$ be the connected component of $\mathcal{F}_k$-leave in $M(\hat{a})$ on which $T_{\hat{a}}^{-1} z_0(\hat{a})$ is located. Let $l = T_{\hat{a}}^{-k+1} l_k$. Then l is a leaf of $\mathcal{F}_1$, and $\mathcal{F}_1$ does not depend on a. From the last paragraph, we know that $T_a^k(l)$ pierces through $Q(a)$ for all a. Let $z_0(a)$ be constructed on $T_a^k(l)$. This construction guarantees the continuity of $a \mapsto z_0(a)$ for all critical points of generation $k+1$ and consequently properties (2)(ii) and (3)(ii) in Definition 10.1.

2. Smoothness of $a \mapsto z_0(a)$ and estimate on $|\frac{d}{da} z_0(a)|$ for z_0 of generation $k+1$:

Continuing to use notation from the last paragraph, we let $x \mapsto \gamma(x,a) = (x, \psi(x,a))$ be the curve $T_a^k l$ in $Q^{(k+1)}(a)$, and let $z_0(a) = (\bar{x}(a), \bar{y}(a))$. For each (x,a), we consider in $X_{\gamma(x,a)}$ the 2D plane $S = S(\partial_x \gamma(x,a), \mathbf{v})$ with orthonormal basis $\{u, v\}$ where $u = \partial_x \gamma / |\partial_x \gamma|$ and v points in roughly the same direction as $\mathbf{v}$. Let e_{k+1} be the most contracted direction of DT_a^{k+1} in S. As in Sect. 3.6, we write e_{k+1} as a linear combination of u and v, and let η_{k+1} denote its v-component. Then $\bar{x}(a)$ is defined implicitly by $\eta_{k+1}(\bar{x}(a), a) = 0$, and therefore is C^2 as a function of a. Likewise, $\bar{y}(a) = \psi(\bar{x}(a), a)$ is a C^2 function of a.

The following lemma is proved in Appendix A.23.

Lemma 10.1 *As functions of x and a,*
(a) $\|u\|_{C^2},\ \|v\|_{C^2} < K^{k+1}$*;*
(b) $\|\eta_{k+1}\|_{C^2} < K^{k+1}$.

Corollary 10.1

$$\left|\frac{dz_0(a)}{da}\right|,\ \left|\frac{d^2 z_0(a)}{da^2}\right| \leq K^{k+1}. \tag{13}$$

Proof of Corollary 10.1: Differentiating $\eta_{k+1}(\bar{x}(a), a) = 0$, we obtain

$$\frac{d\bar{x}(a)}{da} = -\frac{\partial_a \eta_{k+1}}{\partial_x \eta_{k+1}}(\bar{x}(a), a). \tag{14}$$

Observe that $|\partial_x \eta_{k+1}| > K^{-1}$: This follows from Lemma 3.7 and the fact that derivative growth along the orbit of $z_0^*(Q^{(k)}(a))$ is passed on to that of $z_0(a)$ via Lemma 3.2. Our claim on the first derivative follows directly from Lemma 10.1(b) and the fact that $\frac{d\bar{y}}{da} = \partial_x \psi \frac{d\bar{x}}{da} + \partial_a \psi$. To estimate the second derivative we differentiate (14) one more time with respect to a, and use again Lemma 10.1. □

3. Proofs of (A2)(n+1) and (A4)(n+1): Let z_0 be a critical point of generation $k+1$. We give details only for step $n+1$: By Corollary 10.1, $|z_0(a) - z_0(\hat{a})| \leq K^{k+1}(2\rho^{n+1})$ for all $a \in J_{n+1}$, so that

$$|z_{n+1}(a) - z_{n+1}(\hat{a})| \ < \ K\|DT\|^{n+1}|z_0(a) - z_0(\hat{a})| \ << \ e^{-\alpha(n+1)} \tag{15}$$

provided ρ is sufficiently small relative to $\|DT\|^{-1}$ and K^{-1}. To finish, we need to deal with the differences between $d_{\mathcal{C}(a)}(\cdot)$ and $d_{\mathcal{C}(\hat{a})}(\cdot)$. Suppose $z_{n+1}(\hat{a}) \in \mathcal{C}^{(1)}$, and $\phi(z_{n+1}(\hat{a})) = z_0^*(Q^{(j)}(\hat{a}))$. Then $j \leq \alpha^*(n+1)\theta$, $z_{n+1}(\hat{a}) \in Q^{(j)}(\hat{a})$, and $T_{\hat{a}}$ has a horizontal section $H(\hat{a})$ extending considerably beyond $Q^{(j)}(\hat{a})$ on both sides. We conclude from the continuous deformation of structures of generation j, the estimate $|z_0^*(Q^{(j)}(\hat{a})) - z_0^*(Q^{(j)}(a))| \leq K^j|\hat{a} - a|$ and (15) above that $z_{n+1}(a)$ is either in $Q^{(j)}(a)$ or it is in $H(a)$ and just outside of $Q^{(j)}(a)$, and $d_{\mathcal{C}_a}(z_{n+1}(a)) > \frac{1}{2} d_{\mathcal{C}_{\hat{a}}}(z_{n+1}(\hat{a})) > e^{-\alpha(n+1)}$.

To prove (A4)(n+1), we first convert the problem to one involving $|w_i|$, thereby picking up some factors of $e^{\alpha i}$. The comparability of $|w_i(a)|$ and $|w_i(\hat{a})|$ is given by the following lemma, the proof of which follows closely that of Lemma 3.2 and is omitted.

Lemma 10.2 *Let $z_0(\hat{a})$ be of generation $j \leq \theta(n+1)$. For $i \leq n+1$, let $w_i(\hat{a}) = (DT^i_{\hat{a}})_{z_0(\hat{a})}\mathbf{v}$, and $w_i(a) = (DT^i_a)_{z_0(a)}\mathbf{v}$, $a \in J_{n+1}$. Then $|w_i(a)| \geq \frac{1}{2}|w_i(\hat{a})|$.*

This completes the proof of (i) and (ii) for step $n+1$. □

10.2 Properties of $a \mapsto z_0(a)$, $z_0 \in \cup\Gamma_j$

Unlike the situation in 1D, $\cup_{j\geq 1}\Gamma_j$ is an infinite set, and the domains of definition of $a \mapsto z_0(a)$ decrease as the generation of z_0 increases. For $T_{\hat{a}} \in \mathcal{G}_N^{\#}$ and $z_0 \in \Gamma_{\theta N}(\hat{a})$ of generation θn, $n \leq N$, we guarantee the continuation of $z_0(\hat{a})$ only to the interval $J_n = [\hat{a} - \rho^n, \hat{a} + \rho^n]$. We claim, however, that there is a uniform bound on $\frac{d}{da}z_0(a)$ that is valid on for all z_0 (independent of generation) on their respective intervals of continuation.

Lemma 10.3 *Let $\hat{a}$, J_n and $\Gamma_{\theta n}$ be as in Proposition 10.1, and let $a \mapsto z_0^{(k)}(a), a \in J_n$, be a curve of critical points of generation $k \leq [\theta n]$. Then there is $z_0^{(k')}$ of generation k', $k' < k \leq k'(1+2\theta)$, with $z_0^{(k)} \in Q^{(k')}(z_0^{(k')})$ such that*

$$\left|\frac{d}{da}(z_0^{(k)}(a) - z_0^{(k')}(a))\right| < b^{\frac{k'}{9}}.$$

A proof of Lemma 10.3 is given in Appendix A.23.

Corollary 10.2 *Under the hypotheses of Lemma 10.3, there exists K_1 such that for every curve of critical points $a \mapsto z_0^{(k)}(a)$, $k \le \theta n$, if $z_0^{(k)} = (x_0, y_0)$, then*

$$|\frac{d}{da}x_0(a)| \le K_1, \qquad |\frac{d}{da}y_0(a)| \le b^{\frac{1}{10}}.$$

Proof: Let $z_0 = z_0^*(Q^{(k)})$, and suppose $Q^{(k)} = Q^{(k_0)} \subset Q^{(k_1)} \subset \cdots \subset Q^{(1)}$ are consecutive critical regions. We obtain, by comparing $a \mapsto z_0^{(k_i)}$ and $a \mapsto z_0^{(k_{i-1})}$, that $|\frac{d}{da}(z_0^{(k)}(a) - z_0^{(1)}(a))| < b^{\frac{k'}{10}}$. The assertions in this corollary now follow immediately from properties of the finitely many critical points $z_0^{(1)} = (x_0^{(1)}, y_0^{(1)})$ of generation one, namely $\frac{d}{da}y_0^{(1)} = 0$ and $|\frac{d}{da}x_0^{(1)}| < \frac{1}{2}K_1$ for some K_1. □

Remark Lemma 10.3 and Corollary 10.2 together imply the following: (1) The speeds of movement of all critical points are uniformly bounded. (2) While Γ_n as a whole moves with speed $\mathcal{O}(1)$, the *relative speed* of motion of z_0' and z_0 for $z_0' \in Q^{(k)}(z_0)$ decreases exponentially very fast with k.

10.3 Setting for the analysis to follow

Recall that for a single map $T \in \mathcal{G}_N$, whether or not T is in $\mathcal{G}_{\frac{1}{\alpha^*}N}$ is determined by whether (A2) and (A4) are satisfied up to time $\frac{1}{\alpha^*}N$. We now consider a family $\{T_a, a \in J\}$ where $T_a \in \mathcal{G}_N$ for all a. In addition to asking whether $T_a \in \mathcal{G}_{\frac{1}{\alpha^*}N}$ for each individual a, we will also want to know for what fraction of $a \in J$ is $T_a \in \mathcal{G}_{\frac{1}{\alpha^*}N}$.

This leads us to study the evolution of $\zeta_i : a \mapsto z_i(a)$ where ζ_0 is a coherent choice of $z_0 \in \Gamma_{\frac{1}{\alpha^*}\theta N}$. The analysis is highly inductive: For each T_a, critical points are defined inductively, and the presence of certain structures is needed to track their orbits. We wish now to track not single orbits but entire curves. Precise conditions under which this analysis will be carried out are as follows:

Assumptions in the inductive analysis of critical curves

Let $J \subset [a_0, a_1]$ be a parameter interval.

(C1) *$\{T_a, a \in J\}$ is a continuation (of some $T_{\hat{a}}$) in $\mathcal{G}_N$.*

(C2) *A coherent choice of $\Gamma_{\frac{1}{\alpha^*}\theta N}(a)$, $a \in J$, has been made.*

Clarification 1. (C1) and (C2) are related as follows: If, in addition to (C1), we have $T_{\hat{a}} \in \mathcal{G}_N^{\#}$ for some $\hat{a} \in J$, then steps 1 and 2 in Proposition 10.1 can be carried out for critical points of generation k for all $k \le \frac{1}{\alpha^*}\theta N$. In Sections 11 and 12 we are not concerned with how (C2) comes about, but note that once a coherent choice of $\Gamma_{\frac{1}{\alpha^*}\theta N}(a)$ is made, the estimates in Lemma 10.3 and Corollary 10.2 are valid for the critical points in question. Justifications for this claim follow *verbatim* those in Sects. 10.1 and 10.2.

2. Nothing is assumed or claimed at this point about the behavior of critical orbits beyond time N. That *is* the objective of the investigation in the pages to follow. More precisely, we will be concerned with the evolution of

$$\zeta_i : a \mapsto z_i(a) \quad \text{for} \quad a \in J, \quad z_0 \in \Gamma_{\frac{1}{\alpha^*}\theta N} \setminus \Gamma_{\theta N}, \quad i \le \frac{1}{\alpha^*}N$$

with particular interest in the time range $N < i \le \frac{1}{\alpha^*}N$.

3. As we will see, this analysis requires that all critical structures of generation $\le \theta N$ exist and vary with parameter in a certain way. This is provided by (C1) and Corollary 10.2. We remark also that critical structures beyond generation θN are not relevant for this analysis.

11 Dynamics of Curves of Critical Points

The aim of this section is to bring to light a certain resemblance between the evolution of ζ_i and that of certain "horizontal" curves under the iteration of T_a for a fixed a. The reason behind this resemblance is that $|\frac{d}{da}\zeta_i|$ is comparable to $|DT^i(\mathbf{v})|$. We will show that for as long as $T_a \in \mathcal{G}_N$, this comparability self-perpetuates once we get it going, and the start-up mechanism is provided by the parameter transversality condition in the Standing Hypotheses in Section 1. This is discussed in Sect. 11.1. Basic properties in the evolution of ζ_i, such as bound and free periods, are discussed in Sect. 11.2.

Conditions (C1) and (C2) in Sect. 10.3 are assumed throughout.

11.1 Equivalence of space- and a-derivatives

We use the notation $z_i(a) = T_a^i(z_0(a))$, $w_i(a) = (DT_a^i)_{z_0(a)}(\mathbf{v})$ and $\tau_i(a) = \frac{d}{da}z_i(a)$.

Proposition 11.1 *There exist $\hat{K} > 0$ and $i_0 \in \mathbb{Z}^+$ (both depending only on $\{F_a\}$), such that the following holds for all (a,b) sufficiently close to $(a^*,0)$: Fix $z_0 \in \Gamma_{\frac{1}{\alpha^*}\theta N} \setminus \Gamma_{\theta N}$. We assume that for some $n \le \frac{1}{\alpha^*}N$, $d_{\mathcal{C}}(z_i(a)) \ge \min(\delta, e^{-\alpha i})$ for all $a \in J$ and $i < n$. Then*

$$\hat{K}^{-1} \le \frac{|\tau_i|}{|w_i|} \le \hat{K} \quad \text{for all} \quad i_0 < i \le n.$$

We remark that the required proximity of (a,b) to $(a^*,0)$ depends also on i_0, and that under the conditions above, the pair (z_0, w_0) is controlled by $\Gamma_{\theta N}$ up to time n (see Proposition 6.2).

Proposition 11.1 is a consequence of Standing Hypothesis (b) in Section 1 and can be viewed as the higher dimensional analog of Proposition 2.3 in Sect. 2.4. Recall from Proposition 2.3 that

$$\lim_{k\to\infty} \sum_{s=1}^{k} \frac{\frac{d}{da}(f_a(x_{s-1}))(a^*)}{(f^{s-1})'(x_1(a^*))} \;=\; \left[\frac{d}{da}f_a(\hat{x}(a)) - \frac{d}{da}q(a)\right]_{a=a^*} \;:=\; \hat{c}. \tag{16}$$

The constant $\hat{K}$ in Proposition 11.1 is derived from $\hat{c}$ together with angle and other considerations.

Proof: Letting $\psi(z) = \frac{\partial}{\partial a}(T_a z)$, we write

$$\tau_i = DT_{z_{i-1}}\tau_{i-1} + \psi(z_{i-1}) = DT_{z_0}^i\tau_0 + \sum_{s=1}^{i} DT_{z_s}^{i-s}\psi(z_{s-1}) := I + II$$

where

$$I = DT_{z_0}^i\tau_0 + \sum_{s=1}^{i_0} DT_{z_s}^{i-s}\psi(z_{s-1}) \qquad \text{and} \qquad II = \sum_{s=i_0+1}^{i} DT_{z_s}^{i-s}\psi(z_{s-1}),$$

i_0 being a number to be determined. We will show there exist K_0 (depending only on $\{T_a\}$) and i_0 such that if (a,b) is sufficiently near $(a^*,0)$, then for $i_0 < i \le n$,

- $K_0^{-1}|\hat{c}| < \frac{|I|}{|w_i|} < K_0|\hat{c}|$ and
- $\frac{|II|}{|w_i|}$ is as small as we wish.

These estimates together give the desired result.

Estimate on $\frac{|II|}{|w_i|}$: Since $|\psi(\cdot)| < K$, it follows from Lemma 5.6 that

$$|II| \;\le\; K\sum_{s=i_0+1}^{i} \|DT_{z_s}^{i-s}\| \;\le\; K\sum_{s=i_0+1}^{i} Ke^{-\hat{\lambda}s}|w_i|.$$

Choosing i_0 large enough, we can make $K^2 \sum_{s=i_0+1}^{\infty} e^{-\hat{\lambda} s} << K_0^{-1}|\hat{c}|$.

Estimate on $\frac{|I|}{|w_i|}$*:* Increase i_0 if necessary so that with $k = i_0$, the sum on the left side of (16) is $< \frac{1}{2}\hat{c}$ from its limit; i_0 is fixed from here on. Let V be such that $I = DT_{z_{i_0}}^{i-i_0} V$, i.e.

$$V = DT_{z_0}^{i_0} \tau_0 + \sum_{s=1}^{i_0} DT_{z_s}^{i_0 - s} \psi(z_{s-1}).$$

The verification of Lemma 11.1 is given in Appendix A.24.

Lemma 11.1 *As* $(a, b) \to (a^*, 0)$*,*

$$\frac{|w_1|}{|w_{i_0}|} \, V \;\to\; \left(\pm \sum_{s=1}^{i_0} \frac{\frac{d}{da}(f_a(x_{s-1}))(a^*)}{(f^{s-1})'(x_1(a^*))}, \; 0 \right).$$

It is important to note that the convergence above is uniform among all critical curves. This is evident from the uniform bound on $\frac{d}{da} z_0$ and the proof of Lemma 11.1.

To finish, we write

$$\frac{|I|}{|w_i|} \;=\; \frac{|DT_{z_{i_0}}^{i-i_0}(V)|}{|DT_{z_{i_0}}^{i-i_0}(w_{i_0})|} \;=\; \frac{|DT_{z_{i_0}}^{i-i_0}(\frac{V}{|V|})|}{|DT_{z_{i_0}}^{i-i_0}(\frac{w_{i_0}}{|w_{i_0}|})|} \cdot \frac{|V|}{|w_{i_0}|}.$$

Notice first that by Lemma 11.1, $\frac{1}{3}\frac{\hat{c}}{|w_1|} < \frac{|V|}{|w_{i_0}|} < 2\frac{\hat{c}}{|w_1|}$. We claim that

$$K^{-1} \;<\; \frac{|DT^{i-i_0}(\frac{V}{|V|})|}{|DT^{i-i_0}(\frac{w_{i_0}}{|w_{i_0}|})|} \;<\; K \tag{17}$$

provided (a, b) is sufficiently near $(a^*, 0)$. Assume $\angle(V, w_{i_0}) \neq 0$, and let $e_{j-i_0} = e_{j-i_0}(S)$ be the most contracted vector of order $j - i_0$ where $S = S(V, w_{i_0})$. To prove (17), it suffices to show that e_{j-i_0} is well-defined, $\angle(e_{j-i_0}, w_{i_0}) > K^{-1}$, and $\angle(V, w_{i_0}) = \mathcal{O}(b)$. With (a, b) sufficiently near $(a^*, 0)$, we may assume for some $s_0 >> i_0$ that $d_{\mathcal{C}}(z_s) > \frac{1}{2}\delta_0$ for all $s < s_0$ (δ_0 is as in Definition 1.1). This together with our assumption that $d_{\mathcal{C}}(z_j) \geq \min(\delta, e^{-\alpha j})$ implies that for all $j > i_0$, $|w_j|/|w_{i_0}| > K^{-1} e^{(\lambda' - 2\alpha)(j - i_0)}$, proving e_{j-i_0} is well-defined. Since w_{i_0} is b-horizontal, we have $\angle(e_1, w_{i_0}) > K^{-1}$ by Lemma 3.7. This together with $\angle(e_1, e_{j-i_0}) < (Kb)^{j-i_0}$ (Lemma 3.1) gives $\angle(e_{j-i_0}, w_{i_0}) > K^{-1}$. As for V, Lemma 11.1 tells us its slope is small as we wish. Hence $\angle(V, w_{i_0}) = \mathcal{O}(b)$. □

We remark that all the estimates in the proof above – and hence the constants in the statement of the proposition – are independent of N. For as long as both τ_i and w_i grow in magnitude, the angles between them must shrink by rank one arguments. The assumptions in the next lemma are as in Proposition 11.1. A detailed proof is given in Appendix A.24

Lemma 11.2 *If* z_i *is a free return, then* $\angle(\tau_i, w_i) < \frac{K}{|\tau_i|}$.

The following are assumed for the rest of this paper: *(i)* i_0 *is sufficiently large, (ii)* (a, b) *is sufficiently close to* $(a^*, 0)$*, and (iii) all critical points stay at distances* $> \frac{1}{2}\delta_0$ *away from* $\mathcal{C}$ *for* $>> i_0$ *iterates – where "sufficiently large", "sufficiently close" and ">>" are as required in the proof of Proposition 11.1.*

11.2 Resemblance to phase-space dynamics

In addition to (C1) and (C2), we now fix $z_0 \in \Gamma_{\frac{1}{\alpha^*}\theta N} \setminus \Gamma_{\theta N}$ and impose on it

(C3) *For some* $i_0 < n \leq \frac{1}{\alpha^*}N$, $d_{\mathcal{C}}(z_i) > \min(\delta, e^{-\alpha i})$ *for all* $i \leq n$.

We discuss below 4 aspects of the dynamics of $\zeta_i : a \mapsto z_i(a), i = 0, 1, 2, \cdots$. The notation is as in Sect. 11.1; in particular, $w_i(a) = (DT_a^i)_{z_0(a)}(\mathbf{v})$ and $\tau_i(a) = \frac{d}{da}z_i(a)$. Recall also that $s(u) = \frac{|u_y|}{|u_x|}$ where $u = (u_x, u_y)$ is a vector in $\mathbb{R}^m = \mathbb{R} \times \mathbb{R}^{m-1}$.

A. Outside of $\mathcal{C}^{(1)}$

For the first i_0 iterates, we do not have a great deal of information on τ_i. Let $\hat{\varepsilon} := \hat{K}c_2 e^{-\frac{1}{4}\lambda_0 i_0} + \mathcal{O}(b)$ where $\hat{K}$ is as in Proposition 11.1. We may assume $\hat{\varepsilon} << \delta$.

Lemma 11.3 *The following hold for every a:*
(a) If z_n is free, then $s(\tau_n) < \hat{\varepsilon}$.
(b) If z_n is free, and $z_{n+j} \notin \mathcal{C}^{(1)}$ $\forall\ 0 \leq j < j_0$, then
(i) $|\tau_{n+j}| > K^{-1}\delta e^{\frac{1}{4}\lambda_0 j}|\tau_n|$ for $j \leq j_0$; and
(ii) if in addition $\gamma_{n+j_0} \in \mathcal{C}^{(1)}$, then $|\tau_{n+j_0}| > K^{-1}e^{\frac{1}{4}\lambda_0 j_0}|\tau_n|$.

Proof: (a) follows from Lemma 11.2 and the b-horizontal property of w_n. As for (b), since $|\tau_{n+j}| >> 1$, we have $|\tau_{n+j+1}|/|\tau_{n+j}| \approx |f'(x_{n+j})|$, $z_i = (x_i, y_i)$. The assertions follow by a proof similar to that of Lemma 3.5. □

Remark We do not claim that free segments of ζ_i are $C^2(b)$, only that they are roughly horizontal (because $\hat{\varepsilon} << 1$). No effort will be made to control (z_0, τ_0). Information on τ_i is obtained instead through comparisons with w_i via Proposition 11.1 and Lemma 11.2.

B. Geometry of critical curves inside $Q^{(1)}$

Let ω be a subinterval of J. We assume $\zeta_n(\omega)$ is free (meaning $z_n(a)$ is free for each a), and $\zeta_n(\omega) \subset Q^{(1)}$. For each individual a, we have seen in Part II how $z_n(a)$ is related to the critical structure of T_a. We now describe the geometric relationship between the *curve* ζ_n and the 1-parameter family of critical structures.

By Lemma 11.3(a), ζ_n is roughly horizontal. Consider an arbitrary point $\hat{a} \in \omega$, and assume that $\zeta_n(\hat{a})$ lies in the interior of $Q^{(j)}(\hat{a})$ for some $j \leq n$. We assume, for definiteness, that as a increases, we move right along ζ_n. By continuity, for all a in a neighborhood of $\hat{a}$, $\zeta_n(a)$ also lies in the interior of $Q^{(j)}(a)$. Let $\bar{a}$ be the first a for which the last statement is not valid. Then $\zeta_n(\bar{a}) \in \partial Q^{(j)}(\bar{a})$. Since it cannot be in ∂R_j (because $n \geq j$), it has to lie in the right (vertical) boundary of $Q^{(j)}(\bar{a})$.

We claim that as a increases, ζ_n crosses $Q^{(j)}(a)$ in exactly one point, i.e. for all $a > \bar{a}$, $\zeta_n(a) \notin Q^{(j)}(a)$. This is because for $a \in \omega$, $|\frac{d}{da}\zeta_n| > \hat{K}^{-1}|w_n| > K^{-1}e^{\lambda n}$ (Proposition 11.1 and Lemma 5.2), while $|\frac{d}{da}z_0^*(Q^{(j)})| < K_1$ (Corollary 10.2). Since $z_0^*(Q^{(j)})$ and the "vertical" boundaries of $Q^{(j)}$ move in the horizontal direction at the same speed, and we may assume n is large enough that $K^{-1}e^{\lambda n} >> K_1$, $a \mapsto \zeta_n(a)$ crosses $\cup_a \partial Q^{(j)}(a)$ transversally in $R_1 \times J$ in exactly one point.

The picture can therefore be summarized as follows. Let j_0 be the largest $j \leq \frac{1}{\alpha^*}\theta n$ such that ζ_n meets $Q^{(j)}$, and assume $\zeta_n(\hat{a})$ lies in the interior of $Q^{(j_0)}$. Let $Q^{(j_0)} \subset Q^{(j_1)} \subset Q^{(j_2)} \cdots$ be consecutive critical regions starting from $Q^{(j_0)}$ for $T_{\hat{a}}$. This structure, as we know, is identical for all the T_a. As a increases (or decreases), this nested structure moves with speed $\mathcal{O}(1)$, which is *very slow* relative to the speed of $a \mapsto \zeta_n(a)$. Thus from the point of view of ζ_n, the critical structure appears stationary, and the picture resembles that of a single map.

C. Bound period and recovery

The setting is as in B above. For each a, we have defined for the map T_a the notion of $\phi(z)$ for $z = \zeta_n(a)$, $d_{\mathcal{C}}(z) := |\phi(z) - z|$, and $p(z)$. To emphasize their dependence on the map T_a, we now write $\phi_a(z)$, $d_{\mathcal{C}(a)}(z)$ and $p_a(z)$. For purposes of studying the evolution of ζ_n, we can, if we so choose, use the definitions associated with each individual a for $\zeta_n(a)$. For aesthetic as well as practical reasons (to become clear in the next section), we prefer to have some coherence along ζ_n, even if this involves some small modifications in the definitions above. We explain how this can be done:

Step 1. Choosing a common guiding critical point $\phi(\omega)$. The choice is quite arbitrary. Let j_0 be the largest $j \leq \frac{1}{\alpha^*}\theta n$ such that ζ_n meets $Q^{(j)}$, and pick $\hat{a}$ with $\zeta_n(\hat{a}) \in Q^{(j_0)}$. Let $\phi(\omega) := z_0^*(Q^{(j_0)})(\hat{a})$, and define $d_{\mathcal{C}}(z) = |z - \phi(\omega)|$ for $z \in \zeta_n(a)$.

Lemma 11.4 *For $z = \zeta_n(a)$, let j be the generation of $\phi_a(z)$. Then*

$$|d_{\mathcal{C}(a)}(z) - d_{\mathcal{C}}(z)| \leq b^{\frac{j}{4}} + K_1|a - \hat{a}|.$$

Proof: Let $j_0 > j_1 > \cdots$ be as in B. Then $z \in Q^{(j_i)}(a)$ for some $i \geq 0$. By definition, $j \leq j_i$, and so $Q^{(j_0)} \subset Q^{(j)}$. We then have

$$\begin{aligned} |d_{\mathcal{C}(a)}(z) - d_{\mathcal{C}}(z)| &= |z_0^*(Q^{(j)})(a) - z_0^*(Q^{(j_0)})(\hat{a})| \\ &\leq (|z_0^*(Q^{(j)})(a) - z_0^*(Q^{(j_0)})(a)|) + (|z_0^*(Q^{(j_0)})(a) - z_0^*(Q^{(j_0)})(\hat{a})|) \\ &\leq b^{\frac{j}{4}} + K_1|a - \hat{a}|, \end{aligned}$$

the first term in the last inequality is by Lemma 4.1 and the second by Corollary 10.2. □

The first term in the error above is innocuous. Since $|a - \hat{a}| < \hat{K}e^{-\lambda n}|\zeta_n(a) - \zeta_n(\hat{a})|$, the second term is negligible if $d_{\mathcal{C}(a)}(z)$ is $>> K_1\hat{K}e^{-\lambda n}\delta$. In particular, for z with $d_{\mathcal{C}(a)}(z) > e^{-\alpha n}$, we have $d_{\mathcal{C}(a)}(z) \approx d_{\mathcal{C}}(z)$.

Step 2. Definition of a new bound period $p(\cdot)$. Let $\mathcal{P}^\omega$ be the partition $\mathcal{P}$ in Sect. 2.2 centered at $\pi_x(\phi(\omega))$, and let $\hat{\omega}$ be such that $\pi_x(\zeta_n(\hat{\omega})) \approx I_{\mu j}$. We define

$$p(\hat{\omega}) \;=\; \min_{a \in \hat{\omega}} p_a(\zeta_n(a)).$$

For this definition to be meaningful, we must verify that it has the properties of bound periods in the sense of $p_a(\cdot)$ for a single map. The next lemma, the proof of which is given in Appendix A.25, assures us this is the case.

Lemma 11.5 *Let $\hat{\omega}$ be as above, and let $p = p(\hat{\omega})$. Then*
(a) $K^{-1}\mu < p < K\mu$;
(b) for $a, a' \in \hat{\omega}$ and $j < p$, $|\zeta_{n+j}(a) - \zeta_{n+j}(a')| < 2e^{-\beta j}$;
(c) $|\tau_{n+p}(a)| > K^{-1}e^{\frac{p}{4}}|\tau_n(a)|$ for all $a \in \hat{\omega}$;
(d) $\zeta_{n+p}(a)$ is out of all splitting periods, $s(\tau_{n+p}(a)) < \hat{\varepsilon}$;
(e) $|\zeta_{n+p}(\hat{\omega})| \geq \frac{1}{\mu^2}e^{-K\alpha\mu}$.

D. Decomposition into bound and free states

With bound periods defined, we may now assign a bound or free state to each $\zeta_i(a)$ in the evolution of ζ_i, namely that $\zeta_i(a)$ is free if it is not in a bound period as defined in C. We remark that this notion is not necessarily consistent with the one for a single map T_a. Indeed, we must now go back to rectify the following statements: In Lemma 11.3, and in the setting of B and C, the word "free" as stated refers to "free" in the sense of individual maps. We leave it to the reader to check that these statements are, in fact, valid if "free" is given the meaning in this paragraph.

12 Derivative growth via Statistics

This section is about how to deal with (A4) (see Sect. 4.1). We focus on one critical point at a time, and discuss (i) what it takes to maintain regular derivative growth along its orbit, and (ii) why one should expect the conditions guaranteeing this growth to be satisfied by a positive measure set of parameters.

12.1 Estimating $|w_i^*|$ in terms of itinerary

We return in this subsection to the dynamics of a single map to motivate a few ideas. As explained earlier, (A4) is not a self-perpetuating property. We now give a condition in terms of the itinerary of z_i that guarantees sustained exponential growth of $|w_i^*(z_0)|$.

Consider for definiteness $T \in \mathcal{G}_N$, and assume that for some $z_0 \in \Gamma_{\frac{1}{\alpha^*}\theta N} \setminus \Gamma_{\theta N}$, $d_{\mathcal{C}}(z_i) > e^{-\alpha i}$ for all $i \le n$, $n \le \frac{1}{\alpha^*}N$. We know from Corollary 6.1 that for $w_0 = \mathbf{v} \in X_{z_0}$, $|w_i^*| > K^{-1}e^{(\frac{1}{3}\lambda - 2\alpha)i}$ for all $i \le n$. Let us examine more closely how this growth comes about. Let $t_1 < t_1 + p_1 \le t_2 < t_2 + p_2 \le t_3 < t_3 + p_3 \le \cdots$ be such that t_i are the consecutive free return times up to time n and p_i the lengths of the ensuing bound periods. Then we have

$$\text{(i)} \quad \frac{|w^*_{t_{i+1}}|}{|w^*_{t_i+p_i}|} \ \ge\ c_2 e^{\lambda(t_{i+1}-(t_i+p_i))}, \qquad \text{(ii)} \quad \frac{|w^*_{t_i+p_i}|}{|w^*_{t_i}|} \ \ge K^{-1}e^{\frac{1}{3}\lambda p_i}.$$

Observe that in (i), the exponent is, in reality, the "outside exponent" $\frac{1}{4}\lambda_0$, which is strictly $> \lambda$. As for (ii), the guaranteed growth rate of $\frac{1}{3}\lambda$ does not contribute much to maintaining a Lyapunov exponent of λ. It is no significant loss if we replace it by the weaker estimate $\frac{|w^*_{t_i+p_i}|}{|w^*_{t_i}|} \ge c_2^{-1}$, which is what we will do.

We continue to reason as follows: If the fraction of time z_i spends in bound periods between time 0 and n is $< \sigma$, and z_n is not in a bound period, then $|w_n^*| > \mathit{const}\ e^{(1-\sigma)\frac{1}{4}\lambda_0 n}$. This number is $> e^{\lambda n}$ if σ is sufficiently small. Since the "outside exponent" does not decrease as δ decreases (Lemma 2.1), it is logical to attempt to decrease σ by decreasing δ, the idea being that some of the time intervals that are bound periods for the original δ would no longer be counted as bound periods for a smaller δ. We summarize the conclusion of this reasoning in the following lemma:

Let $B(\hat{\delta}; 0, n)$ be the total time between 0 and n during which z_i spends in bound periods initiated from visits to the region $d_{\mathcal{C}}(\cdot) < \hat{\delta}$.

Lemma 12.1 *Let z_0 be as above, and let $\sigma > 0$ and $0 < \hat{\delta} \le \delta$. If $B(\hat{\delta}; 0, n) < \sigma n$, then*

$$|w_n^*| > K^{-1}\hat{\delta}e^{[(1-\sigma)\frac{1}{4}\lambda_0 - 3\alpha]n}.$$

Proof: Consider first the case where z_n is free. Let $\hat{t}_1 < \hat{t}_1+\hat{p}_1 \le \hat{t}_2 < \hat{t}_2+\hat{p}_2 \le \cdots \le \hat{t}_k+\hat{p}_k \le n$ be such that $\hat{t}_1, \cdots, \hat{t}_k$ are the consecutive free return times to $\{d_{\mathcal{C}}(\cdot) < \hat{\delta}\}$. Then

$$|w_n| \ =\ \frac{|w_n|}{|w_{\hat{t}_k+\hat{p}_k}|} \ \cdots\ \frac{|w_{\hat{t}_2}|}{|w_{\hat{t}_1+\hat{p}_1}|} \ \frac{|w_{\hat{t}_1+\hat{p}_1}|}{|w_{\hat{t}_1}|}\ |w_{\hat{t}_1}|.$$

We have $\frac{|w_n|}{|w_{\hat{t}_k+\hat{p}_k}|} > c_2\hat{\delta}e^{\frac{1}{4}\lambda_0(n-\hat{t}_k-\hat{p}_k)}$ by Lemma 3.5(i) with $\hat{\delta}$ in the place of δ, and $\frac{|w_{\hat{t}_{i+1}}|}{|w_{\hat{t}_i+\hat{p}_i}|} > c_2e^{\frac{1}{4}\lambda_0(\hat{t}_{i+1}-\hat{t}_i-\hat{p}_i)}$ by Lemma 3.5(ii). To cancel c_2 we use $\frac{|w_{\hat{t}_i+\hat{p}_i}|}{|w_{\hat{t}_i}|} > c_2^{-1}$, which is a trivial consequence of (P2')(ii). This gives $|w_n^*| > K^{-1}\hat{\delta}e^{[(1-\sigma)\frac{1}{4}\lambda_0]n}$ since $\hat{p}_1 + \cdots + \hat{p}_k \le \sigma n$ by assumption. The factor $-3\alpha n$ is needed if n is not free; see Lemma 5.4. $\square$

In this lemma, we think of $\hat{\delta}$ as possibly $<< \delta$, and the factor $\hat{\delta}$ as the price to pay due to the greater nonuniformness in growth properties "outside" – where "outside" now means $d_{\mathcal{C}}(\cdot) > \hat{\delta}$. As we will see, this factor will be absorbed into the initial growth if the critical orbit remains outside of $\mathcal{C}^{(1)}$ for a long time.

We point out that a quantity similar to $B_n(\hat{\delta}; 0, n)$ has already appeared in the context of 1D maps; see Proposition 2.2. Our next step is to make a connection to this proposition.

12.2 Processes defined by curves of critical points

In Sect. 9.3, we borrowed some results from Sect. 2.3 for the dynamics of $T \in \mathcal{G}$ on unstable manifolds – after establishing a strong resemblance between $T^k|_{W^u}$ and iterated 1D maps. Sect. 11.2 suggests that this similarity can perhaps be extended to $\zeta_i : a \mapsto z_i(a)$. But the relation between the "dynamics" of critical curves and iterated 1D maps is a little more tenuous. For one thing, there is no reasonable description of global geometry for critical curves: even though ζ_0 is defined on an interval, it is inevitable that one will lose control of ζ_n on parts of this interval as n increases.

We claim, nevertheless, that the statistical results in Sect. 2.3 are valid. To see that, we return to Section 2 to examine the situation more closely:

Observations 1. In order to apply the results in Sect. 2.3 to critical curves, we must verify for them estimates analogous to (P1)–(P3) in Sect. 2.2 and fix a definition of canonical subdivision by itinerary with properties identical to that in Sect. 2.3.

2. Once that is done, we may proceed letting $\gamma_i = \zeta_i$ wherever it makes sense. It is not important in Sect. 2.3 for γ_i to be $= f^i$ or the ith iterate of any map; $\{\gamma_i\}$ could have been a *process*, meaning a sequence of maps from J to I.

3. Finally, as noted at the end of Sect. 2.3, the stated results are entirely unaffected if we choose to stop considering any element of $\mathcal{Q}_i$ at any time by simply setting $\zeta_j = *$ for all $j \geq i$. Here, the symbol $*$ will correspond to deleted parameters.

Definition of a process $\{\gamma_i\}$ associated with $\zeta_i : a \mapsto z_i(a)$

We assume (C1) and (C2) on an interval J, and fix $z_0 \in \Gamma_{\frac{1}{\alpha^*}\theta N} \setminus \Gamma_{\theta N}$. Associated with z_0, we seek to define a sequence of maps

$$\gamma_i : J \to R_1 \cup \{*\} \qquad \text{for} \qquad 0 \leq i \leq \frac{1}{\alpha^*} N$$

with the property that $\gamma_i(a) = \zeta_i(a) = z_i(a)$ whenever $\gamma_i(a) \neq *$. Here as in Sect. 2.3, $*$ is the "garbage symbol": once $\gamma_i(a) = *$ for some i, $\gamma_j(a) = *$ for all $j > i$; that is to say, we stop considering $a \in J$ from that point on. To facilitate the description of γ_i, we first introduce the following language:

Let $\gamma : \omega \to R_1$ be such that $\frac{d}{da}\gamma$ is nonzero and roughly horizontal. We introduce on ω a partition that will be referred to as the "*canonical partition* defined by γ". This partition is the pullback of the following partition on $\gamma(\omega)$: First we divide $\gamma(\omega) \setminus \mathcal{C}^{(1)}$ into intervals of length $\approx \delta$ each (except possibly for the end interval(s), which may be shorter), and partition each component of $\gamma(\omega) \cap \mathcal{C}^{(1)}$ into $\{\pi_x^{-1} I_{\mu j}\}$ using the guiding critical orbit $\phi(\omega)$ to center the partition $\mathcal{P}$ as is done in Sect. 11.2C. The final partition on $\gamma(\omega)$ is obtained by adjoining the end intervals in the partition above to their neighbors. We say a canonical partition is nontrivial if at least one partition point is introduced. Note that in a nontrivial partition, each element ω' has the property that either $\gamma(\omega') \cap \mathcal{C}^{(1)} = \emptyset$ and $\delta \leq |\gamma(\omega')| \leq 3\delta$, or $\gamma(\omega') \subset \mathcal{C}^{(1)}$ and $\pi_x(\gamma(\omega')) \approx I_{\mu j}$ for some $I_{\mu j}$.[14]

[14]Some fuzziness is allowed in boundary situations due to the adjoining of end intervals.

The canonical subdivision by itinerary for γ_i proceeds as follows. We ignore the first i_0 iterates as they are not particularly meaningful. In general, for $\omega \in \mathcal{Q}_i$, we use the language "delete ω" and "set $\gamma_i|\omega = *$" interchangeably. Let $\mathcal{Q}_{i_0}$ be the canonical partition on J defined by γ_{i_0}. Here is how we go from one step to the next:

Case 1. Consider $\omega \in \mathcal{Q}_i$ where $\zeta_i(\omega)$ is free and outside of $\mathcal{C}^{(1)}$. We take the canonical partition defined by ζ_{i+1} on ω, and set $\zeta_{i+1}|_{\omega'} = *$ for elements ω' of this partition for which $d_{\mathcal{C}}(\zeta_{i+1}(\omega')) < e^{-\alpha(i+1)}$. On the part of ω not deleted, we set $\gamma_{i+1} = \zeta_{i+1}$, and call the restriction of the canonical partition on it $\mathcal{Q}_{i+1}$.

Case 2. Consider $\omega \in \mathcal{Q}_i$ for which $\gamma_i(\omega)$ is free and inside $\mathcal{C}^{(1)}$. It follows from the previous step that $\pi_x(\gamma_i(\omega)) \subset I^+_{\mu j}$. A bound period $p(\omega)$ is set as in Sect. 11.2C. We put $\omega \in \mathcal{Q}_{i+j}$ for all $j < p$, and at step $i + p$, we do as in Case 1, i.e. consider the canonical partition defined by ζ_{i+p}, delete those elements with $d_{\mathcal{C}}(\cdot) < e^{-\alpha(i+p)}$, set $\gamma_{i+p} = \zeta_{i+p}$ on the rest and call the resulting partition $\mathcal{Q}_{i+p}$.

This completes the definition of the canonical subdivision by itinerary. We remark that by virtue of (C1), $d_{\mathcal{C}}(z_i(a)) > e^{-\alpha i}$ for all $i \leq N$ and $a \in J$, so that no deletions take place before time N. For $N < n \leq \frac{1}{\alpha^*}N$, the construction above is designed to guarantee that if $\gamma_n(a) \neq *$, then (C3) is satisfied for z_0 up to time n on the parameter interval $\omega \in \mathcal{Q}_n$ containing a.

Verification of estimates analogous to (P1)–(P3) for γ_n. Letting $\tau_n = \frac{d}{da}\gamma_n$, Lemmas 11.3 and 11.5 play the role of (P1) and (P2). (P3) at i_0 follows from (a) the corresponding result for w_{i_0}, (b) the fact that all $\zeta_n(a), a \in J$, remain very close to each other for all $n \leq i_0$, and (c) Proposition 11.1. The following distortion estimate for parameters is needed to take the place of (P3) for $n > i_0$. Its proof is given in Appendix A.26.

Lemma 12.2 *There exists $K > 0$ such that for any $\omega \in \mathcal{Q}_{n-1}$ such that $\gamma_n(\omega)$ is free,*

$$K^{-1} < \frac{|\tau_n(a)|}{|\tau_n(a')|} < K \qquad \text{for all } a, a' \in \omega.$$

12.3 Large deviation estimate

Having established the resemblance between γ_i and iterated 1D maps, we now state the analog of Proposition 2.2 for parameters. For $i_1 < i_2$, let $B(a, \hat{\delta}; i_1, i_2)$ denote the number of $i \in (i_1, i_2]$ such that $\gamma_i(a)$ is in a bound period initiated at a previous time j where $d_{\mathcal{C}}(z_j(a)) < \hat{\delta}$. Built into this definition is the implication that if i is one of the times counted in $B(a, \hat{\delta}; i_1, i_2)$, then $\gamma_i(a) \neq *$. The proof of Corollary 12.1 follows closely that of Proposition 2.2.

Corollary 12.1 *Assume (C1) and (C2), and let $z_0 \in \Gamma_{\frac{1}{\alpha^*}\theta N} \setminus \Gamma_{\theta N}$. Then Proposition 2.2 holds for $\{\gamma_i, i \leq \frac{1}{\alpha^*}N\}$, where $\{\gamma_i\}$ is the process associated with $a \mapsto z_i(a)$ defined above. More precisely, given any $\sigma > 0$, there exists $\hat{\varepsilon}_1 > 0$ such that the following holds for all sufficiently small $\hat{\delta} > 0$: Let t_0 and $\omega \in \mathcal{Q}_{t_0}$ be such that $\gamma_{t_0}(\omega)$ is free and $\approx I_{\mu_0 j_0}$ (in particular, $\gamma_{t_0}|_\omega \neq *$), and let n be such that (i) $t_0 + n \leq \frac{1}{\alpha^*}\theta N$ and (ii) $n > K\sigma^{-1}\log|\mu_0|$. Then*

$$|\{a \in \omega : B(a, \hat{\delta}; t_0, t_0 + n) > \sigma n\}| < e^{-\hat{\varepsilon}_1 n}|\omega|.$$

This result holds also if $\gamma_{t_0}(\omega)$ is outside of $\mathcal{C}^{(1)}$, free and has length $\geq \delta$. In this case condition (ii) for n is replaced by $n > K\sigma^{-1}\log\frac{1}{\delta}$.

12.4 In preparation for the selection of good parameters

The main ingredients for dealing with (A4) are treated in the last 3 subsections. The results as stated, however, are not quite in a form that can be applied directly. This subsection contains the adjustments needed to render Lemma 12.1 and Corollary 12.1 ready for use in the construction in Section 13. We also specify the desired relations between σ, $\hat{\delta}$ etc. and constants chosen earlier.

Setting We assume (C1) and (C2), and continue to focus on a single critical point $z_0 \in \Gamma_{\frac{1}{\alpha^*}\theta N} \setminus \Gamma_{\theta N}$. Let $\gamma_i, i \leq \frac{1}{\alpha^*}\theta N$, be the process associated with z_0. In what follows, we may assume also that $d_{\mathcal{C}}(z_i) > \delta_0$ for all $i \leq n_0$ where n_0 is as large as we need, and that at least one subdivision of the parameter interval takes place before $\gamma_i(J)$ meets $\mathcal{C}^{(1)}$.

(1) Measure of parameters deleted in connection with (A4)

The procedure in Section 13 requires that we work with time intervals of the type $[n, 2n]$.

Corollary 12.2 *Assume $\sigma >> \alpha$, and let $\hat{\varepsilon}_1 = \hat{\varepsilon}_1(\frac{1}{2}\sigma)$ be given by Corollary 12.1 (with $\frac{1}{2}\sigma$ in the place of σ). Let $\hat{\delta}$ be small enough to satisfy the requirement in Corollary 12.1. Then for all $\omega \in \mathcal{Q}_n$ with $\gamma_n|_\omega \neq *$,*

$$|\{a \in \omega : B(a, \hat{\delta}; n, 2n) > \sigma n\}| < e^{-\hat{\varepsilon}_1 n}|\omega|.$$

Proof: We explain the modifications necessary to apply Corollary 12.1. If $\gamma_n(\omega)$ is free and is either $\approx I_{\mu j}$ or is outside and has length $\geq \delta$, then we apply Corollary 12.1 directly with $t_0 = n$. Note that the lower bound on n in Corollary 12.1 is satisfied: if $\gamma_n(\omega) \approx I_{\mu j}$, then $\mu \leq \alpha n$, so that $n \geq \frac{1}{\alpha}\mu >> \sigma^{-1}\mu$; we may assume $n > K\sigma^{-1}\log\frac{1}{\hat{\delta}}$ since n_0 can be arbitrarily large. If $\gamma_n(\omega)$ is not free or is shorter than required, we back up to step n_1 when ω was first created as an element of some $\mathcal{Q}_{n_1}$. Notice first that there is such an n_1, for by assumption a subdivision occurred before time n. Moreover, $\gamma_{n_1}(\omega)$ is free, and it is either $\approx I_{\mu j}$ or is outside and has length $\geq \delta$. By the parameter version of Lemma 2.3, $n \leq (1 + K\alpha)n_1$. We may then apply Corollary 12.1 with $t_0 = n_1$ and $\frac{1}{2}\sigma$ in the place of σ. Since $\sigma >> \alpha$, a parameter a with $B(a, \hat{\delta}; n_1, 2n) > \sigma(2n - n_1)$ clearly satisfies $B(a, \hat{\delta}; n, 2n) > \frac{1}{2}\sigma n$. □

(2) Growth of $|w_i^*(z_0)|$ for good parameters

Let $n = 2^{j_1}n_0$, and consider the following procedure repeated on time intervals $[n_0, 2n_0], \cdots, [2^j n_0, 2^{j+1}n_0], \cdots, [2^{j_1-1}n_0, n]$: On each time interval $[2^j n_0, 2^{j+1}n_0]$, in addition to the deletions corresponding to (C3) (see Sect. 11.2), we delete at time $2^{j+1}n_0$ all $\omega \in \mathcal{Q}_{2^{j+1}n_0}$ on which $B(a, \hat{\delta}; 2^j n_0, 2^{j+1}n_0) > \sigma 2^j n_0$. The following corollary gives a lower bound on $|w_i^*(z_0)|$ for $T = T_a$ where a survives these deletions up to time n.

Corollary 12.3 *Assume (a) $d_{\mathcal{C}}(z_i) > \min(\delta, e^{-\alpha i})$ for all $i \leq n$;*
(b) $B(\hat{\delta}; 2^j n_0, 2^{j+1}n_0) < \sigma 2^j n_0$ for all $j < j_1$.
Then for all $i \leq n$, $|w_i^*(z_0)| > c_2 e^{[\frac{1}{4}(1-2\sigma)\lambda_0 - 3\alpha]i}$.

Proof: Assumptions (a) and (b) together with Lemma 12.1 imply that at times $2^j n_0$, $|w_i^*| > K^{-1}\hat{\delta}e^{[\frac{1}{4}(1-\sigma)\lambda_0 - 3\alpha]i}$. Between times i and $2i$, the worst-case scenario is that all the close returns (i.e. returns to $\{d_{\mathcal{C}}(\cdot) < \hat{\delta}\}$) occur at the beginning of this time block. Even so, we guarantee easily that for all $k < i$, $|w_{i+k}^*| > K^{-1}\hat{\delta}e^{[\frac{1}{4}(1-2\sigma)\lambda_0 - 3\alpha](i+k)}$. Observe finally that the factor $K^{-1}\hat{\delta}$ can be replaced by c_2, i.e. it is absorbed into the initial stretch if n_0 is sufficiently large depending on $\hat{\delta}$. □

(3) Choice of constants

The exponents directly related to derivative growth are λ_0, λ, λ^* and α. We review briefly what they represent. First, outside of $\mathcal{C}^{(1)}$, b-horizontal vectors grow at rate $\frac{1}{4}\lambda_0$; see Lemma 3.5. The first constant chosen in this paper, $\lambda < \frac{1}{5}\lambda_0$, is the minimum growth rate along critical orbits guaranteed by (A4); see Sect. 6.1. This growth rate is lowered by up to -3α during bound periods; see Lemma 5.4. Recall also the relationship $\alpha << \min\{\lambda, 1\}$. Finally, because more stringent estimates are needed for reasons to be explained, we fix a slightly larger target Lyapunov exponent $\lambda^* = \lambda + \frac{1}{100}\lambda_0$; see Sect. 10.1.

Next we come to σ, which is chosen so that $\frac{1}{4}(1-2\sigma)\lambda_0 - 3\alpha$, the exponent in Corollary 12.3, is $> \lambda^*$. For example, $\sigma = \frac{1}{100}$ will work. We may assume this is in agreement with the relation $\sigma >> \alpha$ as required in Corollary 12.2. Once σ is fixed, we choose $\hat{\delta}$ small enough to satisfy Corollary 12.1.

Summary: If σ and $\hat{\delta}$ are as in the last paragraph, and the hypotheses of Corollary 12.3 are satisfied, then $|w_i^*(z_0)| > c_2 e^{\lambda^* i}$ for all $i \le n$. Moreover, between times $2^j n_0$ and $2^{j+1}n_0$, the measure of parameters in violation of Corollary 12.3(b) is, by Corollary 12.2, $< e^{-\hat{\varepsilon}2^j n_0}|J|$.

13 Positive Measure Sets of Good Parameters

The purpose of this section is to construct, for a given family $\{T_a\}$ satisfying the Standing Hypotheses in Section 1 and with b sufficiently small, a sequence of sets

$$\Delta_0 \supset \Delta_{n_0} \supset \Delta_{2n_0} \supset \Delta_{2^2 n_0} \supset \cdots$$

in parameter space with the properties that

(i) $\{T_a, a \in \Delta_{2^j n_0}\} \subset \mathcal{G}^{\#}_{2^j n_0}$ (where $\mathcal{G}^{\#}_n$ is as defined in Sect. 10.1) and

(ii) $\Delta := \cap_{j\ge 0}\Delta_{2^j n_0}$ has positive Lebesgue measure.

Together with the material in Section 9, this construction brings to completion the proof of our Main Theorem.

We remark that the construction in this section requires more stringent conditions on the global constants in Sect. 6.1 than are imposed in Part II. See the end of Sect. 11.1 and Sect. 12.4(3).

13.1 Getting Started

The two properties required of the start-up interval Δ_0 are:

(1) *For all $a \in \Delta_0$ and $z_0 \in \Gamma_1$, $d_{\mathcal{C}}(z_i) > \delta_0$ for all $i \le n_0$ where n_0 is a very large number to be prescribed.*

(2) *For each $z_0 \in \Gamma_1$, a subdivision occurs in the process $a \mapsto z_i(a)$ before $\gamma_i(\Delta_0)$ meets $\mathcal{C}^{(1)}$.*

Here δ_0 is as in Definition 1.1; recall that $d(f^i_{a^*}(\hat{x}), C) > 2\delta_0$ for all $i > 0$. Lower bounds have been placed on n_0 a finite number of times in previous sections: among the more important places where this condition appeared are (i) to provide time for hyperbolicity of T_a to build up initially (see Part II); (ii) to allow the comparability of space and a-derivatives to take hold (see Proposition 11.1); and (iii) to absorb the small constant $\hat{\delta}$ from Lemma 12.1 (see Corollary 12.3). A few more conditions on n_0 will be imposed in this section. The process referred to in (2) is the one in Sect. 12.2. The purpose of (2) is to ensure that the entire parameter interval is not lost in the first deletion: Let n_1 be the first time Δ_0 is subdivided. Then $|\gamma_{n_1}(\omega)| \ge \delta$ for every $\omega \in \mathcal{Q}_{n_1}$, and if $\omega \in \mathcal{Q}_j$ is such that $\gamma_i(\omega) \cap \mathcal{C}^{(1)} = \emptyset$ for all $i < j$ and $\gamma_j(\omega) \cap \mathcal{C}^{(1)} \neq \emptyset$, then

$|\gamma_j(\omega)| > \hat{K}^{-1}c_2\delta$. Thus assuming n_0 is large enough that $e^{-\alpha n_0} << \hat{K}^{-1}c_2\delta$, we are guaranteed that only a small fraction of the measure is deleted.

We claim that for *any* Δ_0 *containing* a^* short enough for (1) to be satisfied, (2) is automatically satisfied if b is sufficiently small. To see this, let $\{\hat{x}_0^k\}$ be the critical points of the 1D maps f_a, and let $a \mapsto \hat{x}_i^k(a), i = 0, 1, \cdots$, be the critical curves defined by the 1D maps. For each k, let $\hat{n}_1^k \geq n_0$ be the first time $|\hat{x}_i^k(\Delta_0)| > 3\delta$, and let $\hat{n}_1$ be the maximum of the $\hat{n}_1^k$. Now let $\zeta_i^k : a \mapsto z_i^k(a)$ where z_0^k is the critical point of T_a near $(\hat{x}_0^k, 0)$. We choose b small enough that $|\hat{x}_i^k(a) - \pi_x(z_i^k(a))| << \delta$ for $i \leq \hat{n}_1$ for all $a \in \Delta_0$. Then for $i < \hat{n}_1^k$, $\zeta_i^k(\Delta_0) \cap \mathcal{C}^{(1)} = \emptyset$, and $|\zeta_{\hat{n}_1^k}^k(\Delta_0)| > 2\delta$, so a subdivision occurs at or before time $\hat{n}_1^k$ in the process associated with z_0^k.

In the rest of this section, let λ^* and $\mathcal{G}_n^{\#}$ be as defined in Sect. 10.1. For clarity of presentation, we first describe the construction up to time θ^{-1} (where the situation is simpler) before giving it in full generality.

13.2 Construction of Δ_N for $N \leq \theta^{-1}$

A. Outline of scheme

This time period is characterized by the fact that the only relevant critical points are those in $\Gamma_1 := \{z_0^1, \cdots, z_0^q\}$. Associated with each z_0^k, we construct a sequence of parameter sets $\Delta_0 = \Delta_0^k \supset \Delta_1^k \supset \Delta_2^k \supset \cdots \supset \Delta_{\theta^{-1}}^k$ with the property that for $a \in \Delta_i^k$, $z_j^k(a)$ has the desired properties for all $j \leq i$. The parameter sets $\Delta_i := \cap_{1 \leq k \leq q} \Delta_i^k$ consist, therefore, of parameters for which all the critical orbits have the desired properties up to time i.

The sets $\{\Delta_i^k\}$ are constructed in the following order. First, we set $\Delta_i^k = \Delta_0$ for all $i \leq n_0$ and all k. Then we proceed with an N-to-$2N$ scheme, i.e. we go from step n_0 to step $2n_0$, $2n_0$ to $4n_0$, $4n_0$ to $8n_0$, and so on, until θ^{-1}, which we may assume is $= 2^{\ell_0}n_0$, is reached. Within each stage, i.e. from $N = 2^\ell n_0$ to $2N$, we construct for each k the parameter sets Δ_i^k, $N < i \leq 2N$. Which k goes first is immaterial, but it is important that all the critical orbits be treated up to time $2N$ before we go to the next stage.

Remark. The number 2 in our N-to-$2N$ scheme is somewhat arbitrary; the idea of updating all the critical orbits to order $\sim N$ simultaneously (as opposed to treating one to an arbitrarily large time before beginning on a second) is not. This is because the derivative estimate (A4) for the q critical orbits cannot be developed independently of each other: when z_i^k visits $Q^{(1)}(z_0^{k'})$, it relies on the orbit of $z_0^{k'}$ to guide it through its derivative recovery, and parameters that are favorable for z_0^k may have been deleted for $z_0^{k'}$. As we will see, the times $2^\ell n_0, \ell = 1, 2, \cdots$, are designated times for different critical orbits to communicate to each other their selected parameter sets.

B. Processes $\{\gamma_i^k\}$ defined on Δ_0

In Sect. 12.2, we considered a parameter interval J on which all T_a are assumed to be in $\mathcal{G}_N$, and introduced for each critical point a process γ_i defined up to time $\frac{1}{\alpha^*}N$. In a similar manner, we now wish to define for each z_0^k a process

$$\gamma_i^k : \Delta_0 \to R_1 \cup \{*\}, \quad i = 0, 1, 2, \cdots, \theta^{-1}.$$

Sect. 12.2 does not guarantee that such a process is well defined, for it is not likely that $T_a \in \mathcal{G}_{\alpha^*\theta^{-1}}$ for all $a \in \Delta_0$. Here is how we circumvent the problem: we use the procedure in Sect. 12.2 to *extend* γ_i^k from step $N = 2^\ell n_0$ to step $2N$ *whenever it is feasible*, and to set $\gamma_i = *$ whenever it is not. More precisely, for fixed k and N, we assume γ_i^k is defined on Δ_0 for all $i \leq N$. Associated with γ_i is its canonical subdivision by itinerary $\mathcal{Q}_i$. For each $\omega \in \mathcal{Q}_N$, we set $\gamma_{N+1}|_\omega = *$ unless $T_a \in \mathcal{G}_{2\alpha^* N}$ for all $a \in \omega$. Thus when $\gamma_{N+1}|_\omega \neq *$, the construction in Sect. 12.2 can legitimately be carried out on ω up to time $2N$.

There is one other difference between the construction here and that in Sect. 12.2: In Sect. 12.2, $\gamma_i = *$ is set only to achieve $d_{\mathcal{C}}(z_i) > e^{-\alpha i}$. Here we permit the setting of $\gamma_i|\omega = *$, $\omega \in \mathcal{Q}_i$, for a wider range of reasons as we will see in paragraph C.

C. Formal procedure from step $N = 2^\ell n_0$ to step $2N$

At time N, assume we are handed the following objects: For each $k = 1, 2, \cdots, q$, there is a process $\gamma_i^k : \Delta_0 \to R_1 \cup \{*\}$ well defined up to time N. The set $\Delta_N^k := \{\gamma_N^k \neq *\}$ has the property that for all $a \in \Delta_N^k$,

(i) $d_{\mathcal{C}}(z_i^k(a)) > 3e^{-\alpha i}$ for all $i \leq N$;

(ii) $B(a, \hat{\delta}; 2^j n_0, 2^{j+1} n_0) < \sigma 2^j n_0$ for all $j < \ell$.

Observe, by Corollary 12.3, that $\Delta_N := \cap_k \Delta_N^k \subset \mathcal{G}_N^{\#}$.

How to go from step N to step $2N$: The following steps are taken for each k.

(1) First we set $\gamma_{N+1}^k|_{\hat{\omega}} = *$ on those $\hat{\omega} \in \mathcal{Q}_N^k$ with $\hat{\omega} \cap \Delta_N = \emptyset$.

(2) On the rest of the $\hat{\omega} \in \mathcal{Q}_N^k$, we extend the process γ_i^k to $2N$ (see justification below), deleting all $\omega \in \mathcal{Q}_i|_\omega$ with $d_{\mathcal{C}}(\gamma_i(\omega)) < 3e^{-\alpha i}$.

(3) Set $\gamma_{2N}^k|_\omega = *$ on those $\omega \in \mathcal{Q}_{2N}^k$ with the property that $B(a, \hat{\delta}; N, 2N) > \sigma N$ for $a \in \omega$.

Step (1) stipulates that unless some $a \in \hat{\omega}$ is good for all q critical points, the entire parameter interval will be abandoned.

Justification for step (2): We need to show that $T_a \in \mathcal{G}_{2\alpha^* N}$ for all $a \in \hat{\omega}$. By assumption, there exists $\hat{a} \in \hat{\omega}$ such that $T_{\hat{a}} \in \mathcal{G}_N^{\#}$. It follows from Proposition 10.1 that $T_{\hat{a}}$ has a continuation in $\mathcal{G}_{2\alpha^* N}$ on the interval $[\hat{a} - \rho^{-2\alpha^* N}, \hat{a} + \rho^{-2\alpha^* N}]$. On the other hand, Proposition 11.1 gives $|\hat{\omega}| < \hat{K} e^{-\hat{\lambda} N}$, which is $<< \rho^{-2\alpha^* N}$.

Note that steps (2) and (3) lead directly to (i) and (ii) above at time $2N$.

D. Measure deleted from step N to step $2N$

Consider one z_0^k at a time. We wish to estimate the contribution to $\Delta_N \setminus \Delta_{2N}$ by the orbit of z_0^k (this is not to be confused with $\Delta_N^k \setminus \Delta_{2N}^k$).

Deletions in Step (1): We have no control on the total measure of all the $\hat{\omega} \in \mathcal{Q}_N^k$ removed in this step, but all the $\hat{\omega}$ removed have the property that $\hat{\omega} \cap \Delta_N = \emptyset$: the very fact that $\hat{\omega} \cap \Delta_N = \emptyset$ means that all the parameters in $\hat{\omega}$ have been deleted earlier due to violations on the part of critical orbits other than that of z_0^k. Thus from the point of view of $\Delta_N \setminus \Delta_{2N}$, no measure is deleted in this step.

Deletions in Step (2): For i with $N < i \leq 2N$, we consider $\omega \in \mathcal{Q}_{i-1}^k$, and give an upper bound on the fraction of ω that may be deleted at the ith iterate. Let i_0 be the smallest $j < i$ such that $\omega \in \mathcal{Q}_j^k$, i.e. i_0 is the time when the partition interval ω is created. There are two possibilities:

(i) $\gamma_{i_0}^k(\omega)$ is outside of $\mathcal{C}^{(1)}$ and $\delta < |\gamma_{i_0}^k(\omega)| < 3\delta$. In this case, $|\gamma_i^k(\omega)| > K^{-1}\delta$, and not knowing the location of $\gamma_i^k(\omega)$, we assume the worst-case scenario, i.e. $\gamma_i^k(\omega)$ crosses entirely a forbidden region $d_{\mathcal{C}}(\cdot) < 3e^{-\alpha i}$. The fraction of ω deleted is then $< K\delta^{-1} 6 e^{-\alpha i} < K e^{-\frac{1}{2}\alpha i}$. Here K is the distortion constant (Lemma 12.2) as one transfers the length ratio on $\gamma_i^k(\omega)$ back to ω.

(ii) $\pi_x(\gamma_{i_0}^k(\omega)) \approx I_{\mu j}$. Let p be the bound period initiated at time i_0. Then $p \leq K|\mu|$, so that $|\gamma_i^k(\omega)| > K^{-1} |\gamma_{i_0+p}^k(\omega)| > \frac{K^{-1}}{\mu^2} e^{-K\alpha|\mu|} > e^{-2K\alpha|\mu|}$. For the first inequality above, we use first Proposition 11.1, then $|w_i| \geq c_2 |w_{i_0+p}|$. For the second inequality, we use Lemma 11.5(e). Thus the fraction of ω deleted is $< K 6 e^{-\alpha i} e^{2K\alpha^2 n} < K e^{-\frac{1}{2}\alpha i}$. (We note here the significance of the rule that in canonical subdivisions no partition point is introduced that would result in an element ω with $\pi_x(\gamma_{i_0}^k(\omega)) \subset I_{\mu j}$ and $|\gamma_{i_0}^k(\omega)| << |I_{\mu j}|$. For such an element, we would not be able to control the fraction of parameters deleted.)

We conclude that between times N and $2N$, the total measure deleted in the course of executing step (2) on the orbit of each z_0^k is $< \sum_{N<i\leq 2N} Ke^{-\frac{1}{2}\alpha i}|\Delta_0|$.

Deletions in Step (3): By Corollary 12.2, on each $\hat{\omega} \in \mathcal{Q}_N^k$, the total measure deleted is $< e^{-\hat{\varepsilon}_1 N}|\hat{\omega}|$. Thus the total measure deleted at time $2N$ on account of executing step (3) on the orbit of z_0^k is $< e^{-\hat{\varepsilon}_1 N}|\Delta_0|$.

Summary: Let $\mathcal{D}_{N,2N}^k$ denote the set of $a \in \Delta_N$ deleted on account of the orbit of z_0^k as we carry out our procedure from time N to time $2N$. Then

$$|\mathcal{D}_{N,2N}^k| \ < \ \left(K \sum_{N<i\leq 2N} e^{-\frac{1}{2}\alpha i} \ + \ e^{-\hat{\varepsilon}_1 N} \right) \ |\Delta_0|.$$

Writing the quantity in parenthesis as $K'e^{-\varepsilon' N}$, and letting $\mathcal{D}_{N,2N}$ denote the set of all parameters deleted between time N and time $2N$, we have the estimate

$$|\mathcal{D}_{N,2N}| \ \leq \ q \ Ke^{-\varepsilon' N} \ |\Delta_0|.$$

13.3 Construction of Δ_N for $N > \theta^{-1}$

A. Outline of scheme

Our basic strategy is as before, i.e. we work with cycles that go from time N to time $2N$, treating all relevant critical orbits in each cycle before going to the next and making deletions with the aid of processes of the type in the last subsection. There are two new aspects in the situation: the number of distinguishable critical orbits grows with time, and the critical structures of the maps T_a are not uniform for all $a \in \Delta_0$. The processes we consider must reflect this reality; they are discussed in part B below.

Parts C and D follow their counterparts in Sect. 13.2. Except for treating the new complexities brought to light in part B, they do not differ substantially from before.

B. Processes defined by critical orbits: two new aspects

(1) *Processes associated with critical blobs*

Relabeling the processes $\{\gamma_i^k\}$ in Sect. 13.2B as $\{\gamma_i^{z_0^k}\}$, we seek to explain what is meant by a process $\{\gamma_i^{z_0}\}$ where z_0 is an arbitrary critical point. Questions surrounding the domains of definition of $\{\gamma_i^{z_0}\}$ are treated in item (2) below. We discuss here the more basic question of whether to set $\gamma_i^{z_0} = z_i$ when it is $\neq *$.

Setting $\gamma_i^{z_0} = z_i$ for all i is only natural, but it has the following drawback: Let $z_0' \in B^{(j)}(z_0)$. Even though the orbits of z_0 and z_0' stay together for a long time, if treated independently, the canonical subdivisions accompanying $a \mapsto z_i(a)$ and $a \mapsto z_i'(a)$ are likely to produce slightly different partitions, and rules of deletion such as those in Sect. 13.2C may yield slightly different results. However inconsequential, these differences are a technical nuisance.

To avoid this technical nuisance, we have elected to view $\{\gamma_i^{z_0}\}$ as associated with critical blobs. More precisely, let z_0 be a critical point of generation j. We let $B^{(j_1)} \supset B^{(j_2)} \supset \cdots \supset B^{(j_n)}$ be the complete chain of critical blobs containing $B^{(j)}(z_0)$, i.e. $j_1 = 1, j_n = j$, and for each i, there is no $B^{(k)}$, $j_i < k < j_{i+1}$, such that $B^{(j_i)} \supset B^{(k)} \supset B^{(j_{i+1})}$. (See Sect. 7.3 for the geometry of critical regions.) We say $B^{(j_i)}$ is *visible* on the time interval $(j_{i-1}\theta^{-1}, j_i\theta^{-1}]$, thinking of it as "hidden" inside lower-generation critical blobs before time $j_{i-1}\theta^{-1}$ and no longer active after time $j_i\theta^{-1}$. During the time period when $B^{(j_i)}$ is visible, we set $\gamma_\ell^{z_0} = z_\ell^{(j_i)}$

or $*$ where $z_0^{(j_i)} = z_0^*(B^{(j_i)})$. That is to say, $B^{(1)}(z_0)$ is visible in the first θ^{-1} iterates, and for $\ell \leq \theta^{-1}$, $\gamma_\ell^{z_0} = z_\ell^{(1)}$ or $*$. Since $z_0^{(1)} = z_0^k$ for some k, $\{\gamma_i^{z_0}\}$ is identical to one of the processes defined in the last subsection for $i \leq \theta^{-1}$. At time θ^{-1}, the critical blob $B^{(1)}$ retires, and $B^{(2)}$ becomes visible. We have $\gamma_\ell^{z_0} = z_\ell^{(2)}$ or $*$ for $\theta^{-1} < \ell \leq 2\theta^{-1}$, and so on. (Note that some of the orbit segments are visible for more than θ^{-1} steps due to the "skipping of generations"; see Sect. 7.3).

Assuming for the moment that all definitions are legitimate and all rules for deletion are as in Sect. 13.2C, we make the following observation: During the period when $\gamma_\ell^{z_0} = z_\ell^{(j_i)}$ or $*$, $d_{\mathcal{C}}(z_\ell^{(j_i)}) > 3e^{-\alpha\ell}$ implies that $d_{\mathcal{C}}(\xi_\ell) > 2e^{-\alpha\ell}$ for all $\xi_0 \in B^{(j_i)}$. This follows from earlier estimates on the sizes of critical blobs; see Sect. 4.2. Moreover, if $d_{\mathcal{C}}(\gamma_\ell^{z_0}) > 3e^{-\alpha\ell}$ for all $\ell \leq j\theta^{-1}$, then all $\xi_0 \in B^{(j)}(z_0)$ satisfy $d_{\mathcal{C}}(\xi_\ell) > 2e^{-\alpha\ell}$. In particular, all critical points inside $B^{(j)}(z_0)$ obey (A2)$^\#$ up to time $j\theta^{-1}$. The same conclusion is valid for (A4)$^\#$ since up to time $j\theta^{-1}$, all $\xi_0 \in B^{(j)}(z_0)$ can be regarded as having the same itinerary. Hence they have the same fraction of "bad iterates" in the sense of $B(\cdot, \hat{\delta}; n, 2n)$.

(2) *Stabilization of critical structures and extending the processes* $\{\gamma_i^{z_0}\}$

In Sect. 13.2, we considered processes defined by $z_0 \in \Gamma_1$, which has a continuation on all of Δ_0. Critical structures of higher generations do not have such continuations. To stabilize these structures, we introduce an increasing sequence of partitions $\mathcal{J}_{\theta^{-1}} < \mathcal{J}_{2\theta^{-1}} < \mathcal{J}_{4\theta^{-1}} < \cdots$ on Δ_0 with the following properties: $\mathcal{J}_{\theta^{-1}} = \{\Delta_0\}$; for each $N = 2^\ell\theta^{-1}, \ell \geq 1$, $\mathcal{J}_N$ is a refinement of $\mathcal{J}_{\frac{1}{2}N}$ and partitions Δ_0 into intervals of length $\approx \rho^{\alpha^* N}$. Leaving precise rules of deletion to part C, we explain here the relation between these partitions and the processes defined in (1).

For $N = 2^\ell\theta^{-1}, \ell = 1, 2, \cdots$, the picture is as follows:

(i) There is a decreasing sequence of "good sets" Δ_N with the property that $T_a \in \mathcal{G}_N^\#$ for all $a \in \Delta_N$. (This is not the definition of Δ_N, however.)

(ii) There are subcollections of "good" intervals $\mathcal{J}_N^* \subset \mathcal{J}_N$. For each $J_N \in \mathcal{J}_N^*$,
- $J_N \subset J_{\frac{1}{2}N}$ for some $J_{\frac{1}{2}N} \in \mathcal{J}_{\frac{1}{2}N}^*$,
- $J_N \cap \Delta_{\frac{1}{2}N} \neq \emptyset$.

(iii) For each $J_N \in \mathcal{J}_N^*$, let $\Gamma^N(J_N)$ be the set of critical points of $T_a, a \in J_N$, of generations between $\frac{1}{2}\theta N$ and θN. Then for each $z_0 \in \Gamma^N(J_N)$ of generation k, there is a well defined process $\{\gamma_i^{z_0}, i \leq \min(k\theta^{-1}, N)\}$, the domains of definition of which are as follows:
Let $\Delta_0 = J_{\theta^{-1}} \supset J_{2\theta^{-1}} \supset \cdots \supset J_N$ be the elements of $\mathcal{J}_{2^\ell\theta^{-1}}$ containing J_N. Then
  - $\gamma_i^{z_0}, i \leq \theta^{-1}$, is defined on $\Delta_0 = J_{\theta^{-1}}$ (this is what is constructed in Sect. 13.2);
  - the process above is extended from $i = \theta^{-1}$ to $i = 2\theta^{-1}$ on $J_{2\theta^{-1}}$, then from $i = 2\theta^{-1}$ to $i = 4\theta^{-1}$ on $J_{4\theta^{-1}}$, and so on, up to $i = \frac{1}{2}N$;
  - the product of this last extension is extended from $i = \frac{1}{2}N$ to $i = \min(k\theta^{-1}, N)$ on J_N where k, as we recall, is the generation of z_0.

We explain how to go from step N to step $2N$, clarifying along the way what exactly is meant by some of the statements in (iii) and how they can be achieved:

Elements of $\mathcal{J}_N$ not in $\mathcal{J}_N^*$ are discarded since all parameters in them have been deleted in a previous step (second property of J_N in assumption (ii) above). Let $J_N \in \mathcal{J}_N^*$ be fixed. We consider $\mathcal{J}_{2N}|_{J_N}$, and put into $\mathcal{J}_{2N}^*$ those elements of $\mathcal{J}_{2N}$ that meet Δ_N (as required by (ii)). Consider a (fixed) $J \subset J_N$ such that $J \in \mathcal{J}_{2N}^*$. Since there exists $\hat{a} \in J$ such that $T_{\hat{a}} \in \mathcal{G}_N^\#$ (assumption (ii)), $T_a \in \mathcal{G}_{2\alpha^* N}$ for all $a \in J$ (Proposition 10.1). Thus on J there is a coherent choice of $\Gamma_{2\theta N}$ whose orbits can be treated up to time $\min(k\theta^{-1}, 2N)$ where k is the generation of the critical point.

Fix $z_0 \in \Gamma^{2N}(J)$. Since k, the generation of z_0, is $\geq \theta N$, $\gamma_i^{z_0}$ is defined for all $i \leq N$ in the sense of (iii). As mandated by (iii), we now seek to extend this process to all $i \leq \min(k\theta^{-1}, 2N)$

on the interval J. Such an extension is carried out on one $\omega \in \mathcal{Q}_N^{z_0}$ at a time. Fix ω such that $\gamma_N^{z_0}|\omega \neq *$. If $\omega \subset J$, then we consider $\gamma_i^{z_0}$ for $i = N+1, N+2, \cdots$ starting from ω as explained in Sect. 13.2B. If $\omega \cap J = \emptyset$, then ω is not our concern. It remains to consider the case $\omega \cap \partial J \neq \emptyset$.

Observation: If for all $a \in \omega$, $z_0(a)$ satisfies the hypotheses of Proposition 11.1 up to time N, then $|\omega| << |J|$.

Indeed, $\omega \in \mathcal{Q}_N^{z_0}$ and $J \in \mathcal{J}_N$ have exponentially different length scales. This is because by Proposition 11.1, $|\omega| < \hat{K}e^{-\hat{\lambda}N}$, which is $<< \rho^{\alpha^* N} = |J|$, and our rules of deletion, which are stated precisely in part C below, are built to ensure that the hypotheses of Proposition 11.1 are satisfied for $z_0(a)$ for every $a \in \{\gamma_N^{z_0} \neq *\}$. To deal with those ω that intersect some $J \in \mathcal{J}_N^*$ but are not completely contained in it, we let J^+ be, say, 10% longer than J, and treat all $\omega \in \mathcal{Q}_N^{z_0}$ that are completely contained in J^+. The properties of J continue to be valid in J^+, and these overlapping intervals lead to an overcount by a factor of at most 2. This completes the qualitative description of the extension of (iii).

We finish with the inductive definition of Δ_N, even though the following acquires meaning only after the deletion rules are specified. We let

$$\Delta_{2N}^{z_0} := \{\gamma_{\min(k\theta^{-1},2N)}^{z_0} \neq *\}, \qquad \Delta_{2N}(J) := \cap_{z_0 \in \Gamma^{2N}(J)} \Delta_{2N}^{z_0}$$

and

$$\Delta_{2N} := \Delta_N \cap \left(\cup_{J \in \mathcal{J}_{2N}^*} \Delta_{2N}(J)\right).$$

Our deletion rules are designed to ensure that Δ_N as defined above has the property in assumption (i), and that $\cap_N \Delta_N$ has positive measure.

Remarks (1) The intersection with Δ_N in the definition of Δ_{2N} may seem redundant, for on $J \in \mathcal{J}_{2N}^*$, all critical blobs corresponding to $z_0 \in \Gamma^{2N}(J)$ have already been treated up to time $\min\{k^{-1}\theta, 2N\}$. Consequently, it is tempting to claim that $T_a \in \mathcal{G}_{2N}^{\#}$ for all $a \in \cup_{J \in \mathcal{J}_{2N}^*} \Delta_{2N}(J)$. This is not true in general, for not every critical blob has offsprings (meaning smaller critical blobs inside), and the definition of $\cup_{J \in \mathcal{J}_{2N}^*} \Delta_{2N}(J)$ does not take into consideration the behavior of critical blobs that expired without reproducing before time N. To ensure that $T_a \in \mathcal{G}_{2N}^{\#}$ for all $a \in \Delta_{2N}$, we require that $\Delta_{2N} \subset \Delta_N$, and use the inductively obtained fact that for all $a \in \Delta_N$, all $z_0 \in \Gamma^N(J)$ are well behaved.

(2) To deal with the phenomenon called "skipping of generations", we need to work with slightly overlapping intervals of generations to ensure that all critical behaviors are represented. For example, we should have included in the definition of $\Gamma^N(J)$ all critical points from generation θN to generation $2\theta N(1+2\theta)$, and the elements of $\mathcal{J}_N$ should have been taken to be of length $\approx \frac{1}{2}\rho^{\alpha^* N(1+2\theta)}$, and so on. We have omitted – and will continue to omit – all of these factors of $(1+2\theta)$ to simplify slightly the discussion. The problem is easy to rectify (and should probably ignored on first pass).

C. Formal procedure from step $N = 2^\ell\theta^{-1}$ to step $2N$

We now give the formal procedure at a generic step N. The following should not be thought of as induction hypotheses, but rather as a summary of the situation as we arrive at step N following the procedure described in part B.

At time N we assume we have the following:

(a) A subcollection $\mathcal{J}_N^*$ of $\mathcal{J}_N$ with the property that on each $J \in \mathcal{J}_N^*$, there is a coherent choice of $\Gamma_{\theta N}$; $\Gamma^N(J)$ and $\Delta_N(J)$ are as defined above.

(b) On each $J \in \mathcal{J}_N^*$, associated with each $z_0 \in \Gamma^N(J)$ of generation k is a process $\gamma_i^{z_0} : J \to R_1 \cup \{*\}$ for which the following hold: For all $a \in \Delta_N^{z_0}$ and $i \leq \min(k\theta^{-1}, N)$,
  (i) $d_{\mathcal{C}}(\gamma_i^{z_0}(a)) > 3e^{-\alpha i}$;
  (ii) $B(a, \hat{\delta}; 2^j\theta^{-1}, 2^{j+1}\theta^{-1}) < \sigma 2^j \theta^{-1}$.

(c) A subset of $\cup_{J \in \mathcal{J}_N^*} \Delta_N(J)$ in $\{a : T_a \in \mathcal{G}_N^{\#}\}$ called Δ_N.

As noted in B(1) above, all $z_0 \in \Gamma^N(J)$ of generation k obey (A2)$^\#$ and (A4)$^\#$ up to time $\min(k\theta^{-1}, N)$. Step (c) is needed because $\cup_{J \in \mathcal{J}_N^*} \Delta_N(J)$ is not necessarily in $\{a : T_a \in \mathcal{G}_N^{\#}\}$ (see Remark (1) in B(2) above).

What is done from time N to time $2N$:

(0) First we introduce the partition $\mathcal{J}_{2N}$, and let $\mathcal{J}_{2N}^* \subset \mathcal{J}_{2N}$ be the collection of J with $J \cap \Delta_N \neq \emptyset$. Elements of $\mathcal{J}_{2N} \setminus \mathcal{J}_{2N}^*$ are excluded from further consideration.

We then treat one $J \in \mathcal{J}_{2N}^*$ at a time, carrying out for it steps (1)–(4) below. Steps (1)–(3) are carried out for each $z_0 \in \Gamma^{2N}(J)$, beginning with the z_0 of the lowest generations.

(1) Set $\gamma_{N+1}^{z_0}|_\omega = *$ on those $\omega \in \mathcal{Q}_N^{z_0}$ with $\omega \cap \Delta_N = \emptyset$.

(2) On the rest of the $\omega \in \mathcal{Q}_N^{z_0}$, we continue the process to time $2N$ in the manner described above, deleting along the way all $\omega' \in \mathcal{Q}_i^{z_0}$ with $d_{\mathcal{C}}(\gamma_i^{z_0}(\omega')) < 3e^{-\alpha i}$.

(3) Set $\gamma_{2N}^{z_0}|_\omega = *$ on those $\omega \in \mathcal{Q}_{2N}^{z_0}$ with the property that $B(a, \hat{\delta}, N, 2N) > \sigma N$ for $a \in \omega$.

(4) Define $\Delta_{2N}^{z_0} = \{\gamma_{\min(k\theta^{-1}, 2N)}^{z_0} \neq *\}$ and $\Delta_{2N}(J) = \cap_{z_0 \in \Gamma^{2N}(J)} \Delta_{2N}^{z_0}$ as in Part B.

Finally, after all the $J \in \mathcal{J}_{2N}^*$ are treated, we set $\Delta_{2N} = \Delta_N \cap (\cup_{J \in \mathcal{J}_{2N}^*} \Delta_{2N}(J))$.

Step (0) is to ensure the existence of a coherent choice of $\Gamma_{2\theta N}$ on each selected J. Step (1) is to ensure that the process can legitimately be extended on those ω on which $\gamma_{N+1}^{z_0} \neq *$. Note also that every z_0 has an ancestor, so all $\gamma_i^{z_0}$ are extensions of previously constructed processes. Since many of the $z_0 \in \Gamma^{2N}(J)$ are related to each other via ancestry, the steps above in fact contain many duplications.

It is evident that the steps above lead to (a)–(c) at the beginning of Part B for time $2N$.

D. Measure deleted from time N to time $2N$

First we estimate the measure deleted on account of a fixed $J \in \mathcal{J}_{2N}^*$ and a fixed $z_0 \in \Gamma^{2N}(J)$: Step (0) does not contribute to $\Delta_N \setminus \Delta_{2N}$ since no $a \in \cup\mathcal{J}_N^* \setminus \cup\mathcal{J}_{2N}^*$ belongs in Δ_N. The same remark holds for step (4). Explanations and estimates for steps (1)–(3) are exactly as before, except that $|\Delta_0|$ should be replaced by $|J^+|$. Thus we have

$$|\mathcal{D}_{N,2N}^{z_0}| \ \leq \ K' e^{-\varepsilon' N} \cdot 2|J|.$$

Since the cardinality of Γ^{2N} is $\leq 2N\theta K_1^{2\theta N(1+2\theta)}$ (Corollary 7.1), we have

$$|\mathcal{D}_{N,2N}| \ \leq \ \sum_{J \in \mathcal{J}_{2N}^*} \ \sum_{z_0 \in \Gamma^{2N}} |\mathcal{D}_{N,2N}^{z_0}| \ \leq \ 2N\theta K_1^{2\theta N(1+2\theta)} \cdot K' e^{-\varepsilon' N} \cdot 2|\Delta_0|.$$

13.4 The final count

From Sects. 13.1, 13.2C and 13.3D, we see that the total measure deleted at the end of the procedure is

$$\leq \ \left(K\delta_0^{-1} e^{-\alpha n_0} \ + \ qK' \sum_{N = 2^\ell n_0, \ell \in \mathbb{Z}^+} N\theta K_1^{2\theta N(1+2\theta)} \cdot e^{-\varepsilon' N} \cdot \right) \ |\Delta_0|.$$

As $(a,b) \to (a^*,0)$, $n_0 \to \infty$ and $\theta \to 0$, but none of the other constants is affected. Thus with (a,b) sufficiently near $(a^*,0)$, the quantity in parenthesis can be made arbitrarily small. In other words, $|\Delta|$ can be made as large a fraction of $|\Delta_0|$ as we wish. This completes the proof of our main result.

APPENDICES

A.1 Properties of "good" 1D maps (Sects. 2.1 and 2.2)

Proof of Lemma 2.1: Let x be such that $f^i(x) \notin C_\delta$ for $i \in [0,n)$. We divide $[0,n]$ into maximal time intervals $[i, i+k]$ such that $f^{i+j}(x) \notin C_{\delta_0}$ for $0 < j < k$, and estimate $|(f^k)'(f^i x)|$ as follows:

Case 1. $f^i(x), f^{i+k}(x) \in C_{\delta_0}$. Definition 1.1(b)(ii) and (c)(ii) together guarantee that $|(f^k)'(f^i x)| \geq e^{\frac{1}{3}\lambda_0 k}$.

Case 2. $f^i(x) \notin C_{\delta_0}, f^{i+k}(x) \in C_{\delta_0}$. Same as Definition 1.1(b)(ii).

Case 3. $f^i(x), f^{i+k}(x) \notin C_{\delta_0}$. If $k \geq M_0$, then $|(f^k)'(f^i x)| > e^{\lambda_0 k}$ from Definition 1.1(b)(i). If $k < M_0$, we let $\hat{k}$ be the smallest integer $> k$ such that $f^{i+\hat{k}}(x) \in C_{\delta_0}$. Using Definition 1.1(b)(i) for $\hat{k} \geq M_0$ and Definition 1.1(b)(ii) for $\hat{k} < M_0$, we conclude that $|(f^k)'(f^i x)| > c_0 K_0^{-M_0} e^{\lambda_0 k}$ where $K_0 = \max |f'(x)|$.

Case 4. $f^i(x) \in C_{\delta_0}, f^{i+k}(x) \notin C_{\delta_0}$. Same as Case 3, with an extra factor $\geq (\min_{y \in C_{\delta_0}} |f''(y)|)\,\delta$.

Cases 3 and 4 are relevant only for part (a). □

Proof of Lemma 2.2: Proceed as in the proof of Lemma 2.1. From Definition 1.1(b)(i) and (c)(ii), we see that for f, the estimates in all 4 cases are determined by $|(f^j)'(y)|$ for $y \notin C_\delta$ and $j \leq N := \max(M_0, K \log \frac{1}{\delta})$. Choose g sufficiently near f that $|g^j(y) - f^j(y)|$ is sufficiently small for all $y \notin C_\delta$ and $j \leq N$. □

We will use the notation $x_i = f^i(x)$.

Proof of Proposition 2.1: (P1) is Lemma 2.2. Let $x \in C_\delta(\hat{x})$.

Sublemma A.1.1 *For all $y \in [\hat{x}, x]$ and $k < p$, we have*

$$\frac{1}{2} \leq \frac{(f^k)'(y_1)}{(f^k)'(\hat{x}_1)} \leq 2$$

provided that δ and ε are sufficiently small.

Proof: We write

$$\log \frac{(f^k)'(y_1)}{(f^k)'(\hat{x}_1)} \leq \sum_{j=1}^{k} \frac{|f'(y_j) - f'(\hat{x}_j)|}{|f'(\hat{x}_j)|} \leq K \sum_{j=1}^{k} \frac{|y_j - \hat{x}_j|}{d(\hat{x}_j, C)}.$$

We first choose h_0 large enough that $\frac{1}{\delta_0}\sum_{i=h_0+1}^{\infty} e^{-2\alpha j} << 1$, followed by δ small enough that $\delta \sum_{j=1}^{h_0} \frac{1}{\delta_0} K^j << 1$. We then require ε to be sufficiently small so that $d(\hat{x}_j, C) > \delta_0\ \forall j < n_0$ for some n_0 satisfying $e^{-\alpha n_0} < \delta$. These choices ensure that

$$\sum_{j=1}^{k} \frac{|y_j - \hat{x}_j|}{d(\hat{x}_j, C)} < \sum_{j=1}^{h_0} \frac{1}{\delta_0} K^j \delta + \sum_{j=h_0+1}^{n_0} \frac{1}{\delta_0} e^{-2\alpha j} + \sum_{j=n_0+1}^{k} e^{-(2\alpha - \alpha)j} << 1.$$

◊

Proof of (P2): Suppose $|x-\hat{x}| = e^{-h}$. Then (G2) together with the sublemma above imply that $|x_p - \hat{x}_p| = |(f^{p-1})'(y_1)||x_1-\hat{x}_1| \geq K^{-1}e^{\lambda(p-1)}(x-\hat{x})^2$. ¿From $|x_p-\hat{x}_p| < 1$, we read off the upper bound $p < \frac{3}{\lambda}h$ for h sufficiently large. For the lower bound, we write $|x_p - \hat{x}_p| < K^{p-1}e^{-2h}$ and recall that p is defined such that $|x_p - \hat{x}_p| \geq e^{-2\alpha p}$. That $p >$ constant$\cdot h$ follows directly from $K^{p-1}e^{-2h} > e^{-2\alpha p}$.

To prove (P2)(ii) we again write $|x_p - \hat{x}_p| < K|(f^{p-1})'(y_1)|(x-\hat{x})^2$, so that

$$K|(f^{p-1})'(\hat{x}_1)|^{\frac{1}{2}}\ |x-\hat{x}| > e^{-\alpha p}. \tag{18}$$

We also have $|(f^p)'(x)| = |(f^{p-1})'(x_1)||f'x| > (K^{-1}|(f^{p-1})')(\hat{x}_1)|) \cdot (K^{-1}|x-\hat{x}|)$. Combined with (18) this gives $|(f^p)'(x)| > K^{-3}|(f^{p-1})'(\hat{x}_1)|^{\frac{1}{2}}\ e^{-\alpha p} > \hat{c}_1^{\frac{1}{2}}K^{-3}e^{\frac{1}{2}\lambda(p-1)}e^{-\alpha p}$, which we may assume is $> e^{-\frac{1}{3}\lambda p}$ if p is sufficiently large, or equivalently, δ is sufficiently small.

It remains to prove (P2)(iii). From (P2)(i), (ii) and Sublemma A.1.1, it follows that for $I_{\mu j} \in \mathcal{P}|_{C_\delta(\hat{x})}$,

$$|f^p(I_{\mu j})|\ \geq K^{-1}\frac{|f(I_{\mu j})|}{|f([\hat{x},\hat{x}+e^{-\mu}])|}\ |f^p([\hat{x},\hat{x}+e^{-\mu}])|\ \geq\ K^{-1}\frac{1}{\mu^2}e^{-2\alpha p}\ > e^{-K\alpha|\mu|}.$$

□

Proof of (P3): We write $\sigma_0 = [x,y]$, $\sigma_k = f^{t_k}\sigma_0$, and assume for definiteness that $\sigma_0 \subset C_\delta$ and $n \geq t_q + p_q$. Then

$$\log\frac{(f^n)'(x)}{(f^n)'(y)}\ \leq\ \sum_{j=0}^{n-1}\frac{|f'(y_j)-f'(x_j)|}{|f'(y_j)|}\ \leq\ K\sum_{k=1}^{q}(S'_k + S''_k)$$

where

$$S'_k = \sum_{j=t_k}^{t_k+p_k-1}\frac{|y_j - x_j|}{d(y_j,C)} \quad \text{and} \quad S''_k = \sum_{t_k+p_k}^{t_{k+1}-1}\frac{|y_j-x_j|}{d(y_j,C)}$$

except for S''_q which ends at index $n-1$.

I. Bound on $\sum_{k=1}^{q} S''_k$

For $k < q$ and $t_k + p_k \leq j < t_{k+1}-1$, we have, by (P1)(ii), $|\sigma_{k+1}| \geq c_1 e^{\lambda(t_{k+1}-j)}|x_j - y_j|$, so $S''_k \leq K\frac{|\sigma_{k+1}|}{\delta}$. Also, by combining (P2)(ii) and (P1)(ii), we have $|\sigma_{k+1}| \geq e^{\frac{1}{3}\lambda(t_{k+1}-t_k)}|\sigma_k|$ $\geq \tau|\sigma_k|$ for some $\tau > 1$, so $\sum_{t=0}^{q-1} S''_k \leq K\frac{|\sigma_q|}{\delta}$.

The term S''_q is treated differently because x_n may not be a return. Observe the following: (i) If $[x_n, y_n] \subset C_{\delta_0}$, then (P1)(ii) gives, as before, $S''_q \leq \frac{1}{\delta}K|y_n - x_n|$ which is $\leq K$ since $|y_{n-1} - x_{n-1}| \lesssim \delta$ by definition. (ii) If for $t_q + p_q \leq j < n$, $[x_j, y_j] \cap C_{\delta_0} = \emptyset$, then (P1)(i) with δ_0 in the place of δ gives $S''_q \leq \frac{1}{\delta_0}\frac{K}{\delta_0}|y_n - x_n| \leq K\delta\delta_0^{-2} \leq K$. In general, if there is $\hat{n} \geq t_q + p_q$ such that $\hat{n}$ is the last return to C_{δ_0} before time n, then we apply (i) to $\sum_{t_q+p_q}^{\hat{n}}$ and (ii) to $\sum_{\hat{n}+1}^{n-1}$.

II. Bound on $\sum_{k=1}^{q} S'_k$

First we estimate S'_k. Suppose $y_{t_k} \in C_\delta(\hat{x})$. For $t_k < j < t_k + p_k$ we write

$$\frac{|y_j - x_j|}{d(y_j,C)} = \frac{|y_j - x_j|}{|y_j - \hat{x}_{j-t_k}|}\cdot\frac{|y_j - \hat{x}_{j-t_k}|}{d(y_j,C)}.$$

By Sublemma A.1.1 and the usual estimates near $\hat{x}$, the first factor on the right is

$$<\ K\frac{|y_{t_k+1} - x_{t_k+1}|}{|y_{t_k+1} - \hat{x}_1|}\ <\ K\frac{|f'(x_{t_k})|\ |y_{t_k} - x_{t_k}|}{|y_{t_k} - \hat{x}|^2}\ <\ K\frac{|\sigma_k|}{d(y_{t_k},C)}\ .$$

Thus

$$S_k' \;=\; \frac{|y_{t_k}-x_{t_k}|}{d(y_{t_k},C)} + \sum_{j=t_k+1}^{t_k+p_k-1} \frac{|y_j-x_j|}{d(y_j,C)} \;\le\; K\frac{|\sigma_k|}{d(y_{t_k},C)}\left(1+\sum_{j=t_k+1}^{t_k+p_k-1}\frac{|y_j-\hat{x}_{j-t_k}|}{d(y_j,C)}\right) \;<\; K\frac{|\sigma_k|}{d(y_{t_k},C)}.$$

Now let $\mathcal{K}_\mu = \{k \le q : \ \sigma_k \subset I_{\mu,j} \text{ for some } j\}$. Then

$$\sum_{k\in\mathcal{K}_\mu} S_k' < \sum_{k\in\mathcal{K}_\mu} K\frac{|\sigma_k|}{e^{-|\mu|}} < K\frac{1}{\mu^2}.$$

The first inequality is from above. The second follows from the following two facts: (i) $|\sigma_{k+1}| \ge \tau|\sigma_k|$ for $k < q$, and (ii) the term with the largest index is bounded above by $|I^+_{\mu,j}|$, which is $< K\frac{1}{\mu^2}e^{-\mu}$. To finish, we sum over all μ to obtain $\sum S_k' < K$. □

A.2 Growth estimates and large deviations (Sect. 2.3)

To avoid cumbersome notation, we write μ instead of $|\mu|$ in all estimates.

Proof of Lemma 2.3: Since points in ω are assumed to have the same itinerary up to time n, $[0,n]$ is divided into bound intervals (t_k, t_k+p_k) and free intervals $[t_k+p_k, t_{k+1}]$. From (P2)(ii), we have $|\gamma_{t_k+p_k}(\omega)| > e^{\frac14\lambda p_k}|\gamma_{t_k}(\omega)|$ and from (P1)(ii), we have $|\gamma_{t_{k+1}}(\omega)| > c_1 e^{\frac14\lambda_0(t_{k+1}-t_k-p_k)}|\gamma_{t_k+p_k}(\omega)|$. Thus for any time j such that $\gamma_j(\omega)$ is free, $|\gamma_j(\omega)| > e^{\frac15\lambda j}|\omega|$. Now $|\gamma_n(\omega)| < 1$ forces n to be $< K\mu_0$. □

Proof of Lemma 2.5: Let $s \in \omega$ be such that $S(s) > n$. We define the *essential return times* $t_1 < t_2 < \cdots$ and corresponding *return addresses* $I^{i_1}_{\mu_1 j_1}, I^{i_2}_{\mu_2 j_2}, \cdots$ for s as follows: Let t_1 be the smallest $i > 0$ when either (a) $\gamma_i(\omega)$ is out of bound period and $|\gamma_i(\omega)| > \delta$ or (b) i is the extended bound period of $\gamma_0(\omega)$, whichever happens first. If (a) happens first, then $S|_\omega = t_1$, and we stop iterating. If not, then we may assume $\gamma_{t_1}(\omega) \subset C_\delta$, and the return address of s at time t_1 is $I^{i_1}_{\mu_1 j_1}$ if $\gamma_{t_1}(\mathcal{Q}_{t_1}(s)) \approx I_{\mu_1 j_1} \subset C_\delta(\hat{x}_{i_1})$. Similarly, $t_2(s)$ is the first $i > t_1(s)$ when either (a) $\gamma_i(\mathcal{Q}_{t_1}(s))$ is out of bound period and $|\gamma_i(\mathcal{Q}_{t_1}(s))| > \delta$ or (b) i is the extended bound period of $\mathcal{Q}_{t_1}(s)$, whichever happens first. Again if (a) happens first, then $S|_{\mathcal{Q}_{t_1}(s)} = t_2$ and we stop considering $\mathcal{Q}_{t_1}(s)$; otherwise $\gamma_{t_2}(\mathcal{Q}_{t_2}(s)) \approx I_{\mu_2 j_2} \subset C_\delta(\hat{x}_{i_2})$, and so on.

Let $A_q = \{s \in \omega : S(s) > n$, and $\gamma_i(s)$ makes a total of exactly q essential returns before time $n\}$. Then $|\{S > n\}| = \sum_q |A_q|$. We write $A_q = \cup_R A_{q,R}$ where $A_{q,R} = \{s \in A_q :$ if $(\mu_1, \cdots, \mu_q)$ are the μ-coordinates of its first q return addresses, then $|\mu_1| + |\mu_2| + \cdots + |\mu_q| = R\}$. We further decompose $A_{q,R}$ into intervals σ consisting of points whose first q return addresses are identical. For σ with return addresses $(I_{\mu_1 j_1}, \cdots, I_{\mu_q j_q})$, we let $Q_{t_i} = \mathcal{Q}_{t_i}(s)$ for $s \in \sigma$. Then

$$|\sigma| \;=\; \frac{|Q_{t_q}|}{|Q_{t_{q-1}}|}\frac{|Q_{t_{q-1}}|}{|Q_{t_{q-2}}|}\cdots\frac{|Q_{t_1}|}{|\omega|}|\omega| \;\le\; K^q\frac{|\gamma_{t_q}(Q_{t_q})|}{|\gamma_{t_q}(Q_{t_{q-1}})|}\cdots\frac{|\gamma_{t_1}(Q_{t_1})|}{|\gamma_{t_1}(\omega)|}|\omega|$$

where K is the distortion constant in (P3). Now $|\gamma_{t_{q+1}}(Q_{t_q})| < 1$, and by (P2)(iii) and (P1)(ii), we have

$$\frac{|\gamma_{t_k}(Q_{t_k})|}{|\gamma_{t_{k+1}}(Q_{t_k})|} \;\le\; K\frac{|I_{\mu_k j_k}|}{|\gamma_{t_k+p_k}(Q_{t_k})|} \;\le\; Ke^{-(1-K\alpha)\mu_k}\ .$$

Thus

$$|\sigma| \;<\; K^q e^{-\sum_{k=1}^q \frac{9}{10}\mu_k + K\alpha\mu_0}|\omega| \;=\; K^q e^{-\frac{9}{10}R + K\alpha\mu_0}|\omega| \;:=\; |\sigma|_R.$$

(For $q = 0$, this estimate presumes that $\gamma_i(\omega)$ has completed its initial bound period, i.e. $n > K\mu_0$.) We estimate $|\{S > n\}|$ by

$$|\{S>n\}| \;=\; \sum_{q,R}|A_{q,R}| \;\le\; \sum_R (\text{number of } \sigma \text{ in } \cup_q A_{q,R}) \cdot |\sigma|_R\ .$$

There are $\begin{pmatrix} R-1 \\ q \end{pmatrix}$ ways of decomposing R into a sum of $q+1$ integers. For a fixed q-tuple $(\mu_1, \cdots, \mu_q)$, we claim there are $\leq 2^q \mu_1^2 \mu_2^2 \cdots \mu_q^2$ possibilities for σ with these data. This is because $\gamma_{t_k}(\mathcal{Q}_{t_k}(\sigma))$ is short enough that it can meet at most one $C_\delta(\hat{x})$, which contains $\leq 2\mu_k^2$ intervals of the form $I_{\mu_k j}$. Furthermore, for $(\mu_1, \cdots, \mu_q)$ with $\mu_1 + \mu_2 + \cdots + \mu_q = R$, we have $\mu_1^2 \mu_2^2 \cdots \mu_q^2 \leq (\frac{R}{q})^{2q}$.

There is one other piece of information that is crucial to us, namely that all bound periods are $\geq \Delta := K^{-1} \log \frac{1}{\delta}$. This means that for a given R, the only feasible q are $\leq \frac{R}{\Delta}$. For a fixed R, then, the number of σ in $\cup_q A_{q,R}$ is

$$\leq \sum_q \begin{pmatrix} R-1 \\ q \end{pmatrix} \cdot 2^q \left(\frac{R}{q}\right)^{2q} \leq \frac{R}{\Delta} \cdot \begin{pmatrix} R \\ \frac{R}{\Delta} \end{pmatrix} \cdot 2^{\frac{R}{\Delta}} \Delta^{2\frac{R}{\Delta}} ,$$

which, by Sterling's formula, is $\sim \frac{R}{\Delta} \left(e^{\varepsilon(\frac{1}{\Delta})} 2^{\frac{1}{\Delta}} \Delta^{\frac{2}{\Delta}} \right)^R$ where $\varepsilon\left(\frac{1}{\Delta}\right) \to 0$ as $\delta \to 0$. Calling the expression above $(1+\eta(\delta))^R$, we have $\eta(\delta) \to 0$ as $\delta \to 0$. Observe also that $n \leq KR + K\mu_0$ by Lemma 2.3, so $R \geq K^{-1}n - \mu_0$. Thus

$$|\{S > n\}| < \sum_{R \geq K^{-1}n - \mu_0} K^q (1+\eta(\delta))^R e^{-\frac{9}{10}R + K\alpha\mu_0} |\omega| < e^{-\frac{4}{5}K^{-1}n + \mu_0} |\omega| < e^{-\frac{1}{2}K^{-1}n} |\omega|$$

provided that $n > 3K\mu_0$. □

Proof of Corollary 2.1: Let $t_1 \geq 0$ be the smallest i such that there are points $s, s' \in \omega$ with different itineraries in their first i iterates. Then either $t_1 = 0$, or $t_1 < K \log \delta^{-1}$ and $|\gamma_{t_1}(\omega)| > K^{-1}\delta$ by (P1). Let n be an arbitrary integer $> t_1$. We partition ω into $\{\hat{\omega}\} \cup \{\omega_{\mu j}\}$ where $s \in \hat{\omega}$ if $\gamma_{t_1}(\mathcal{Q}_{t_1}(s))$ is outside of C_δ and $\gamma_{t_1}(\omega_{\mu j}) \approx I_{\mu j}$. Then $S|_{\hat{\omega}} = t_1$. For μ with $n > 3K\mu$, $|\omega_{\mu j} \cap \{S > n + t_1\}| < Ke^{-\frac{1}{2}K^{-1}n} |\omega_{\mu j}|$ by Lemma 2.4; K here is the distortion constant in (P3). Note also that the total length of $I_{\mu j}$ with $n \leq 3K\mu$ is $\leq 2e^{-\frac{1}{3}K^{-1}n}$. It follows therefore that

$$|\{s \in \omega, S(s) > n + t_1\}| < Ke^{-\frac{1}{2}K^{-1}n} |\omega| + K \left(\frac{2e^{-\frac{1}{3}K^{-1}n}}{K^{-1}\delta} \right) |\omega| < e^{-\hat{K}^{-1}(n+t_1)} |\omega|$$

provided $\hat{K}$ is sufficiently large and $n > \hat{K} \log \delta^{-1}$. □

In the next proof, it is advantageous to take a probabilistic viewpoint, with (ω, P), P being normalized Lebesgue measure, as the underlying probability space.

Proof of Proposition 2.2: Let $\hat{\delta} > 0$ be a small number to be determined, and let B_n be as in the statement of the Proposition. The idea of this proof is to introduce random variables $\hat{X}_i, i = 0, 1, \cdots$, with the property that (i) $B_n \leq \sum_{i \leq n} \hat{X}_i$ and (ii) the conditional expectations of $\hat{X}_i$ are dominated by certain exponential random variables.

Step I. Reformulation of problem as one involving $\sum_{i \leq n} \hat{X}_i$

Define a sequence of random variables $t_1 < t_2 < \cdots$ marking certain intersection times with $C_{\hat{\delta}}$ as follows: If $I_{\mu_0 j_0} \subset C_{\hat{\delta}}$, let $t_1 = 0$, and let S_1 be the stopping time S defined in Sect. 2.3. If $I_{\mu_0 j_0} \cap C_{\hat{\delta}} = \emptyset$, let t_1 be the smallest i for which $\gamma_i(\mathcal{Q}_{i-1}(s)) \cap C_{\hat{\delta}} \neq \emptyset$, and define S_1 on each element of $\mathcal{Q}_{t_1}$ as follows: If $\gamma_{t_1}(\mathcal{Q}_{t_1}(s)) \cap C_{\hat{\delta}} = \emptyset$, set $S_1(s) = 0$. If $\gamma_{t_1}(\mathcal{Q}_{t_1}(s)) \approx I_{\mu j} \subset C_{\hat{\delta}}$, let S_1 be the stopping time S on $\mathcal{Q}_{t_1}(s)$ for the sequence $\gamma_{t_1}, \gamma_{t_1+1}, \cdots$ (instead of $\gamma_0, \gamma_1, \cdots$); that is to say, $S_1(s)$ is the smallest i such that $\gamma_{t_1+i}(\mathcal{Q}_{t_1+i-1}(s))$ is not in a bound period and $|\gamma_{t_1+i}(\mathcal{Q}_{t_1+i-1}(s))| > \delta$. Then on each element of $\mathcal{Q}_{t_1}$, we define t_2 to be the smallest $i \geq t_1 + S_1$

such that $\gamma_i(\mathcal{Q}_{i-1}(s)) \cap C_{\hat{\delta}} \neq \emptyset$, and on each $\mathcal{Q}_{t_2}(s)$, define S_2 to be either S or 0 as before depending on whether $\gamma_{t_2}(\mathcal{Q}_{t_2}(s)) \subset C_{\hat{\delta}}$ or not, and so on.

Before proceeding further, we record the following lower estimate on $|\gamma_{t_i}(\mathcal{Q}_{t_i-1}(s))|$. Let t be the time $\mathcal{Q}_{t_i-1}(s)$ is created. By definition, $t_{i-1} + S_{i-1} \leq t < t_i$, and $\gamma_t(\mathcal{Q}_{t_i-1}(s)) \approx P$ for some $P \in \mathcal{P}$. Moreover, there are only two possibilities: either P is outside and $|P| \geq \delta$, or it is $\approx I_{\mu j}$ for some $I_{\mu j} \subset C_\delta \setminus C_{\hat{\delta}}$. By (P1) and (P2)(iii), $|\gamma_{t_i}(\mathcal{Q}_{t_i-1}(s))| > \hat{\delta}' := \min(c_2\delta, \hat{\delta}^{K_1\alpha})$. Note that if $\hat{\delta} << \delta$, then $\hat{\delta}' >> \hat{\delta}$.

We now head toward the promised random variables. For $i = 0, 1, 2, \cdots$, let $X_i(s) = 1$ for $i \in [t_k, t_k + S_k)$, any k, and $= 0$ otherwise. Then $B_n \leq \sum_{i \leq n} X_i$; in fact, this is likely to be an overcount, for S_k goes beyond bound periods. It is thus sufficient to show that $P\{\sum_{i\leq n} X_i > \sigma n\}$ decreases exponentially with n.

As we will see, it simplifies the discussion slightly if we "speed up time" to skip over the intervals $[t_k, t_k + S_k)$. Let $T_{-1} = -1$. With T_i defined, we let $T_{i+1} = T_i + 1$ except when $T_i + 1 = t_k$, in which case we let $T_{i+1} = T_i + 1 + S_k$. We let $\hat{X}_0 = S$ if $\omega \subset C_{\hat{\delta}}$, 0 otherwise, and let $\hat{X}_{i+1} = S_k$ if $T_i + 1 = t_k$, 0 otherwise. Let $\mathcal{Q}_{T_i}$ be the partition defined by $\mathcal{Q}_{T_i}(s) = \mathcal{Q}_{T_i(s)}(s)$, and note that $\hat{X}_i$ is measurable with respect to $\mathcal{Q}_{T_i}$. Since $X_i \leq \hat{X}_i$, it is all the more true that $B_n \leq \sum_{i\leq n} \hat{X}_i$.

Step II. Conditional distribution of $\hat{X}_{i+1}$ given $\mathcal{Q}_{T_i}$

Let $i \geq 0$, and consider $Q \in \mathcal{Q}_{T_i}$. On most Q, $\hat{X}_{i+1}$ is identically equal to 0. The only time when this is not the case is when $\gamma_{T_i(Q)+1}(Q)$ meets $C_{\hat{\delta}}$. We note that

(1) for all $s, s' \in Q$, $\gamma'_{T_i(Q)+1}(s)/\gamma'_{T_i(Q)+1}(s') < K$;

(2) $|\gamma_{T_i(Q)+1}(Q)| > \hat{\delta}'$.

(1) follows from (P3); (2) is from Step I. From (1) and (2), we deduce that (i) $P(\hat{X}_{i+1} = 0 \mid Q) \geq 1 - K\hat{\delta}\hat{\delta}'^{-1}$ and (ii) $P(\hat{X}_{i+1} > n \mid Q \cap \{\gamma_{T_i+1} \in I_{\mu j}\}) < Ke^{-\frac{1}{2}K^{-1}n}$ if $n \geq 3K\mu$ (Lemma 2.5); for $n < 3K\mu$, there is no information. It follows that for all $n \geq 0$,

$$P(\hat{X}_{i+1} > n \mid Q) < K\hat{\delta}'^{-1} \min(\hat{\delta}, e^{-(3K)^{-1}n}) + K\hat{\delta}\hat{\delta}'^{-1}e^{-\frac{1}{2}K^{-1}n}. \tag{19}$$

A simple computation shows that if $\varepsilon < \frac{1}{6}K^{-1}$ (where K is as in the exponents above), then $E[e^{\varepsilon \hat{X}_{i+1}}|Q] < \infty$. We note further that by decreasing $\hat{\delta}$ (keeping ε fixed), $E[e^{\varepsilon \hat{X}_{i+1}}|Q]$ can be made arbitrarily close to 1. Let $\eta > 0$ be a number to be determined shortly, and choose $\hat{\delta} = \hat{\delta}(\eta)$ sufficiently small that $E[e^{\varepsilon \hat{X}_{i+1}}|Q] < e^\eta$. Observing that the upper bound in (19) and hence that for $E[e^{\varepsilon \hat{X}_{i+1}}|Q]$ do not depend on i or on Q, we conclude that with $\hat{\delta} = \hat{\delta}(\eta)$ as above, $E[e^{\varepsilon \hat{X}_{i+1}}|\mathcal{Q}_{T_i}] < e^\eta$ for every $i \geq 0$.

Step III. Large deviation estimate for $\sum_{i\leq n} \hat{X}_i$

To finish, we write

$$E\left[e^{\varepsilon \sum_{i\leq n} \hat{X}_i}\right] = E\left[E[e^{\varepsilon \sum_{i\leq n} \hat{X}_i}|\mathcal{Q}_{T_{n-1}}]\right] = E\left[e^{\varepsilon \sum_{i<n} \hat{X}_i} E[e^{\varepsilon \hat{X}_n}|\mathcal{Q}_{T_{n-1}}]\right] < e^\eta E\left[e^{\varepsilon \sum_{i<n} \hat{X}_i}\right],$$

giving inductively $E[e^{\varepsilon \sum_{i\leq n} \hat{X}_i}] < e^{n\eta} E[e^{\varepsilon \hat{X}_0}]$. Since $E[e^{\varepsilon \hat{X}_0}] < e^{K\varepsilon\mu_0}$, we arrive at

$$P\{B_n > \sigma n\} < P\left\{\sum_{i\leq n} \hat{X}_i > \sigma n\right\} < e^{\eta n - \varepsilon\sigma n + K\varepsilon\mu_0}.$$

This is $< e^{-\frac{1}{2}\varepsilon\sigma n}$ if η is chosen $< \frac{1}{4}\varepsilon\sigma$ and n is $> K\mu_0\sigma^{-1}$. □

A.3 Parameter transversality (Sect. 2.4)

Proof of Lemma 2.6: First we show that $\cup_{i\geq 0} f^{-i}C$ is dense in I. If not, there would be an interval ω with the property that $\phi(x)$ is identical for all $x \in \omega$. Let ω be a maximal interval of this type. Then either (i) $f^{n+k}(\omega) \subset f^n(\omega)$ for some n, k, or (ii) $f^k(\omega), k = 0, 1, \cdots$, are pairwise disjoint. Case (i) cannot happen since it implies the presence of a periodic point x with $|(f^k)'x| \leq 1$. Case (ii) is equally absurd, for it implies the existence of $\{k_i\}$ where $f^{k_i}(\omega)$ are arbitrarily short and arbitrarily close to C, a scenario not permitted by Definition 1.1 (c)(ii) and Lemma 2.1.

For each n, let $l_n(\hat{x})$ and $r_n(\hat{x})$ be the two points in $\cup_{0\leq i\leq n} f^{-i}C$ closest to $\hat{x} \in C$. In the case $I = S^1$, let $\Lambda^{(n)} = \{x \in I : f^i x \notin \cup_{\hat{x}\in C}(l_n(\hat{x}), r_n(\hat{x}))\ \forall i \geq 0\}$. If I is an interval, we may assume n is large enough that $f(I) \subset (z_n^1, z_n^2)$ where z_n^1 and z_n^2 are the two points in $\cup_{0\leq i\leq n} f^{-i}C$ closest to the ends of I. We then define $\Lambda^{(n)}$ as in the circle case but with I replaced by $[z_n^1, z_n^2]$. In both cases, $\Lambda^{(n)}$ is compact and $f(\Lambda^{(n)}) \subset \Lambda^{(n)}$. Clearly, $\cup\Lambda^{(n)}$ is dense in I since $\cup_{i\geq 0} f^{-i}C$ is dense in I and the gaps in $\Lambda^{(n)}$ decrease in size as n increases.

For part (a), it remains to show that $f|_{\Lambda^{(n)}}$ is conjugate to a shift of finite type. Let $\mathcal{J}^{(n)} = \{J_i^{(n)}\}$ be the partition of I by $\cup_{0\leq i\leq n} f^{-i}C$. Observe that for $J_i(n) \neq (l_n(\hat{x}), \hat{x})$ or $(\hat{x}, r_n(\hat{x}))$, $f(J_i^{(n)})$ is equal to the union of a finite number of elements of $\mathcal{J}^{(n)}$. Let $\Lambda_i^{(n)} = \Lambda^{(n)} \cap J_i^{(n)}$. Then the alphabet of the shift in question is $\{i : \Lambda_i^{(n)} \neq \emptyset\}$, and the transition $i \to j$ is admissible if $f(\Lambda_i^{(n)}) \supset \Lambda_j^{(n)}$.

Assertion (b) follows from our construction. □

Proof of Corollary 2.2: Fix n large enough that for all $i \geq 0$, $f^i(q) \notin (l_n(\hat{x}), r_n(\hat{x}))$ for all $\hat{x} \in C$, and let $\Lambda = \Lambda^{(n)}$. Let $B = \cup_i \partial\bar{\Lambda}_i$ where $\bar{\Lambda}_i$ is the shortest interval containing Λ_i. Since B is a finite set with $f(B) \subset B$, it consists of pre-periodic points. From Lemma 2.1, these periodic points are repelling. Thus if g is sufficiently near f, there is a unique set B_g with $g(B_g) \subset B_g$ such that $g|_{B_g}$ is conjugate to $f|_B$. Using B_g, we recover a set Λ_g on which g is conjugate to $f|_\Lambda$. The uniqueness of q_g follows from the expanding property of g away from C (Lemma 2.2). □

Proof of Proposition 2.3: (i) We prove $a \mapsto q(a)$ is differentiable with

$$\frac{d}{da}q(a) \ = \ -\sum_{i=1}^{\infty} \frac{\partial_a f_a(f_a^{i-1}(q))}{(f_a^i)'(q)}. \tag{20}$$

Here all objects depend on a, mention of which is often suppressed (e.g. $f = f_a, q = q(a)$). Continuing to use the notation in Corollary 2.2, we let $\Lambda_{i_0,i_1,\cdots,i_n} = \{x \in I : f^j(x) \in \Lambda_{i_j}, 0 \leq j \leq n\}$, and let $\Lambda_{i_0,i_1,\cdots,i_n}(q)$ be the cylinder set containing q. For each n, choose $q_n \in \partial\bar{\Lambda}_{i_0,i_1,\cdots,i_n}(q)$. We will show that as functions of a, $\frac{d}{da}q_n$ converges uniformly to the right side of (20). This requires in particular the uniform bound $\max_{i_0,i_1,\cdots,i_n} |\bar{\Lambda}_{i_0,i_1,\cdots,i_n}| < Ke^{-\frac{1}{4}\lambda_0 n}$ for all $n > 0$ (see Lemma 2.1).

Introduce $G(x,a) = (f_a(x), a)$, and let $G^n(q_n, a) = (p_n, a)$. Then $p_n \in B$. Differentiating, we obtain $\frac{d}{da}p_n = \sum_{i=1}^n \partial_x f^{n-i}(f^i q_n)\partial_a f(f^{i-1}q_n) + \partial_x f^n(q_n)\ \frac{d}{da}q_n$. Hence we have

$$\frac{d}{da}q_n \ = \ \frac{\frac{d}{da}p_n}{\partial_x f^n(q_n)} \ - \ \sum_{i=1}^{n} \frac{\partial_a f(f^{i-1}q_n)}{\partial_x f^i(q_n)} \ . \tag{21}$$

Since B is a finite set, $\frac{d}{da}p_n$ is uniformly bounded for all n. With $|(f^n)'(q_n)|$ growing exponentially, the first term on the right is exponentially small. It remains to check that the second term converges uniformly to the right side of (20). In addition to the growth of f^n, this uses our estimates on $\max|\bar{\Lambda}_{i_0,i_1,\cdots,i_n}|$ above and a distortion estimate for f^i. We leave it as an exercise.

(ii) This is a direct application of (20) to q defined by $q(a^*) = f_{a^*}(\hat{x})$:

$$\frac{\frac{d}{da}f^n(\hat{x})}{\partial_x f^{n-1}(\hat{x}_1)} \;=\; \frac{1}{\partial_x f^{n-1}(\hat{x}_1)}\left(\partial_x f^{n-1}(\hat{x}_1)\frac{d\hat{x}_1}{da} + \partial_a f^{n-1}(\hat{x}_1)\right) \;=\; \frac{d\hat{x}_1}{da} + \frac{\partial_a f^{n-1}(\hat{x}_1)}{\partial_x f^{n-1}(\hat{x}_1)}\;.$$

Thus the limit as $n \to \infty$ at $a = a^*$ differs from $\frac{dq}{da}(a^*)$ by $\frac{d\hat{x}_1}{da}(a^*)$, which is also easily seen to be the term corresponding to $i = 0$ in the sum in Proposition 2.3(ii). □

A.4 Most contracted directions: Preliminaries (Sect. 3.1)

We record in this Appendix some elementary estimates in preparation for the proof of Lemma 3.1.

I. Area growth

Let $\{E_1, E_2, \cdots, E_m\}$ denote the usual basis of $\mathbb{R}^m$. Recall that if $u = \sum u_i E_i$ and $v = \sum v_i E_i$, then $u \wedge v = \sum_{i<j} I_{ij} E_i \wedge E_j$ where $I_{ij} = u_i v_j - v_i u_j$, and the area of the parallelogram spanned by u and v is equal to

$$|u \wedge v| \;=\; \sqrt{|u|^2|v|^2 - \langle u, v\rangle^2} \;=\; \left(\sum_{i<j} I_{ij}^2\right)^{\frac{1}{2}}.$$

Sublemma A.4.1 *Let M and $\tilde{M}$ have the properties in (H1) in Sect. 3.1. Then*
(a) $|Mu \wedge Mv| < Kb\ |u \wedge v|$;
(b) $|Mu \wedge \tilde{M}v| \le Kb\ |u||v|$.

Consider next linearly independent unit vectors $u^{(0)}$ and $v^{(0)}$ parameterized by $s = (s_1, s_2)$. For $n = 1, 2, \cdots$, let $u^{(n)} = M_n u^{(n-1)}$ and $v^{(n)} = M_n v^{(n-1)}$, and let $u^{(n)} \wedge v^{(n)} = \sum I_{ij}^{(n)} E_i \wedge E_j$.

Sublemma A.4.2 *Assume $M_i, u^{(0)}$ and $v^{(0)}$ satisfy (H2). Then for $k = 0, 1, 2,$*

$$|\partial^k (u^{(n)} \wedge v^{(n)})| \;<\; (Kb)^n.$$

It follows that if $\Delta_n = |u^{(n)} \wedge v^{(n)}|$, then $|\partial^k \Delta_n^2| < (Kb)^{2n}$.

II. Formulas for e and f

We fix $M \in \mathrm{L}(m, \mathbb{R})$ and $S = S(u, v)$ where for simplicity we assume u and v are unit vectors with $u \perp v$. The formulas below all pertain to $M|_S$; mention of S is suppressed (e.g. we write $e = e(S)$) except where ambiguity arises. The following formulas are results of elementary computations:

First, we write down the squares of the singular values of $M|_S$:

$$|Me|^2 = \frac{1}{2}(B - \sqrt{B^2 - 4C}) := \lambda, \qquad |Mf|^2 = \frac{1}{2}(B + \sqrt{B^2 - 4C})$$

where $B = |Mu|^2 + |Mv|^2, C = |Mu \wedge Mv|^2$. (Note that the formulas above are in agreement with $|Mu \wedge Mv| = |Me||Mf|$.) We then write $e = \alpha_0 u + \beta_0 v$, and solve for $|Me| = \sqrt{\lambda}$ subject to $\alpha_0^2 + \beta_0^2 = 1$. There are two solutions (a vector and its negative): either $e = \pm v$, or the solution with a positive u-component is given by

$$e = \frac{1}{Z}(\alpha u + \beta v) \tag{22}$$

with $\alpha = |Mv|^2 - \lambda, \beta = -\langle Mv, Mu\rangle$ and $Z = \sqrt{\alpha^2 + \beta^2}$. From this we deduce that a solution for f is $f = \frac{1}{Z}(-\beta u + \alpha v)$.

A.5 Most contracted directions: Proof of Lemma 3.1 (Sect. 3.1)

I. Proof of Lemma 3.1(a)

We assume M_i satisfies (H1) and let $S = S(u,v)$. As before, mention of S is suppressed. Recall that $\Delta_i := |M^{(i)}u \wedge M^{(i)}v|$.

Sublemma A.5.1 *(i)* $\Delta_i < (Kb)^i$*;*

(ii) $|M^{(i)}e_i| < \left(\frac{Kb}{\kappa}\right)^i$*;*

(iii) $|M^{(i+1)}f_i| = |M^{(i+1)}f_{i+1}| \pm \mathcal{O}((\frac{Kb}{\kappa})^i) \gtrsim K_0^{-1}\kappa^i$*;*

(iv) If we substitute $u = e_i, v = f_i$ *and* $M = M^{(i+1)}$ *into the formulas in Appendix A.4, part II,*

and let $\alpha_{i+1}, \beta_{i+1}$ *and* Z_{i+1} *be the resulting quantities, then* $Z_{i+1} \approx |\alpha_{i+1}| \approx |M^{(i+1)}f_i|^2$*.*

Proof: (i) is the $k = 0$ case of Sublemma A.4.2. For (ii), write $|M^{(i)}e_i| = \frac{\Delta_i}{|M^{(i)}f_i|}$; the assertion follows from (i) and our assumption on $|M^{(i)}f_i|$. Now make the substitution in (iv). From the formula for $|Mf|$, we see that

$$|M^{(i+1)}f_{i+1}|^2 = B_{i+1} + \mathcal{O}(C_{i+1}) = |M^{(i+1)}f_i|^2 + |M_{i+1}M^{(i)}e_i|^2 + \mathcal{O}(C_{i+1});$$

estimates for the last two terms given by (ii) and (i). This proves (iii). (iv) is now obvious. $\Diamond$

We now prove Lemma 3.1(a). Continuing to substitute $u = e_i$ and $v = f_i$ in the formulas in Appendix A.4, we have, from (22),

$$e_{i+1} - e_i = \frac{1}{Z_{i+1}} \left(\frac{-\beta_{i+1}^2}{\alpha_{i+1} + Z_{i+1}} e_i + \beta_{i+1} f_i \right). \tag{23}$$

To estimate $|e_{i+1} - e_i|$, then, we need to obtain a suitable upper bound for $|\beta_{i+1}|$ and lower bounds for $|\alpha_{i+1}|$ and Z_{i+1}. Sublemma A.5.1 gives

$$|\beta_{i+1}| \le |M^{(i+1)}e_i||M^{(i+1)}f_i| < \left(\frac{Kb}{\kappa}\right)^i \sqrt{Z_{i+1}} \tag{24}$$

and $|\alpha_{i+1}| \approx Z_{i+1}$. These estimates together with $Z_{i+1} > K_0^{-1}\kappa^{2i}$ tell us $|e_{i+1} - e_i| \approx \frac{|\beta_{i+1}|}{Z_{i+1}} < \left(\frac{Kb}{\kappa^2}\right)^i$. The second assertion follows easily from $|M^{(i)}e_n| \le |M^{(i)}(e_n - e_{n-1})| + \cdots + |M^{(i)}(e_{i+1} - e_i)| + |M^{(i)}e_i| < \left(\frac{Kb}{\kappa^2}\right)^i$. $\square$

II. Proof of Lemma 3.1(b): First derivative estimates

For this part we assume M_i and S satisfy (H2) with C^2 norms replaced by C^1 norms. Let ∂ denote a fixed partial derivative.

Sublemma A.5.2 $|\partial e_1|, |\partial f_1| < K_1$ *for some* K_1*.*

Proof: Switching u and v in (H2) if necessary, we may assume $|M_1 v| \ge |M_1 u|$. Then from Appendix A.4.II we have $Z_1 > \alpha \ge |M_1 v|^2 - Kb > \frac{1}{2}B - Kb > \frac{1}{4}K_0^{-2}$. Differentiating (22) gives the desired result. $\Diamond$

Our plan of proof is as follows: For $k = 1, 2, \cdots$, we assume for all $i \le k$

(*) $|\partial e_i|, |\partial f_i| < 2K_1$ where K_1 as in Sublemma A.5.2,

and prove for all $i \le k$:

(A) $|\partial(M^{(i)}f_i)| < K^i$, $|\partial(M^{(i)}e_i)| < \left(\frac{Kb}{\kappa^2}\right)^i$;

(B) $|\partial(e_{i+1}-e_i)|$, $|\partial(f_{i+1}-f_i)| < (\frac{Kb}{\kappa^3})^i$.

Observe that for $i=1$, (*) is given by Sublemma A.5.2. It is easy to see that (B) above implies (*) with $i=k+1$, namely $|\partial f_{k+1}| \le |\partial(f_{k+1}-f_k)| + \cdots + |\partial(f_2-f_1)| + |\partial f_1|$. ¿From (B), we have $|\partial(f_{i+1}-f_i)| < \left(\frac{Kb}{\kappa^3}\right)^i$, and from Sublemma A.5.2, we have $|\partial f_1| < K_1$. Hence $|\partial f_{k+1}| < \frac{Kb}{\kappa^3} + K_1$, which, for b sufficiently small, is $< 2K_1$. The computation for e_{k+1} is identical.

Proof of (*) $\Longrightarrow$ (A): First we prove the estimate for $\partial(M^{(i)}f_i)$. Writing $\partial(M^{(i)}f_i) = \sum_{j=1}^i M_i\cdots(\partial M_j)\cdots M_1 f_i + M^{(i)}\partial f_i$, we obtain easily $|\partial(M^{(i)}f_i)| \le \sum_{j=1}^i |M_i\cdots(\partial M_j)\cdots M_1 f_i| + \|M^{(i)}\||\partial f_i| \le iK^i + K^i(2K_1)$.

This estimate is used to estimate $\partial(M^{(i)}e_i)$.[15] Write $\partial(M^{(i)}e_i) = (I)+(II)$ where (I) is its component in the direction of $M^{(i)}f_i$ and (II) is its component orthogonal to $M^{(i)}f_i$. Recall that $\partial\langle M^{(i)}e_i, M^{(i)}f_i\rangle = 0$. We have

$$|(I)| = \left|\langle\partial(M^{(i)}e_i), \frac{M^{(i)}f_i}{|M^{(i)}f_i|}\rangle\right| = \frac{1}{|M^{(i)}f_i|}\ |\langle M^{(i)}e_i, \partial(M^{(i)}f_i)\rangle| < \frac{1}{\kappa^i}\left(\frac{Kb}{\kappa}\right)^i K^i\ ;$$

$$|(II)|\ |M^{(i)}f_i| = |\partial(M^{(i)}e_i)\wedge M^{(i)}f_i| \le |\partial(M^{(i)}e_i\wedge M^{(i)}f_i)| + |M^{(i)}e_i\wedge\partial(M^{(i)}f_i)|.$$

The first term in the last line is $< (Kb)^i$ by Sublemma A.4.2, noting that we have established $|\partial e_i|, |\partial f_i| < 2K_1$; the second term is $< (\frac{Kb}{\kappa})^i\cdot K^i$. This completes the proof of (A). ◊

To prove (B), we first compute some quantities associated with the next iterate. Substitute $u=e_i, v=f_i, M=M^{(i+1)}$ into the formulas in Appendix A.4, and let $B_{i+1}, C_{i+1}, \lambda_{i+1}$ etc. be the resulting quantities. The following is a straightforward computation.

Sublemma A.5.3 *Assume (*) and (A). Then for all $i\le k$:*
(a) $|\partial\lambda_{i+1}| < \left(\frac{Kb}{\kappa^2}\right)^{2(i+1)}$;
(b) $|\partial\beta_{i+1}| < (\frac{Kb}{\kappa^2})^i\sqrt{Z_{i+1}}$;
(c) $|\partial\alpha_{i+1}|, |\partial Z_{i+1}| < K^i\sqrt{Z_{i+1}}$.

Proof of (*), (A) $\Longrightarrow$ (B): We work with e_i; the computation for f_i is similar. From (23) we have $\partial(e_{i+1}-e_i) = (III)+(IV)+(V)$ where

$$\begin{aligned}
|(III)| &= |\frac{1}{Z_{i+1}}(e_{i+1}-e_i)\partial_1 Z_{i+1}| < \frac{K^i\sqrt{Z_{i+1}}}{Z_{i+1}}\cdot\left(\frac{Kb}{\kappa^2}\right)^i < \left(\frac{Kb}{\kappa^3}\right)^i;\\
|(IV)| &= |\frac{1}{Z_{i+1}}\partial(\beta_{i+1}f_i)| < \frac{1}{Z_{i+1}}(|\partial\beta_{i+1}| + |\beta_{i+1}||\partial f_i|) < \left(\frac{Kb}{\kappa^3}\right)^i;\\
|(V)| &= |\frac{1}{Z_{i+1}}\partial\left(\frac{\beta_{i+1}^2}{\alpha_{i+1}+Z_{i+1}}e_i\right)| << \left(\frac{Kb}{\kappa^3}\right)^i.
\end{aligned}$$

To estimate (III), we have used Sublemmas A.5.1, A.5.3(c) and part (a) of Lemma 3.1. To estimate (IV), we have used Sublemma A.5.3(b), (*) and $|\beta_{i+1}| < (\frac{Kb}{\kappa})^i$. The estimate for (V) is easy. ◊

This completes the proofs of the first derivative estimates in Lemma 3.1(b). □

[15]We thank O. Lanford for showing us this argument.

III. Proof of Lemma 3.1(b): Second derivative estimates

We now assume the full force of (H2). The proof proceeds in a manner entirely analogous to that for first derivatives: We first prove $|\partial^2 e_1|, |\partial^2 f_1| < K_1'$ for some K_1'. Then for $k = 1, 2, \cdots$, we assume for all $i \leq k$

(*') $|\partial^2 e_i|, |\partial^2 f_i| < 2K_1'$,

and prove for all $i \leq k$:

(A') $|\partial^2(M^{(i)} f_i)| < K^i$, $|\partial^2(M^{(i)} e_i)| < \left(\frac{Kb}{\kappa^3}\right)^i$;
(B') $|\partial^2(e_{i+1} - e_i)|$, $|\partial^2(f_{i+1} - f_i)| < (\frac{Kb}{\kappa^4})^i$.

Details are left to the reader. □

A.6 A perturbation lemma (Sect. 3.2)

Proof of Lemma 3.2: Assume inductively that $\angle(w_i, w_i') < \eta^{\frac{i+1}{4}}$ for all $i < n$. Let $n = 2j$ (or $2j \pm 1$). Let $u_j = \frac{w_j}{|w_j|}$, $A = DT^j_{z_j}$, and let u_j' and A' be the corresponding quantities for (z_0', w_0'). Since $|w_j| < K^j$ and $|w_{2j}| > K_0^{-1} \kappa^{2j-1}$ by hypothesis, we have

$$|Au_j| = \frac{|w_{2j}|}{|w_j|} > \left(\frac{\kappa^2}{K}\right)^j. \tag{25}$$

We observe first that $|A'u_j'| \gtrsim (\frac{\kappa^2}{K})^j$: Clearly, $|A'u_j'| \geq |Au_j| - \|A\||u_j - u_j'| - \|A - A'\||u_j'|$. The desired estimate follows from the fact that $\|A\||u_j - u_j'| \approx \|A\|\angle(u_j, u_j') \leq K^j \eta^{\frac{j+1}{4}}$, $\|A - A'\| = |DT^j_{z_j} - DT^j_{z_j'}| \leq jK^j \eta^{j+1}$, and both of these quantities are $<< (\frac{\kappa^2}{K})^j$ by the relation imposed on η and κ.

We estimate $\angle(u_{2j}, u_{2j}') \approx |u_{2j} \wedge u_{2j}'|$ by

$$|u_{2j} \wedge u_{2j}'| \leq \frac{|Au_j \wedge Au_j'|}{|Au_j| \cdot |A_j' u_j'|} + \frac{|Au_j \wedge (A - A')u_j'|}{|Au_j| \cdot |A_j' u_j'|}.$$

The first term is $< (Kb)^j \eta^{\frac{j+1}{4}} (\frac{K}{\kappa^2})^{2j}$. The second term is $< K^j (K\eta)^{j+1} (\frac{K}{\kappa^2})^{2j}$. Both are $< \frac{1}{2}\eta^{\frac{n+1}{4}}$ by the relations we imposed on b, η and κ with K_1 appropriately chosen. This completes the proof of (b) for $n = 2j$.

To prove (a), we write

$$\frac{|w_{2j}'|}{|w_j'|} \ \geq \ \frac{|w_{2j}|}{|w_j|} \left(1 - \frac{|w_j|}{|w_{2j}|}(\|A' - A\| + \|A\||u_j' - u_j|)\right).$$

Using the same bounds as before, we see that the factor inside parentheses is $> 1 - [(\frac{\kappa^2}{K})^j((K\eta)^{j+1} + K^j \eta^{\frac{j+1}{4}})] > 1 - (\frac{1}{4})^j$. This proves $|w_{2j}'| \geq |w_{2j}|(\sum_{1 \leq i \leq j} \frac{1}{4^i})$. □

A.7 Temporary stable curves (Sect. 3.3)

Proof of Proposition 3.1: Let B_0 be the ball of radius η in S centered at z_0. Then on B_0 we have, by Lemma 3.2, $\|DT|_S\| \geq \frac{1}{2}K_0^{-1}$, so that $e_1(S)$ is well defined. Let γ_1 be the integral curve to $e_1(S)$ defined for $s \in (-\eta, \eta)$ with $\gamma_1(0) = z_0$. Note that $|DT(e_1)| < Kb$.

To construct γ_2, let B_1 be the $\frac{\eta^2}{2K_0}$-neighborhood of γ_1 in S where K_0 is a constant related to $\|T\|_{C^2}$. For $\xi \in B_1$, let ξ' be a point in γ_1 with $|\xi - \xi'| < \frac{\eta^2}{2K_0}$. Then $|T\xi - Tz_0| \leq |T\xi - T\xi'| + |T\xi' - Tz_0| \leq \frac{\eta^2}{2} + \frac{Kb}{\kappa^2}\eta < \eta^2$. Thus by Lemma 3.2, $||DT^2_\xi|_S\| \geq \frac{1}{2}K_0^{-1}\kappa$. This ensures that $e_2(S)$ is defined on all of B_1. Let γ_2 be the integral curve to $e_2(S)$ with $\gamma_2(0) = z_0$. We verify that γ_2 is defined on $(-\eta, \eta)$ and runs alongside γ_1. More precisely,

$$\begin{aligned} |\frac{d}{ds}(\gamma_2(s) - \gamma_1(s))| &\leq |e_2(\gamma_2(s)) - e_1(\gamma_2(s))| + |e_1(\gamma_2(s)) - e_1(\gamma_1(s))| \\ &\leq |e_2 - e_1| + |\frac{d}{ds}e_1||\gamma_2(s) - \gamma_1(s)| \;\leq\; \frac{Kb}{\kappa^2} + \frac{K}{\kappa^3}|\gamma_2(s) - \gamma_1(s)| \end{aligned}$$

by Lemma 3.1. By Gronwall's inequality, $|\gamma_2(s) - \gamma_1(s)| \leq \frac{Kb}{\kappa^2}|s|e^{\frac{K}{\kappa^3}|s|}$, which is $<< \frac{\eta^2}{2K_0}$ for $|s| < \eta$. This ensures that γ_2 remains in B_1 and hence is well defined for all $s < \eta$.

In general, we inductively construct γ_i by letting B_{i-1} be the $\frac{\eta^i}{2K_0^{i-1}}$-neighborhood of γ_{i-1} in S. Then for all $\xi \in B_{i-1}$, $|T^j\xi - T^j z_0| < \eta^{j+1}$ for $k < i$. Thus by Lemma 3.2, $\|DT^i_\xi|_S\| \geq \frac{1}{2}K_0^{-1}\kappa^{i-1}$, and so e_i is well defined. Integrating and arguing as above, we obtain γ_i with $|\gamma_i(s) - \gamma_{i-1}(s)| < K(\frac{Kb}{\kappa^2})^{i-1}|s| << \frac{\eta^i}{2K_0^{i-1}}$ for all s with $|s| < \eta$. □

A.8 A curvature estimate (Sect. 3.4)

Proof of Lemma 3.3: Recall that

$$k_i(s) = \frac{|\gamma_i'(s) \wedge \gamma_i''(s)|}{|\gamma'(s)|^3}.$$

Since $\gamma_i' = DT_{\gamma_{i-1}}(\gamma_{i-1}')$, we have $\gamma_i'' = (\frac{d}{ds}DT_{\gamma_{i-1}})(\gamma_{i-1}') + DT_{\gamma_{i-1}}(\gamma_{i-1}'')$. Thus $k_i \leq \frac{1}{|\gamma_i'|^3}(I + II)$ where

$$I = |DT(\gamma_{i-1}') \wedge DT(\gamma_{i-1}'')|, \quad II = |DT(\gamma_{i-1}') \wedge (\frac{d}{ds}DT_{\gamma_{i-1}})(\gamma_{i-1}')|.$$

Since $DT = DT_{\gamma_{i-1}}$ has the form in (H1) in Sect. 3.1, we have $I < Kb|\gamma_{i-1}' \wedge \gamma_{i-1}''|$ (see Sublemma A.4.1). Observe that $\frac{d}{ds}DT_{\gamma_{i-1}}$ has the same form with K_0 replaced by $K_0|\gamma_{i-1}'|$, i.e. if $\frac{d}{ds}DT_{\gamma_{i-1}} = M = (\hat{M}^1, \cdots, \hat{M}^m)$, then $\|\hat{M}^1\| < K_0|\gamma_{i-1}'|$ and $\|\hat{M}^j\| < K_0|\gamma_{i-1}'|b$ for $j \geq 2$. Thus $II < Kb|\gamma_{i-1}'|^3$, and so

$$k_i \leq (Kb \cdot k_{i-1} + Kb)\frac{|\gamma_{i-1}'|^3}{|\gamma_i'|^3} = \sum_{j=1}^{i-1}(Kb)^j\frac{|\gamma_{i-j}'|^3}{|\gamma_i'|^3} + (Kb)^i\frac{|\gamma_0'|^3}{|\gamma_i'|^3}\kappa_0 \;\leq\; \frac{Kb}{\kappa^3}.$$

□

A.9 Properties of e_1 in $\mathcal{C}^{(1)}$ (Sect. 3.6)

Proof of Lemma 3.7: Consider $T_0 = (\hat{T}^1, 0, \cdots, 0)$ acting at $(x, 0)$, $x \in C_\delta$, and let $e_1 = e_1(T_0, (x,0); S(\frac{\partial}{\partial x}, \mathbf{v}))$ be the most contracted direction of DT_0 at $(x, 0)$ on the plane indicated. Since $\det(DT_0(e_1)) = 0$, it is easy to see that

$$\langle e_1, \mathbf{v}\rangle \;=\; \pm\,\frac{f'(x)}{\sqrt{|DT_0(\mathbf{v})|^2 + (f'(x))^2}}\;,$$

the sign depending on the orientation of $DT_0(\mathbf{v})$. Using the facts that $|f''| > K^{-1}$ and $|DT_0(\mathbf{v})| > K^{-1}$, one verifies readily that $|\frac{d}{dx}e_1(T_0, (x,0); S(\frac{\partial}{\partial x}, \mathbf{v}))| > K^{-1}$.

Observe that if b is sufficiently small, then by continuity, $e_1(T, \gamma(x); S(\gamma'(x), \mathbf{v}))$ is defined everywhere on γ. We compare it to $e_1(T_0, (x,0); S(\frac{\partial}{\partial x}, \mathbf{v}))$:

First, we continue to focus on $(x, 0)$ and $S = S(\frac{\partial}{\partial x}, \mathbf{v})$, and interpolate between T_0 and T by introducing $T_s := (\hat{T}^1, \frac{s}{\sqrt{b}}\hat{T}^2, \cdots, \frac{s}{\sqrt{b}}\hat{T}^m)$, $s \in [0, \sqrt{b}]$. More precisely, we consider the

2-parameter family $M(s,x) := (DT_s)_{(x,0)}$. Observing that M satisfies (H2) in Sect. 3.1 with $\sqrt{b}$ in the place of b, we obtain, by Lemma 3.1, $|\frac{\partial^2}{\partial x \partial s} e_1| < K$. From this we conclude

$$|\frac{d}{dx} e_1(T,(x,0);S(\frac{\partial}{\partial x},\mathbf{v})) \ - \ \frac{d}{dx} e_1(T_0,(x,0);S(\frac{\partial}{\partial x},\mathbf{v}))| \ = \mathcal{O}(\sqrt{b}).$$

Next we consider T and interpolate between the x-axis and γ. Write $\gamma(x) = (x, \gamma_y(x))$. For $s \in [0,b]$, let $z(s,x) = (x, \frac{s}{b}\gamma_y(x))$, and let $M(s,x) = DT_{z(s,x)}$, $S = S(u(s,x),\mathbf{v})$ where $u(s,x) = (1, \frac{s}{b}\gamma_y'(x))$. Another application of Lemma 3.1 gives

$$|\frac{d}{dx} e_1(T,\gamma(x);S(\gamma'(x),\mathbf{v})) \ - \ \frac{d}{dx} e_1(T,(x,0);S(\frac{\partial}{\partial x},\mathbf{v}))| \ = \ \mathcal{O}(b).$$

The inequality in (3) now follows from $|\frac{d}{ds} e_1| > K^{-1}$ and the fact that both $|\frac{d}{ds}\gamma'|$, which is equal to the curvature of γ, and $|\frac{d}{ds} S(\gamma',\mathbf{v})|$ are $<< 1$. □

A.10 Critical points on $C^2(b)$-curves (Sect. 3.7)

Proof of Corollary 3.1: Let $\gamma : [\hat{x} - \delta, \hat{x} + \delta] \to R_1$ be the $C^2(b)$-curve in question, with $\hat{x} \in C$ and $\gamma(x) = (x, \gamma_y(x))$. Let $\eta = \langle e_1(S), v \rangle$ be as defined in Sect. 3.6. Since $|\frac{d\eta}{dx}| > K_1^{-1}$ (Lemma 3.7), there can be at most one $x \in [\hat{x} - \delta, \hat{x} + \delta]$ with $\eta(x) = 0$. Observe that if we show $\eta(\hat{x}) = \mathcal{O}(b)$, that will force $\eta(x) = 0$ for some x with $|x - \hat{x}| < K_1|\eta(\hat{x})|$. The claim on $\eta(\hat{x})$ follows by interpolating between $(T_0, (x,0), S(\frac{\partial}{\partial x},\mathbf{v}))$ and $(T, \gamma(x), S(\gamma'(x),\mathbf{v}))$ as detailed in Appendix A.9. □

Proof of Lemma 3.8: We obtain by using Lemma 3.2 that for all $i < n$, $DT_z^i(\mathbf{v}) > \hat{K}_0^{-1}$ for all z with $|z - \gamma(0)| < 2b^{\frac{n}{5}}$. This guarantees that $e_n(S)$ with $S = S(\hat{\gamma}',\mathbf{v})$ is defined at $\hat{\gamma}(s)$ for all $s \in [-b^{\frac{n}{5}}, b^{\frac{n}{5}}]$.

Let η_n be defined by using e_n instead of e_1 in the definition of η in Sect. 3.6. We have $|\frac{d}{ds}\eta_n| = |\frac{d}{ds}\eta_1| + \mathcal{O}(b) > \frac{1}{2}K_1^{-1}$ from Lemmas 3.1 and 3.7. This shows that there is at most one point at which $\eta_n = 0$, i.e. a critical point of order n. To see there exists one such point, we first interpolate between $(\gamma(0), S(\gamma'(0),\mathbf{v}))$ and $(\hat{\gamma}(0), S(\hat{\gamma}'(0),\mathbf{v}))$. By Lemma 3.1 and assumption (b) in this lemma,

$$|e_n(\hat{\gamma}(0)) - e_n(\gamma(0))| < Kb^{\frac{n}{4}}. \tag{26}$$

We have

$$|\eta_n(0)| \le |e_n(\hat{\gamma}(0)) - e_n(\gamma(0))| + |e_n(\gamma(0)) - \gamma'(0)| + |\gamma'(0) - \hat{\gamma}'(0)| < Kb^{\frac{n}{4}}$$

because $|e_n(\gamma(0)) - \gamma'(0)| = 0$, and $|\gamma'(0) - \hat{\gamma}'(0)| < b^{\frac{n}{4}}$ from assumption (b) of this lemma. This estimate on $\eta_n(0)$ forces $\eta_n(s) = 0$ for some s with $|s| < Kb^{\frac{n}{4}}$. □

Proof of Lemma 3.9: Since e_{n+1} is defined on a neighborhood of $\gamma(0)$ of radius $>> (Kb)^n$, and $|\eta_{n+1}(0)| < (Kb)^n$ by Lemma 3.1(a), we proceed as in the proof of Lemma 3.8 to obtain a critical point of order $n+1$. This argument is then repeated to obtain successively critical points of order $n+2$, $n+3$, and so on. The distances between critical points of consecutive orders decrease geometrically. (It is not necessary to increase the order by 1 each time, but we may not be able to construct a critical point of order $n+m$ in a single step: for m large, e_{n+m} may not be defined in a neighborhood of order $(Kb)^n$.) □

A.11 Splitting algorithm (Sect. 3.8)

Proof of Lemma 3.10: Consider first I_j with the property that $I_j \not\supset I_{j'}$ for any j'. We observe that (i) for $i = t_j + 1, \cdots, t_j + \ell_{t_j} - 1$, w_i^* is b-horizontal by Lemma 3.4, and (ii) $w^*_{t_j+\ell_{t_j}}$ is b-horizontal by assumption (a) in the lemma and the the single-return argument in Sect. 3.8. We emphasize that the preceding discussion is entirely independent of what happens before time t_j, for assumption (a) guarantees that whatever happens before, $w^*_{t_j}$ splits in a desirable manner.

Consider next I_j with the property that all $I_{j'} \subset I_j$ are of the type in the last paragraph. For definiteness, we label these inner intervals as $I_{j_1}, \cdots, I_{j_k}$ with $j_1 < \cdots < j_k$. Then applying the observations in the last paragraph to each of the inner intervals and Lemma 3.4 to the times in between, we see that the only time i we need to be concerned with is $i = t_j + \ell_{t_j}$. There are two cases: $t_{j_k} + \ell_{t_k} < i$, and $= i$.

If $t_{j_k} + \ell_{t_k} < i$, the b-horizontal property of w_i^* follows from an argument identical to that of the single-return case applied to the time interval I_j; note that when making this argument, one is entirely oblivious to whether or not $\hat{w}_{t_j}$ is split and recombined between times t_j and i.

If $t_{j_k} + \ell_{t_k} = i$, we argue first that the rejoining of $DT^{t_{j_k}}(\hat{E}_{j_{t_k}})$ increases the slope of $DT_{z_{i-1}}(w^*_{i-1}))$ by at most $(Kb)^{\frac{1}{2}\ell_{t_{j_k}}}$. Then we apply the single-return argument to I_j (ignoring the splitting and re-combinations that occurred in between), and note that with the rejoining of $DT^{\ell_{t_j}}\hat{E}_{t_j}$, the slope deteriorates by an additional $(Kb)^{\frac{1}{2}\ell_{t_j}}$. Since $s(DT_{z_{i-1}}(w^*_{i-1})) < \frac{3K_0}{2\delta}b$, the resulting vector w_i^* is still b-horizontal.

Inducting on the number of layers inside an I_j, we see that the only question that remains to be treated is the following: Suppose there exist $j_1 < \cdots < j_k$ such that $j_1 + \ell_{j_1} = \cdots = j_k + \ell_{j_k} = i$. Can we be assured of the b-horizontal property of w_i^* for arbitrary k? We answer in the affirmative, on the grounds that the deterioration in slope caused by recombining $DT^{\ell_{j_t}}(\hat{E}_{j_t})$ is a geometric series of the form $\sum (Kb)^q$. To see this, one must start from the rejoining of the vector that is split off last, and work backwards one step at a time in the estimation of additional deterioration in slope. □

A.12 Estimates on $B^{(k)}$ and $\mathcal{F}_k$ (Sect. 4.2)

Sublemma A.12.1 *For $\varepsilon, a > 0$, let J be an interval containing $[0, \frac{\varepsilon}{a}]$, and let $\psi : J \to \mathbb{R}$ be a C^2 function with $|\psi''| \le a$ and $|\psi(\frac{\varepsilon}{a}) - \psi(0)| \le \frac{1}{2}\frac{\varepsilon^2}{a}$. Then $|\psi'(0)| \le \varepsilon$.*

Proof: Suppose $|\psi'(0)| = \varepsilon' > \varepsilon$. Then $|\psi(\frac{\varepsilon}{a}) - \psi(0)| \ge \varepsilon' \frac{\varepsilon}{a} - \frac{1}{2}a(\frac{\varepsilon}{a})^2 > \frac{1}{2}\frac{\varepsilon^2}{a}$. ◇

Proof of Lemma 4.2: Between $Q^{(k)}$ and $Q^{(\hat{k})}$ we have $Q^{(k)} \supset Q^{(k+1)} \supset \cdots \supset Q^{(k+n)} = Q^{(\hat{k})}$. Let $\hat{z} = z_{k+n}$, and for $0 \le i < n$, choose z_{k+i} so that

- $z_{k+i} \in Q^{(k+i)}$,
- z_{k+i} has the same x-coordinates as $\hat{z}$, and
- z_k lies on the $\mathcal{F}_k$-leaf containing z.

Let γ_{k+i} be the $\mathcal{F}_{k+i}$-leaf containing z_{k+i} and let τ_{k+i} be the tangent to γ_{k+i} at z_{k+i}. We claim that $\angle(\tau_{k+i}, \tau_{k+i+1}) \le K\delta^{-\frac{3}{2}} b^{\frac{k+i}{4}+\frac{1}{2}}$. To see this, regard γ_{k+i} and γ_{k+i+1} as graphs of functions defined on the x-axis, fix l with $1 \le l < m$ (where $m = \dim(X)$), and let $\psi(x) =$ y^l coordinate of $\gamma_{k+i+1}(x) - y^l$ coordinate of $\gamma_{k+i}(x)$. Since the diameter of $Q^{(k+i)}$ is $< b^{\frac{k+i}{2}}$ by (A1)(ii), and the γ_j are $C^2(b)$-curves, we wish to use Sublemma A.12.1 with $a = \frac{Kb}{\delta^3}$ and $\frac{1}{2}\frac{\varepsilon^2}{a} = b^{\frac{k+i}{2}}$ to conclude that $|\psi'| \le \varepsilon = K\delta^{-\frac{3}{2}} b^{\frac{k+i}{4}+\frac{1}{2}}$. To do this, we need to first verify that $\frac{\varepsilon}{a} <<$ the length of $Q^{(k+i+1)}$, i.e. $K\delta^{\frac{3}{2}} b^{\frac{k+i}{4}-\frac{1}{2}} << \min\{\delta, e^{-\lambda(k+i+1)}\}$. This is true for $k+i > 1$. The claim is also valid when $k + i = 1$, for $|\psi'| < \frac{Kb}{\delta}$ by the $C^2(b)$-property of the curves in question.

Thus we have

$$\angle(\tau, \hat{\tau}) \;\le\; \angle(\tau, \tau_k) + \sum_{i=0}^{n-1} \angle(\tau_{i+k}, \tau_{i+k+1}) < K\delta^{-3} b \cdot |z - \hat{z}|_h + b^{\frac{k}{4}},$$

the first term in the last inequality following again from the fact that γ_k is $C^2(b)$. $\square$

Proof of Lemma 4.1: It suffices for us to prove $|z_0^*(Q^{(k)}) - z_0^*(Q^{(k+1)})| < Kb^{\frac{k}{4}}$. The rest follows immediately. To prove the estimate on z_0^*, let γ and $\hat{\gamma}$ be the leaves of $\mathcal{F}_k$ and $\mathcal{F}_{k+1}$ containing $z_0^*(Q^{(k)})$ and $z_0^*(Q^{(k+1)})$ respectively, parametrized so that $\gamma(0) = z_0^*(Q^{(k)})$ and $\hat{\gamma}(0)$ has the same x-coordinate as $\gamma(0)$. We apply Lemma 3.8 to obtain a critical point $\hat{z}$ of order k on $\hat{\gamma}$: the bound for $|\gamma(0) - \hat{\gamma}(0)|$ comes from the diameter bound for $Q^{(k)}$ given by (A1); the one for $|\gamma'(0) - \hat{\gamma}'(0)|$ comes from Lemma 4.2. Lemma 3.8 tells us also that $|\hat{z} - \hat{\gamma}(0)| < Kb^{\frac{k}{4}}$. Lemma 3.9 says that $|z_0^*(Q^{(k+1)}) - \hat{z}| < (Kb)^k$. $\square$

A.13 Correct alignment implies correct splitting (Sect. 4.4)

Proof of Lemma 4.7: Assume that $\phi(z_i) = z_0^*(Q^{(j)})$, and let γ and $\hat{\gamma}$ be the $\mathcal{F}_j$-leaves parameterized by arc-length through $\phi(z_i)$ and z_i respectively. Both γ and $\hat{\gamma}$ are $C^2(b)$-curves by (A1)(ii). Let ℓ_i be the splitting period of z_i.

¿From Lemma 4.3 and Lemma 3.8, there exists a critical point $\hat{z}$ of order j on $\hat{\gamma}$ with $|\hat{z} - \phi(z_i)| < Kb^{\frac{j}{4}} << d_{\mathcal{C}}(z_i)$. From Lemma 3.2, the most contracted direction of order ℓ_i in S is well defined on $\hat{\gamma}$ from $\hat{z}$ to z_i where $S = S(\hat{\gamma}', \mathbf{v})$. By Lemma 3.7, $|\frac{d}{ds} e_1| > K_1^{-1}$. By Lemma 3.1, $|\frac{d}{ds}(e_{\ell_i} - e_1)| < Kb$. Together we have

$$\angle(e_{\ell_i}(S), \hat{\gamma}')(z_i) > K_1^{-1}|\hat{z} - z_i| - |e_j(\hat{z}) - e_{\ell_i}(\hat{z})| > \frac{3}{4} K_1^{-1} d_{\mathcal{C}}(z_i). \tag{27}$$

For the last inequality we used $|e_j(\hat{z}) - e_{\ell_i}(\hat{z})| < \min((Kb)^{\frac{1}{4}j}, (Kb)^{\ell_i}) << d_{\mathcal{C}}(z_i)$.

Next we pass from $e_{\ell_i}(S)$ to $e_{\ell_i}(S^*)$ at z_i where $S^* = S(w_i^*, \mathbf{v})$. This is a straightforward interpolation between $u = \hat{\gamma}'$ and $u = w_i^*$ using Lemma 3.1. Since the difference in u is $\le \varepsilon_0 d_{\mathcal{C}}(z_i)$ by the assumption of correct alignment, we obtain

$$|\angle(e_{\ell_i}(S), e_{\ell_i}(S^*))| < K\varepsilon_0 d_{\mathcal{C}}(z_i) \tag{28}$$

where K is the constant in Lemma 3.1. Finally, $|B_i|/|A_i|$ in Lemma 4.7 is $\approx |\angle(e_{\ell_i}(S^*), w_i^*)|$, and from (27) and (28), we have the following estimates at z_i:

$$\begin{aligned} |\angle(e_{\ell_i}(S^*), w_i^*)| \;&>\; |\angle(e_{\ell_i}(S), \hat{\gamma}')| - |\angle(e_{\ell_i}(S), e_{\ell_i}(S^*))| - |\angle(w_i^*, \hat{\gamma}')| \\ &>\; \frac{3}{4} K_1^{-1} d_{\mathcal{C}}(z_i) - K\varepsilon_0 d_{\mathcal{C}}(z_i) - \varepsilon_0 d_{\mathcal{C}}(z_i) \;>\; \frac{1}{2} K_1^{-1} d_{\mathcal{C}}(z_i). \end{aligned}$$

$\square$

A.14 Comparison of derivatives during bound periods (Sect. 5.1)

The following sublemma is used in a number of places. Its proof is easy and left to the reader:

Sublemma A.14.1 *Let $z_0 \in \Gamma_{\theta N}$ be of generation k. Then for all $i \le \theta^{-1} k$, the size of the longest splitting period z_i is in is $< K\theta\alpha i$.*

Proof of Proposition 5.1: The proof proceeds by induction. Let $i < N$ be the inductive index. We assume that (6) and (7) hold for *all* triples (z_0, ξ_0, ξ_0') in the same component of $\mathcal{C}^{(1)}$ and all $j \le \min(p(z_0; \xi_0, \xi_0'), i-1)$. We then fix a specific triple (z_0, ξ_0, ξ_0') and prove for it step i of these two assertions assuming $i \le p(z_0; \xi_0, \xi_0')$. Note that K_1, the constant in the statement of the proposition, must not be allowed to increase from step to step. It is larger than any other generic constant K that appears in the proof. In particular, K does not depend on K_1.

Let $M_i = |w_i^*(\xi_0)|$, $M_i' = |w_i^*(\xi_0')|$, and $\theta_i(\xi_0, \xi_0') = \angle(w_i^*(\xi_0), w_i^*(\xi_0'))$.

Case 1 No splitting period expires at z_i and $i-1$ is not a return time. In this case $w_i^* = DTw_{i-1}^*$. Writing $C = DT_{\xi_{i-1}}$, $C' = DT_{\xi_{i-1}'}$, $u = \frac{w_{i-1}^*(\xi_0)}{|w_{i-1}^*(\xi_0)|}$ and $u' = \frac{\hat{w}_{i-1}^*(\xi_0')}{|\hat{w}_{i-1}^*(\xi_0')|}$, we have

$$\theta_i(\xi_0, \xi_0') \approx \frac{|Cu \wedge C'u'|}{|Cu||C'u'|} \le K^{-1}\delta^{-2}\left(|Cu \wedge Cu'| + |Cu \wedge (C' - C)u'|\right).$$

By Sublemma A.4.1, $|Cu \wedge Cu'| < Kb\theta_{i-1}$. This together with $|Cu \wedge (C'-C)u'| < Kb|\xi_{i-1} - \xi_{i-1}'|$ gives $\theta_i \le \frac{Kb}{\delta^2}(|\xi_{i-1} - \xi_{i-1}'| + \theta_{i-1}) < b^{\frac{1}{2}}\Delta_{i-1}(\xi_0, \xi_0')$, proving (7) for step i.

To compare magnitudes, we have

$$\frac{M_i'}{M_i} = \frac{M_{i-1}'}{M_{i-1}} \cdot \frac{|C'u'|}{|Cu|} \le \frac{M_{i-1}'}{M_{i-1}}\left(1 + \frac{|C'u' - Cu|}{|Cu|}\right) \le \frac{M_{i-1}'}{M_{i-1}}\left(1 + \frac{\|C' - C\|}{|Cu|} + \frac{|C(u - u')|}{|Cu|}\right).$$

Since $|Cu| > K^{-1}d_{\mathcal{C}}(z_{i-1})$, $\|C - C'\| < K|\xi_{i-1} - \xi_{i-1}'|$ and $|u - u'| \approx \theta_{i-1}(\xi_0, \xi_0') < b^{\frac{1}{2}}\Delta_{i-2}(\xi_0, \xi_0')$, we have

$$\frac{M_i'}{M_i} \le \frac{M_{i-1}'}{M_{i-1}}\left(1 + K\frac{\Delta_{i-1}(\xi_0, \xi_0')}{d_{\mathcal{C}}(z_{i-1})}\right) \le \exp\left\{K_1 \sum_{n=1}^{i-1} \frac{\Delta_n(\xi_0, \xi_0')}{d_{\mathcal{C}}(z_n)}\right\}, \tag{29}$$

the last inequality following from (6) for step $i-1$ and the fact that $K < K_1$.

Case 2 $i - 1$ is a return time. In this case the angle estimate is trivial since $\theta_i(\xi_0, \xi_0') = \angle(C\mathbf{v}, C'\mathbf{v})$. To compare magnitudes, we first recall that

$$w_{i-1}^*(\xi_0) = A(\xi_{i-1}) \cdot e(\xi_{i-1}) + B(\xi_{i-1}) \cdot \mathbf{v}$$

where $e = e(\xi_{i-1}) = e_{\ell(z_{i-1})}(\xi_{i-1}, S(w_{i-1}^*(\xi_0), \mathbf{v}))$; $w_{i-1}^*(\xi_0')$ and $e' = e(\xi_{i-1}')$ are defined similarly. From Lemma 3.1, we have

$$|e - e'| \le K(|\xi_{i-1} - \xi_{i-1}'| + \theta_{i-1}(\xi_0, \xi_0')). \tag{30}$$

Let $B_0 = \frac{B(\xi_{i-1})}{|w_{i-1}^*(\xi_0)|}$. Since $w_i^*(\xi_0) = B(\xi_{i-1}) \cdot C\mathbf{v}$, we have

$$\frac{M_i'}{M_i} = \frac{M_{i-1}'}{M_{i-1}} \cdot \frac{|B_0'|}{|B_0|} \cdot \frac{|C'\mathbf{v}|}{|C\mathbf{v}|}. \tag{31}$$

To estimate $\frac{B_0'}{B_0}$, we let $u = \frac{w_{i-1}(\xi_0)}{|w_{i-1}(\xi_0)|}$, and let $e^\perp$ denote the unit vector orthogonal to e in $S(u, \mathbf{v})$. Then a straightforward computation (using the fact that $\langle \mathbf{v}, e^\perp \rangle \approx 1$) gives

$$|B_0 - B_0'| \;=\; \left|\frac{\langle u, e^\perp \rangle}{\langle \mathbf{v}, e^\perp \rangle} - \frac{\langle u', e'^\perp \rangle}{\langle \mathbf{v}, e'^\perp \rangle}\right| \;\le\; 2(|u - u'| + |e^\perp - e'^\perp|) \;\le\; 2(|u - u'| + |e - e'|). \tag{32}$$

This together with $|B_0| \sim d_{\mathcal{C}}(z_{i-1})$ gives

$$\left|\frac{B_0'}{B_0} - 1\right| < \frac{1}{|B_0|}\left(b^{\frac{1}{2}}\Delta_{i-2} + K|\xi_{i-1} - \xi_{i-1}'|\right) < \frac{K\Delta_{i-1}(\xi_0, \xi_0')}{d_{\mathcal{C}}(z_{i-1})}. \tag{33}$$

For the last ratio,

$$\left|\frac{|C'\mathbf{v}|}{|C\mathbf{v}|}-1\right| \le K|\xi_{i-1}-\xi'_{i-1}| < \frac{K\Delta_{i-1}(\xi_0,\xi'_0)}{d_{\mathcal{C}}(z_{i-1})}. \tag{34}$$

Thus (6) is proved for step i by substituting (33) and (34) into (31) and taking $K_1 > 2K$.

Case 3 At least one splitting period initiated previously expires at time i. Among the splitting periods expiring at this time, let j be the time when the first one is initiated. Then

$$w_i^*(\xi_0) = B(\xi_j)\cdot DT^{i-j}_{\xi_j} w_0 + A(\xi_j)\cdot DT^{i-j}_{\xi_j} e(\xi_j). \tag{35}$$

Let

$$B_0 = \frac{B(\xi_j)}{|w_j^*(\xi_0)|}, \quad A_0 = \frac{A(\xi_j)}{|w_j^*(\xi_0)|}, \quad V = DT^{i-j}_{\xi_j}\mathbf{v}, \quad E = DT^{i-j}_{\xi_j} e(\xi_j).$$

As before, all corresponding quantities for ξ'_0 carry a prime. We shall use $\theta_i(\xi_0,\xi'_0) \le (I)+(II)$ where

$$(I) := \left|\frac{V'}{|V'|}-\frac{V}{|V|}\right| \quad \text{and} \quad (II) := \left|\frac{A'_0E'}{B'_0|V'|}-\frac{A_0E}{B_0|V|}\right|.$$

Assume that z_j is bound to $\eta_0 \in \Gamma_{\theta N}$. We apply our inductive hypotheses to the triple (η_0,ξ_j,ξ'_j) for time $i-j$. From (7), we get $(I) < b^{\frac{1}{2}}\hat{\Delta}_{i-j-1}$ where

$$\hat{\Delta}_n = \sum_{s=1}^{n} b^{\frac{s}{4}} 2^{\hat{\ell}_{n-s}} |\xi_{j+n-s}-\xi'_{j+n-s}| \tag{36}$$

and $\hat{\ell}_{n-s}$ is the longest splitting period η_{n-s} finds itself in. Clearly we have $\ell_{j+n-s} \ge \hat{\ell}_{n-s}+n_0$ where n_0 is the minimum number of iterations between returns to $\mathcal{C}^{(1)}$. Set $n_0 = 2$ if there is no return to $\mathcal{C}^{(1)}$ between time $j+1$ to i. Then $\hat{\Delta}_n < 2^{-n_0}\Delta_{j+n}$ and $(I) < \frac{1}{2}b^{\frac{1}{2}}\Delta_{i-1}$.

For (II), we first write

$$(II) \le \frac{|A_0|}{|B_0|}\cdot\frac{|E'-E|}{|V|} + \left|\frac{A_0}{B_0|V|}-\frac{A'_0}{B'_0|V'|}\right||E'|. \tag{37}$$

From $\frac{|A_0|}{|B_0|} \sim \frac{1}{d_{\mathcal{C}}(z_j)}$, $|E'-E| \le (Kb)^{i-j}(|\xi_j-\xi'_j|+\theta_j(\xi_0,\xi'_0))$ (Lemma 3.1), and $|V|>1$, we obtain

$$\frac{|A_0|}{|B_0|}\frac{|E'-E|}{|V|} < K(Kb)^{\frac{2}{3}(i-j)}\Delta_j << b^{\frac{1}{2}}\Delta_{i-1}. \tag{38}$$

For the second term on the right side of (37), we write

$$\begin{aligned}\left|\frac{A_0}{B_0|V|}-\frac{A'_0}{B'_0|V'|}\right||E'| &\le (Kb)^{i-j}\frac{|A'_0|}{|B'_0|}\cdot\frac{1}{|V|}\cdot\left(\left|\frac{A_0}{A'_0}\cdot\frac{B'_0}{B_0}-1\right|+\left|1-\frac{|V|}{|V'|}\right|\right)\\ &\le \frac{(Kb)^{i-j}}{d_{\mathcal{C}}(z_j)}\left(\frac{|A_0|}{|A'_0|}\left|\frac{B'_0}{B_0}-1\right|+\left|\frac{A_0}{A'_0}-1\right|+\left|1-\frac{|V|}{|V'|}\right|\right).\end{aligned}$$

This is estimated term by term: For the first term,

$$\frac{(Kb)^{i-j}}{d_{\mathcal{C}}(z_j)}\cdot\frac{|A_0|}{|A'_0|}\cdot\left|\frac{B'_0}{B_0}-1\right| \le \frac{(Kb)^{i-j}}{d_{\mathcal{C}}(z_j)}\cdot\frac{\Delta_j}{d_{\mathcal{C}}(z_j)} < (Kb)^{\frac{i-j}{3}}\Delta_j << b^{\frac{1}{2}}\Delta_{i-1} \tag{39}$$

because $b^{\frac{i-j}{3}} \approx d_{\mathcal{C}}(z_j)$ by the definition of splitting period. For the second term,

$$\frac{(Kb)^{i-j}}{d_{\mathcal{C}}(z_j)}\cdot\left|\frac{A_0-A'_0}{A'_0}\right| \le \frac{(Kb)^{i-j}}{d_{\mathcal{C}}(z_j)}K\Delta_j << b^{\frac{1}{2}}\Delta_{i-1} \tag{40}$$

because $A_0' \approx 1$ and $|A_0 - A_0'| \le K(|\xi_j - \xi_j'| + \theta_j(\xi_0, \xi_0'))$. Finally, for the third term, we have

$$\frac{(Kb)^{i-j}}{d_{\mathcal{C}}(z_j)}\left|1 - \frac{|V'|}{|V|}\right| \le \frac{(Kb)^{i-j}}{d_{\mathcal{C}}(z_j)} \sum_{k=1}^{i-j-1} \frac{K_1 \hat{\Delta}_k}{d_{\mathcal{C}}(z_{k+j})} < K_1 (Kb)^{\frac{2}{3}(i-j)} \sum_{k=j}^{i-1} e^{\alpha(k-j)} \Delta_k << b^{\frac{1}{2}} \Delta_{i-1}. \tag{41}$$

Here we have used our inductive assumption (6) for $\frac{|V'|}{|V|}$. Putting (38)-(41) together, we conclude $(II) << b^{\frac{1}{2}} \Delta_{i-1}$. Hence $\theta_i(\xi_0, \xi_0') < (I) + (II) < b^{\frac{1}{2}} \Delta_{i-1}$.

To compare magnitudes, we write

$$\frac{M_i'}{M_i} = \frac{M_j'}{M_j} \cdot \frac{|B_0'V' + A_0'E'|}{|B_0 V + A_0 E|} \le \frac{M_j'}{M_j} \cdot \frac{|V'|}{|V|} \cdot \frac{|B_0'|}{|B_0|} \cdot \left(1 + \frac{\left|\frac{V'}{|V'|} - \frac{V}{|V|} + \frac{A_0'E'}{B_0'|V'|} - \frac{A_0 E}{B_0|V|}\right|}{\left|\frac{V}{|V|} + \frac{A_0 E}{B_0|V|}\right|}\right)$$

Since $\frac{|A_0|}{|B_0|} \sim \frac{1}{d_{\mathcal{C}}(z_j)}$ and $\frac{|E|}{|V|} \le d_{\mathcal{C}}^3(z_j)$ (by the definition of the splitting period at z_j) , it follows that $\frac{|A_0 E|}{|B_0 V|} << 1$, giving

$$\frac{M_i'}{M_i} \le \frac{M_j'}{M_j} \cdot \frac{|V'|}{|V|} \cdot \frac{|B_0'|}{|B_0|} \cdot \left(1 + 2\left|\frac{V'}{|V'|} - \frac{V}{|V|}\right| + 2\left|\frac{A_0'E'}{B_0'|V'|} - \frac{A_0 E}{B_0|V|}\right|\right). \tag{42}$$

We estimate the contributions from the first three ratios on the right side. Applying our inductive assumption (6) to the first ratio, we obtain an upper bound of $\exp\{K_1 \sum_{n=1}^{j-1} \frac{\Delta_n(\xi_0, \xi_0')}{d_{\mathcal{C}}(z_n)}\}$. Let z_j be bound to $\eta_0 \in \Gamma_{\theta N}$. Applying inductive assumption (6) to (η_0, ξ_j, ξ_j'), we obtain

$$\frac{|V'|}{|V|} \le 1 + 2^{-n_0+1} K_1 \sum_{k=1}^{i-j-1} \frac{\Delta_{j+k}}{d_{\mathcal{C}}(z_{k+j})} = 1 + \frac{K_1}{2^{n_0-1}} \sum_{k=j+1}^{i-1} \frac{\Delta_k}{d_{\mathcal{C}}(z_k)}. \tag{43}$$

Observe that without the factor 2^{-n_0+1} in front of K_1, there would be no room for contributions from the remaining terms (unless we allow K_1 to increase). $\frac{|B_0'|}{|B_0|}$ is estimated as in (33), giving

$$\left|\frac{B_0'}{B_0} - 1\right| < \frac{K\Delta_j}{d_{\mathcal{C}}(z_j)}. \tag{44}$$

Since K is independent of K_1, this term is easily absorbed. Finally the terms inside parentheses in (42) sum up to

$$< 1 + 2\theta_i(\xi_0, \xi_0') < 1 + 2b^{\frac{1}{2}} \Delta_{i-1}(\xi_0, \xi_0'). \tag{45}$$

Substituting (43)-(45) into (42), we complete the proof of (6) for the triple (z_0, ξ_0, ξ_0') at step i. □

Proof of Lemma 5.1: From Sublemm A.14.1, we have $\ell_{n-s} < K\alpha\theta(n-s)$, from which it follows that $\Delta_n < \sum_{s=0}^{n} b^{\frac{s}{4}} e^{-\frac{1}{2}\beta(n-s)} < 2e^{-\frac{1}{2}\beta n}$. We now choose b small enough, and follow the last part of the proof of Sublemma A.1.1 in Appendix A.1 to finish. □

A.15 Properties of w_i^* along controlled orbits (Sect. 5.3A,B)

Proof of Lemma 5.2: We may assume that ξ_i is in a splitting period, otherwise there is nothing to prove. Let $i_1 < i \le i_2$ be the longest splitting period containing i. By Sublemma A.14.1 we have $i_2 - i_1 \le K\alpha\theta i$. Let $w_{i_1} = Ae + B\mathbf{v}$ be the usual splitting. An upper bound for $|w_i^*|$ in terms of $|w_i|$ is then given by

$$|w_i^*| \le K^{i-i_1}|B| \le K^{i-i_1}|w_{i_2}^*| = K^{i-i_1}|w_{i_2}| \le K^{i-i_1}(K^{i_2-i}|w_i|) \le K^{\varepsilon i}|w_i|.$$

The first "$\leq$" uses the fact that $\frac{|w^*_{j+1}|}{|w^*_j|} \leq$ some K, the second uses $|w^*_{i_2}| > K^{-1}|B|$, and the third $\|DT\| \leq K$.

To obtain an upper bound for $|w_i|$ in terms of $|w^*_i|$, we let $j_1 < \cdots < j_n$ be the return times between i_1 and i with the property that the splitting period initiated at each j_k extends beyond i. Using the nested structure of splitting periods, and from the way w^*_i is defined, we have $|w_i| \leq K^{i-i_1}|w_{i_1}| \leq K^{i-i_1}|\xi_{i_1} - \phi(\xi_{i_1})|^{-1}|\xi_{j_1} - \phi(\xi_{j_1})|^{-1} \cdots |\xi_{j_n} - \phi(\xi_{j_n})|^{-1}|w^*_i|$. From Sublemma A.14.1 and (A2), we have $i - j_k < K\alpha\theta(i - j_{k-1}) < \cdots < (K\alpha\theta)^{k+1}i_1$ and $|\xi_{j_k} - \phi(\xi_{j_k})| > e^{-\alpha(i-j_{k-1})}$. Hence $|w_i| < K^{i-i_1}e^{\alpha(1+2K\alpha\theta)i_1}|w^*_i| < K^{\varepsilon i}e^{2\alpha i}|w^*_i|$. $\square$

Proof of Lemma 5.3: Observe first that if t is any return, and ℓ_t is its splitting period, then by Corollary 5.1, w^*_i aligns correctly at all returns in the time interval $(t, t+\ell_t)$ with $2\varepsilon_0$-error. This is because before the rejoining of the vector split off at time t, the situation is identical to that in Proposition 5.1.

To prove the lemma, we consider, in the notation of Proposition 5.2, one bound interval $[n_i, n_i + p_i)$ at a time. At time n_i, we have correct alignment by assumption. From the observation in the first paragraph, it suffices to consider returns at time $t \in (n_i, n_i + p_i)$ where ξ_t is not in any splitting period. Write $w^*_{n_i} = B\mathbf{v} + Ae_{p_i}$. Then $w_t = B \cdot DT^{t-n_i}_{\xi_{n_i}}\mathbf{v} + A \cdot DT^{t-n_i}_{\xi_{n_i}}e_{p_i}$. The first of these two vectors has length $> K^{-1}e^{-\alpha(t-n_i)}e^{\lambda(t-n_i)}$ and aligns correctly with $< 2\varepsilon_0$-error at time t by Corollary 5.1. Addition of the second, which has length $< (Kb)^{t-n_i}$, changes the angle of alignment by an insignificant amount relative to $d_{\mathcal{C}}(\xi_t) > e^{-\alpha(t-n_i)}$. Thus w_t aligns correctly with $< 3\varepsilon_0$-error. $\square$

A.16 Derivative growth along controlled orbits (Sect. 5.3C)

Proof of Lemma 5.4: 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 $\mathcal{C}^{(1)}$ 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:

$$I := \frac{|w^*_{i_{s+1}}|}{|w^*_{i_s+p_s}|}, \qquad II := \frac{|w^*_{i_{s+1}+p_s}|}{|w^*_{i_{s+1}}|} \quad \text{and} \quad III := \frac{|w^*_n|}{|w^*_j|}.$$

First we prove the lemma assuming that no splitting periods initiated before time k expires between times k and n. By Lemma 3.5, $I \geq \frac{1}{2}c_2 e^{\frac{1}{4}\lambda_0(i_{s+1}-(i_s+p_s))}$. By Proposition 5.2(ii), $II \geq K^{-1}e^{\frac{\lambda}{3}p_{s+1}}$. Moreover, K^{-1} can be easily absorbed into the exponential estimate for the bound period $[i_s, i_s + p_s]$. For III, let ℓ be the splitting period initiated at time j. If $\ell > n - j$, then $III \geq K^{-1}d_{\mathcal{C}}(\xi_j)e^{\lambda(n-j)}$. If not, we split w^*_j into $w^*_j = Ae_{n-j} + B\mathbf{v}$ where e_{n-j} is the most contracted direction of order $n-j$ at ξ_j in $S = S(\mathbf{v}, w^*_j)$. (Note that e_{n-j} is well-defined.) Then $III \geq K^{-1}d_{\mathcal{C}}(\xi_j)e^{\lambda(n-j)} - (Kb)^{n-j}$. The last term is negligible because $d_{\mathcal{C}}(\xi_j) \sim b^{\frac{\ell}{3}} >> (Kb)^{n-j}$. Altogether, this gives $\frac{|w^*_n|}{|w^*_k|} \geq K^{-1}d_{\mathcal{C}}(\xi_j)e^{\lambda'(n-k)}$ for some $\lambda' > 0$ as claimed.

To finish this proof, we view contributions from splitting periods initiated before time k as perturbations of the estimates above, and verify that they are in fact inconsequential. $\square$

Proof of Lemma 5.5: The case where ξ_k is not in a splitting period is contained in Lemma 5.4. Let j be the starting point of the largest splitting period covering ξ_k. We claim that its length ℓ is $< K\theta(n-j)$. If not, then we would have $|\xi_j - z_0| < b^{K\theta(n-j)}$ where $z_0 = \phi(\xi_j)$, so that for all $\hat{m}$ with $n - j < \hat{m} \leq n$, $|\xi_{\hat{m}} - z_{\hat{m}-j}| < \|DT\|^{\hat{m}-j}b^{K\theta(n-j)} < e^{-\beta(\hat{m}-j)}$, contradicting our assumption that ξ_n is free. Since $n - j > p(\xi_j) >> \ell(\xi_j) > k - j$, it follows that $\ell < 2K\theta(n-k)$. By Lemma 5.4, $|w_n| > K^{-1}\delta e^{\lambda'(n-j)}|w_j| \geq K^{-1}\delta e^{\lambda'(n-k)}K^{-\ell}|w_k|$. $\square$

A.17 $\|DT^{i-s}_{\xi_s}\|$ and $w_i(\xi_0)$ (Sect. 5.3C)

It suffices to show for any (fixed) unit vector $u \in \mathbb{R}^m$ that $|DT^{i-s}_{\xi_s} u| \le K e^{-\hat\lambda s}|w_i|$. Since this involves only two vectors, the problem is a 2-dimensional one.

To understand the result, recall that in 2D, we have, by simple linear algebra,

$$\|DT^i_{\xi_0}\| \;=\; \|DT^s_{\xi_0}\|\|DT^{i-s}_{\xi_s}\| \cdot \angle(e_s(DT^{-s}_{\xi_s}), e_{i-s}(DT^{i-s}_{\xi_s})). \tag{46}$$

Note that $\|DT^i_{\xi_0}\| \sim |w_i|$ and $\|DT^s_{\xi_0}\| \sim |w_s|$. This is because $|w_j| > e^{\lambda'' j}$ for $j = 1, 2, \cdots$ (Lemmas 5.2 and 5.4), so that Lemma 3.1 applies, and since w_0 makes a definite angle with $e_1 = e_1(DT)$, it makes a definite angle with $e_j = e_j(DT^j)$ for all j. Plugging these estimates into (46), we obtain

$$\|DT^{i-s}_{\xi_s}\| \sim \frac{|w_i|}{|w_s|} \cdot \angle(e_s(DT^{-s}_{\xi_s}), e_{i-s}(DT^{i-s}_{\xi_s})).$$

The key, therefore, is to understand the angle in the displayed formula above, and to compare it to $|w_s|$, which is $> e^{\lambda'' s}$. This angle is clearly more delicate during or around splitting periods.

Sublemma A.17.1 *Let t be a return time to $\mathcal{C}^{(1)}$ for ξ_0. We denote its splitting period by ℓ_t, and let $I_t := (t - 5\ell_t, t + \ell_t)$. Then modifying I_t slightly to $\tilde I_t = (t - (5 \pm \varepsilon)\ell_t,\ t + (1 \pm \varepsilon)\ell_t)$, we may assume $\{\tilde I_t\}$ has a nested structure.*

Proof: 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 the proof of Lemma 4.6, the increments in length are exponentially small and the process terminates.

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. ♢

Let us assume this nested structure and write I_t instead of $\tilde I_t$ from here on.

Sublemma A.17.2 *For $s \notin \cup I_t$, we have, for all j with $1 \le j < i - s$, $|w_{s+j}| \ge b^{\frac{j}{9}}|w_s|$.*

Proof: We fix j and let r be such that ξ_r makes the deepest return between times s and $s + j$. Let j' be the smallest integer $\ge j$ such that $\xi_{s+j'}$ is outside of all splitting periods. Then, from Lemma 5.4, it follows that

$$|w_{s+j}| \ge K^{-K\theta(j'-j)}|w_{s+j'}| \ge K^{-K\theta(j'-j)} d_{\mathcal{C}}(z_r)|w_s| \approx K^{-K\theta(j'-j)} b^{\frac{\ell_r}{3}}|w_s|. \tag{47}$$

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 $j' - j \le \ell_r$ because r is the deepest return. The rightmost quantity in (47) is therefore $> K^{-\ell_r} b^{\frac{\ell_r}{3}}|w_s| > b^{\frac{j}{9}}|w_s|$.

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 sublemma. ♢

Proof of Lemma 5.6: Consider the case $s \notin \cup I_t$, and assume for the moment that $d_{\mathcal{C}}(\xi_s) \ge \delta_0$. Then by Sublemma A.17.2 $e_{i-s}(\xi_s)$ is well defined, and since w_s is b-horizontal, we have $\angle(w_s, e_{i-s}(\xi_s)) > K^{-1}$ by Lemma 3.6. Thus $\|DT^{i-s}_{\xi_s}\|\ |w_s| \le K|DT^{i-s}(\xi_s)w_s| = K|w_i|$, which together with $|w_s| > e^{\lambda'' s}$ gives the desired estimate. For s with $s \notin \cup I_t$ and $d_{\mathcal{C}}(\xi_s) < \delta_0$, consider ξ_{s+1}. 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 ξ_0 obeys (A2)). If $i \in I_r$, then $\|DT^{i-s}_{\xi_s}\| < K^{6\ell_r} << e^{\frac{1}{2}\lambda'' i} < e^{-\frac{1}{2}\lambda'' s} e^{\lambda'' i} < e^{-\frac{1}{2}\lambda'' s}|w_i|$. If $i \notin I_r$, let $s' = r + \ell_r$. Then $s' \notin \cup I_t$. This case having been dealt with, we have

$$\|DT^{i-s}_{\xi_s}\| \le \|DT^{s'-s}_{\xi_s}\| \cdot \|DT^{i-s'}_{\xi_{s'}}\| \le K^{6\ell_r} \cdot K e^{-\hat{\lambda} s'}|w_i|.$$

□

A.18 Quadratic turn estimates (Sect. 5.3D)

Proof of Proposition 5.3: We fix $s_1 > 0$, and let $p^* = \min_{0<s\le s_1}\{p(\xi_0(s), z_0), M\}$. All time indices i considered are $\le p^*$, and all s considered are in $(0, s_1)$, with further restrictions indicated where necessary. For $s \in (0, s_1)$ and $S = S(\gamma', \mathbf{v})$, $e_{p^*} = e_{p^*}(S)$ is well defined by Proposition 5.1. Let

$$\gamma'(s) = A_0(s) e_{p^*}(s) + B_0(s)\mathbf{v}. \tag{48}$$

Then

$$\gamma_i'(s) = A_0(s) DT^i e_{p^*}(s) + B_0(s) w_i(s).$$

All splitting periods below are determined by the orbit of z_0; we use them for all the $\xi_0(s)$ in question. Writing

$$w_i(s) = w_i(0) + (w_i^*(s) - w_i^*(0)) + (E_i(s) - E_i(0))$$

where

$$E_i(s) = \sum_{k\in\Lambda_i} A_k(s) DT^{i-k} e_{\ell_k}$$

and Λ_i is the collection of $k > 0$ such that the splitting period begun at time k extends beyond time i, we arrive at the formula

$$\xi_i(s) - z_i \;=\; \int_0^s \gamma'(u)du \;=\; w_i(0)\int_0^s B_0(u)du \;+\; I \;+\; II \;+\; III \tag{49}$$

where

$$I = \int_0^s A_0(u) DT^i(u) e_{p^*}(u) du, \quad II = \int_0^s B_0(u)(w_i^*(u) - w_i^*(0))du,$$

$$III = \int_0^s B_0(u)(E_i(u) - E_i(0))du.$$

Plan of proof: We will prove that for i and s satisfying $i \in [\ell(s), p^*]$, the first term on the right side of (49) dominates, so that assuming s_1 is sufficiently small,

$$\xi_i(s) - z_i \;\approx\; w_i(0)\int_0^s B_0(u)du \;\approx\; \frac{1}{2}B_0'(0)s^2\; w_i(0).$$

The following estimate then completes the proof: Differentiating (48), we obtain $\gamma'' = A_0' e_{p^*} + A_0 \frac{d}{ds} e_{p^*} + B_0'\mathbf{v}$. On the left side, $|\gamma''| = \mathcal{O}(b)$ since γ is $C^2(b)$. On the right side, $|A_0 \frac{d}{ds} e_{p^*}| \approx |\frac{d}{ds} e_1| > K^{-1}$ (Lemma 3.7) and $\langle e_{p^*}, \frac{d}{ds} e_{p^*}\rangle = 0$. It follows therefore that $B_0'\mathbf{v} \approx \frac{d}{ds} e_1$.

We divide the main argument of the proof into the following two steps:

Step I. Estimates on $|I|, |II|$ and $|III|$

We begin with $|I|$. We have $A_0(s) \approx 1$, so that $|I| \le (Kb)^i s << s^2$ provided $b^i < b^{\ell(s)} := s^2$. This is where the lower bound on i is used for each s.

By assumption, z_0 is a critical point of order M. If $p^* = M$, then $B(0) = 0$. For $p^* < M$, $B(0)$ may not be zero but we have $|B(0)| < (Kb)^{p^*}$. Since this error is negligible, we will write $B(0) = 0$ in the computation that follows. For $|II|$, then, we have

$$\begin{aligned} |II| &\leq K|w_i^*(0)| \int_0^s u \left| \frac{|w_i^*(u)|}{|w_i^*(0)|} - 1 \right| du \leq K|w_i^*(0)| \int_0^s u \left(\sum_{j<i} \frac{\Delta_j}{d_{\mathcal{C}}(z_j)} \right) du \\ &\leq K|w_i^*(0)| \int_0^s u \left(\sum_{j<i} (2^{K\alpha\theta j} d_{\mathcal{C}}(z_j))^{-1} \right) \sup_{j<i} |z_j - \xi_j(u)| du \\ &\leq K e^{2\alpha i} |w_i^*(0)| \int_0^s u \sup_{j<i} |z_j - \xi_j(u)| du. \end{aligned}$$

Here we have used Proposition 5.1 for the second inequality, and assumption (2) in Sect. 5.3D and Sublemma A.14.1 for the next two.

To estimate $|III|$, we have, for each $k \in \Lambda_i$,

$$\begin{aligned} & |A_k DT_{\xi_k}^{i-k} e(\xi_k) - A_k(0) DT_{z_k}^{i-k} e(z_k)| \\ \leq\ & (Kb)^{i-k} |A_k - A_k(0)| + |A_k(0)|\, |DT_{\xi_k}^{i-k} e(\xi_k) - DT_{z_k}^{i-k} e(z_k)|. \end{aligned}$$

We claim that the first term can be estimated by

$$|A_k(u) - A_k(0)| < K|w_k^*(0)| e^{4\alpha k} \sup_{j<k} |z_j - \xi_j| < K|w_i^*(0)| e^{5\alpha i} \sup_{j<i} |z_j - \xi_j|.$$

For the first inequality, we use

$$\left| \frac{A_k(u)}{A_k(0)} - 1 \right| \approx \left| \frac{|w_k^*(s)|}{|w_k^*(0)|} (1 + \mathcal{O}(|z_k - \xi_k| + \theta_k(\xi_0(u), z_0))) - 1 \right|,$$

and $|A_k(0)| \leq K|w_k^*(0)| e^{\alpha k}$ because $w_k^*(0)$ aligns correctly at time k. For the second inequality, we use $|w_k^*(0)| \leq K e^{\alpha k} |w_i^*(0)|$ by virtue of Lemma 5.4 and assumption (2) in Sect. 5.3D. Summing over all $k \in \Lambda_i$ is not problematic because of the factor $(Kb)^{i-k}$ in front. For the second term we use

$$|DT_{\xi_k}^{i-k} e(\xi_k) - DT_{z_k}^{i-k} e(z_k)| \leq (Kb)^{i-k} (|\xi_k - z_k| + \theta_k(\xi_0(u), z_0)).$$

This inequality is derived from Lemma 3.1. Altogether, we have proved

$$|II|, |III| < K|w_i^*(0)| e^{5\alpha i} \int_0^s u \sup_{j<i} |z_j - \xi_j(u)| du. \tag{50}$$

Step II. Proof of formula for $|\xi_i(s) - z_i|$

Fix i_0 so that $e^{9\alpha i_0} e^{-\beta i_0} << 1$. We define

$$U_i := K e^{5\alpha i} \sup_{j \leq i} |w_j^*(0)|$$

where K is the constant in the bound for $|II|$ and $|III|$ above. By choosing δ sufficiently small, we may assume $U_{i_0} s^2 \leq U_{i_0} \delta^2 << 1$. We now prove inductively (and in tandem) the following two statements:

(i) $K e^{2\alpha i} U_j s^2 << 1$,

(ii) $|\xi_j(s) - z_j| \approx \frac{1}{2}|\frac{d}{ds}e_1(0)||w_j(0)|s^2$, or, equivalently, $|II|, |III| << |w_i(0)|s^2$.

The first n_0 steps, where the entire action takes place away from $\mathcal{C}^{(1)}$, are trivial. We assume now that (i) and (ii) have been proved for all $j < i$, and prove (i) for step i. Using Lemmas 5.2 and 5.4, one has

$$\sup_{j \leq i} |w_j^*(0)| < e^{\alpha i}|w_i^*(0)| \leq e^{2\alpha i}|w_i(0)| \leq Ke^{2\alpha i}|w_{i-1}(0)|.$$

This combined with (ii) for step $i-1$ gives

$$U_i s^2 \leq Ke^{5\alpha i}\ (Ke^{2\alpha i}|w_{i-1}(0)|)s^2 \approx K^2 e^{7\alpha i}|\frac{d}{ds}e_1(0)|^{-1}|\xi_{i-1}(s) - z_{i-1}|.$$

Thus

$$Ke^{2\alpha i}U_i s^2 \leq K^2 e^{9\alpha i}|\frac{d}{ds}e_1(0)|^{-1}e^{-\beta(i-1)} << 1, \tag{51}$$

proving (i). To prove (ii), first observe that from Lemmas 5.2 and 5.4 we have

$$|w_j(0)| < e^{2\alpha i}|w_i(0)| \tag{52}$$

for all $j < i$. We claim that

$$\begin{aligned} |II| + |III| &\leq 2U_i \int_0^s u \sup_{j<i}|z_j - \xi_j(u)|\ du \ \leq\ KU_i \sup_{j<i}|w_j(0)|s^4 \\ &\leq (Ke^{2\alpha i}U_i s^2)|w_i(0)|s^2\ <<\ |w_i(0)|s^2. \end{aligned}$$

The first inequality above is the conclusion of Step I, the second is obtained by using (ii) for $j < i$, the third is by (52), and the last is step i of (i). This completes the proof of Step II.

Step III. Monotonicity of $s \mapsto p(s)$:

To prove that the distance formula (Step II(ii)) holds for all $i \in [\ell(s), p(s_1)]$ and $0 < s < s_1$, it remains to show that $p(s)$ is monotone in s so p^* introduced at the beginning of this proof is equal to $\min\{p(\xi_0(s_1), z_0), M\}$. To do this, we check that for s and i with $i > \ell(s)$,

$$\left|\frac{d}{ds}I\right|,\ \left|\frac{d}{ds}II\right|,\ \left|\frac{d}{ds}III\right|\ <<\ |B_0(s)||w_i(0)|.$$

Thus $\frac{d}{ds}(\xi_i(s) - z_i) \approx B_0(s)w_i(0) \approx B_0'(0)s\ w_i(0)$, i.e. $|\xi_i(s) - z_i|$ increases monotonically with s. It follows by definition that $p(s)$ is monotone in s. □

A.19 Sectional diameter of $Q^{(k)}$ (Sect. 6.3)

Proof of Lemma 6.2: Let $\xi_k \in Q^{(k)}$ be fixed. We argue as before that $d_{\mathcal{C}}(\xi_i) > 2b^{\frac{i}{5}}$ for $i = 1, \cdots, k-1$. Let S be a 2D subspace through ξ_1 containing ξ_1 and $\xi_1 + \tau_1$. All constructions are in S until the very end of the proof. There is clearly a stable curve γ_1 of order one in S passing through ξ_1. Since $d_{\mathcal{C}}(\xi_1) > 2b^{\frac{1}{5}}$, γ_1 makes an angle $\gtrsim b^{\frac{1}{5}}$ with the x-axis by Lemma 3.7; thus it connects the two components of $\partial(R_1 \cap S)$. We wish to borrow the argument in Appendix A.7 to construct inductively stable curves γ_i of order i, $i = 2, 3, \cdots, k$, through ξ_1, but are prevented from doing so due to the following technical problem: with $\kappa = b^{\frac{1}{5}}$, Lemma 3.2 (a general perturbative result) does not apply. We seek instead to use Lemma 5.4, which relies on the control of (ξ_1, τ_1) for k iterates, to estimate the growth of τ_i. Details of the argument are as follows:

Assume that $\gamma_i = \gamma_i(S)$ with the following properties has been constructed: (i) γ_i is tangent to e_i, passes through ξ_1 and connects the two components of $\partial R_1 \cap S$; and (ii) for $\xi \in \gamma_i$, $d_{\mathcal{C}}(T^j\xi) > \frac{3}{2}b^{\frac{j}{5}}$ for $1 \le j \le i$.

To construct γ_{i+1} we let U_i be the $b^{\frac{i+1}{4}}$-neighborhood of γ_i. Then the following hold for all $\xi \in U_i$: First, $d_{\mathcal{C}}(T^j\xi) > b^{\frac{j}{5}}$, so if τ_1 is tangent to $\mathcal{F}_1$ at ξ, then (ξ, τ_1) is provisionally controlled by Γ_k for i iterates.

Claim: $|\tau_j| > (Kb)^{\frac{j}{5}}$ for $j \le i$.

Proof: If $T^j\xi$ is out of all splitting periods, then $|\tau_j| > (Kb)^{\frac{j}{5}}$ by Lemma 5.4. If not, let $j_1 < j$ be the time at which the longest splitting period extending beyond j is initiated. Since $d_{\mathcal{C}}(T^{j_1}\xi) > b^{\frac{j_1}{5}}$, it follows that l_{j_1}, the splitting period initiated at j_1, is $< \frac{3j_1}{5}$. Thus $|\tau_j| > (\|DT\|)^{-\frac{3}{5}j_1}|\tau_{j_1+\ell_{j_1}}| > (\|DT\|)^{-\frac{3}{5}j_1} b^{\frac{j}{5}}$; the second inequality is obtained by applying Lemma 5.4 to $\tau_{j_1+l_{j_1}}$. $\Diamond$

By Lemma 3.1, $e_{i+1}(\xi)$ is well-defined, $|DT^j_\xi(e_{i+1})| < (Kb^{\frac{3}{5}})^j$ for all $j \le i+1$, $|e_{i+1} - e_i| < (Kb^{\frac{3}{5}})^i$, and $|\frac{d}{ds}(e_{i+1} - e_i)| < (Kb^{\frac{2}{5}})^i$. Let γ_{i+1} be the integral curve of e_{i+1} through ξ_1. We verify following the computation in Appendix A.7 that $|\gamma_{i+1} - \gamma_i| < Kb^{\frac{3}{5}i}$, so γ_{i+1} stays in U_i until it meets $\partial R_1 \cap S$. Properties (i) and (ii) are again valid for γ_{i+1}.

To finish, we let $W = T^{k-1}W_1$ where $W_1 = \cup_S \gamma_k(S)$, the union being taken over all 2D planes S containing ξ_1 and $\xi_1 + \tau_1$. $\square$

A.20 Geometry of monotone branches (Sect. 7.3)

The proof of Lemma 7.1 uses material in Sects. 7.3, 8.1 and 8.2.

Proof of Lemma 7.1: Let $T \in \mathcal{G}_N$. For $k \le \theta N$, let $\hat{R}_{1,k} = \{\xi_1 \in R_1 : \xi_k \in \cup_{M \in \mathcal{T}_k} M^\circ\}$. Then for $\xi_1 \in \hat{R}_{1,k}$ and time indices $\le k$, bound periods $p(\xi_i)$ for $\xi_i \in \mathcal{C}^{(1)}$ are well defined and $\{p(\xi_i)\}$ has a nested structure, i.e., $i + p(\xi_i) \ge j + p(\xi_j)$ for $\xi_i, \xi_j \in \mathcal{C}^{(1)}$ satisfying $i < j < i + p(\xi_i)$. We introduce a function $\mathfrak{b}_k(\xi_1)$ on $\hat{R}_{1,k}$ as follows:

- if ξ_k is free, then $\mathfrak{b}_k(\xi_1) = 0$;
- if ξ_k is bound to some point and the bound period lasts beyond time θN, then having no knowledge of events beyond time θN, we set $\mathfrak{b}_k(\xi_1) = \infty$;
- if $(j, j + p(\xi_j))$ is the longest bound period ξ_k finds itself in, and $j + p(\xi_j) \le \theta N$, then we set $\mathfrak{b}_k(\xi_1) = j + p(\xi_j) - k$.

That is to say, $\mathfrak{b}_k(\xi_1)$ gives the number of iterates it takes for ξ_k to become free – without knowledge of events after time θN. We observe immediately that due to the nested structure of $\{p(\xi_i)\}$, if $\mathfrak{b}_{k-1}(\xi_1) = i$, $0 < i \le \infty$, then $\mathfrak{b}_k(\xi_1) = i - 1$.

Let l be an arbitrary $\mathcal{F}_1$-leaf parametrized by s. Then $\mathfrak{b}_k$ is defined on $l_k := l \cap \hat{R}_{1,k}$, and the T^{k-1}-images of the connected components of l_k are exactly the maximal $\mathcal{F}_k$-segments in M° for $M \in \mathcal{T}_k$. We say $\mathfrak{b}_k$ restricted to $\omega = l(s_1, s_2) \subset l_k$ is a U-shaped function if there exists $s^* \in (s_1, s_2)$ such that $\mathfrak{b}_k$ is non-creasing on $l(s_1, s^*]$ and nondecreasing on $l[s^*, s_2)$. Lemma 7.1 is reduced to the following. We claim that on all connected components of l_k, $\mathfrak{b}_k$ is a U-shaped function, and leave the proof to the reader as an exercise. $\square$

Proof of Corollary 7.1: Corollary 7.1 follows immediately from the arguments above. The numbers K_1 and K_2 are determined from f_0 as follows: Let $\hat{x}_1 < \hat{x}_2 < \cdots < \hat{x}_r = \hat{x}_0$ be the critical points of f_0, and let $I_i = (\hat{x}_{i-1}, \hat{x}_i)$. Then

$$K_1 = \max_{1 \le i \le r} N_i \qquad \text{and} \qquad K_2 = \max_{1 \le i \le r} \sum_{1 \le j \le r} L_{ij}$$

where N_i is the number of I_j-intervals counted with multiplicity $f_0(I_i)$ intersects at least partially, and L_{ij} is the cardinality of $f_0(I_j) \cap \{\hat{x}_i\}$. □

A.21 Branch replacement (Sect. 8.3)

Proof of Lemma 8.2: Let $\frac{1}{3}H$ be the middle third of H. We will show $T^{-i}M^\circ$ is inside $\frac{1}{3}H$ so $T^{-i}M \subset H$. To prove $T^{-i}M^\circ \subset \frac{1}{3}H$, it suffices to show that if we start from $B^{(k-i)} \subset H$ and move right along any $\mathcal{F}_{k-i}$-segment γ, we will get out of $T^{-i}M^\circ$ before we reach the end of $\frac{1}{3}H$. Suppose, to derive a contradiction, that this is not true for some γ. Then every point in γ, which we may assume runs from $B^{(k-i)}$ to the right end of $\frac{1}{3}H$, is controlled for the next i iterates. We will show if this is the case then there exists $j < K\alpha(k-i)$ such that $T^j\gamma$ crosses some $Q^{(1)}$. It follows that $T^j\gamma$ crosses some $B^{(k-i+j)}$.

Let γ_0 be a segment of γ such that $\pi_x(\gamma_0) = I_{\mu j}(\hat{x})$ for some $I_{\mu j}(\hat{x})$ with $\mu \sim 2\alpha(k-i)$ (See Sect. 2.2 for a formal definition of $I_{\mu j}$), and let $\gamma_i = T^i\gamma_0$. We follow the argument in Sect. 9.2 to conclude that γ_i obeys the rules (P1') and (P2') in Sect. 9.2C (leaving details as an exercise to the reader). That $T^j\gamma$ crosses $Q^{(1)}$ for some $j < K\alpha(k-i)$ is then proved by repeating the proof of Lemma 2.4 using (P1') and (P2'). □

Proof of Lemma 8.3: Let n_1 be the smallest ℓ for which $T^{-\ell}S_\ell \subset H$.

First, we observe that if P_1 is well defined and E_1' remains active for at least n_1 generations, then all the offsprings of P_1 survive, i.e. they are not discontinued, for at least n_1 generations. This is because the new ends created as the offsprings of P_1 reproduce are younger than the end originating from $\hat{B}^{(k-i)}$, and so will last longer than it. It follows that $P_j, j \le n_1$, are well defined, and $T^{n_1-1}P_1$ is a union of branches in $\mathcal{T}_{k-i+n_1}$ with adjacent ones overlapping in critical blobs.

We prove next that P_{n_1} is subordinate to S_{n_1}.

Observe that since n_1 is the smallest ℓ with the property that $T^{-\ell}S_\ell \subset H$, it follows that S_{n_1-1} contains a $B^{(k-i+n_1-1)}$ and S_{n_1} is the image of the subset of S_{n_1-1} between $T^{n_1-1}\hat{B}^{(k-i)}$ and this $B^{(k-i+n_1-1)}$. We claim that $T^{n_1-2}P_1$ meets the $Q^{(k-i+n_1-1)}$ containing $B^{(k-i+n_1-1)}$, so that $T^{n_1-2}P_1 \cap B^{(k-i+n_1-1)}$ contains a $B^{(p+n_1-1)}$. To prove this, let $z \in B^{(k-i+n_1-1)}$. By Lemma 6.2, we know there exists a stable manifold $W = W^s_{k-i+n_1-1}$ whose T^{k-i+n_1-1}-image contains z and along which T contracts at a rate $\sim b^{\frac{1}{2}}$. Since $T^{k-i-1}W$ meets every fiber in H and has diameter $< b^{\frac{k-i-1}{2}}$, the desired result follows from the relation between $T^{-n_1}S_{n_1}$, H and P.

We explain why $T^{n_1-1}P_1$ is a monotone branch: From the observation in the first paragraph of this proof, we see that it suffices to show there are no critical blobs between $T^{n_1}\hat{B}^{(p)}$ and $TB^{(p+n_1-1)}$: If there was one, look at when and where it was created, and argue that the corresponding iterate of $T^{-n_1}S_{n_1}$ also crosses some $Q^{(j)}$ containing it, leading to the existence of a monotone branch of generation $k-i+n_1$ contained properly in S_{n_1}, which is absurd.

For $n_1 < j \le i$, the arguments are as above, namely that if S_j subdivides, then so does P_j in corresponding locations; and that no other subdivisions of P_j are possible. □

A.22 Dynamics on unstable manifolds (Sect. 9.2)

Proof of Lemma 9.1: As before, it suffices to prove correct alignment at free returns. Inducting on k, we let ξ_k be a free return, and let τ^* denote the tangent to $\mathcal{F}_j$ at ξ_k where $\phi(\xi_k)$ is of generation j. We need to show $\angle(\tau_k, \tau^*) < \varepsilon_1 d_{\mathcal{C}}(\xi_k)$, and our plan is to deduce that from the control of foliations proved earlier.

Let $n \ge k$ be a sufficiently large number to be determined. We let $\xi_k = \zeta_n$, so that τ_k is a multiple of $DT^n_{\zeta_0}\tau$, $\tau \in X_{\zeta_0}$ being a unit vector tangent to $T^{k-n}l_0$. Let $\hat{\tau}_1 \in X_{\zeta_0}$ be a unit vector tangent to $\mathcal{F}_1$. By the bound on $|\det(DT)|$, we have

$$\angle(DT^n_{\zeta_0}\tau, DT^n_{\zeta_0}\hat{\tau}_1) \ \le (Kb)^n \ \frac{1}{|DT^n_{\zeta_0}\tau|} \ \frac{1}{|DT^n_{\zeta_0}\hat{\tau}_1|} \ .$$

Observe that $|DT^n_{\zeta_0}\tau| > K^{-1}e^{\lambda'' n}$: for the first $n-k$ iterates, Lemma 3.5 applies since ζ_i is essentially outside of $\mathcal{C}^{(1)}$; for the next k iterates, use Lemma 5.4 and the fact that ξ_k is a free return. The constraints on n are as follows: First, ζ_n must be in a monotone branch in $\mathcal{T}_{n+1}$, so that $(\zeta_0, \hat{\tau}_1)$ is controlled through this time, giving $|DT^n_{\zeta_0}\hat{\tau}_1| > K^{-1}e^{\lambda'' n}$. This is not a problem since $\xi_0 \in \Omega$. Second, we assume $n \ge j$, so that our choice of $\phi(\zeta_n)$ in the control of foliations is compatible with the definition of $\phi(\xi_k)$. (A6) then guarantees $\angle(DT^n_{\zeta_0}\hat{\tau}_1, \tau^*) < \varepsilon_1 d_{\mathcal{C}}(\zeta_n)$, and the desired conclusion follows if n is large enough that $\angle(DT^n_{\zeta_0}\tau, DT^n_{\zeta_0}\hat{\tau}_1)$ is negligible. □

Proof of Proposition 9.2: (P1') is an easy exercise. (P2')(iii) and (P3') require the following extensions of Proposition 5.1.

Sublemma A.22.1 *The setting is as in Proposition 5.3. Let $\xi_0, \xi_0' \in \gamma$ be such that $|\xi_0 - \xi_0'| < \frac{1}{10}d_{\mathcal{C}}(\xi_0)$, and let $\ell = \ell(\xi_0)$, $p = p(\xi_0)$. Then*

(a) for $\ell < i \le p$,

$$\frac{|\xi_i - \xi_i'|}{|\xi_i - z_i|} < K^{6\alpha i+1}\frac{|\xi_0 - \xi_0'|}{d_{\mathcal{C}}(\xi_0)};$$

(b) with $w_i = DT^i\mathbf{v}$, we have

$$\frac{|w_p(\xi_0)|}{|w_p(\xi_0')|} < \exp\left\{K\frac{|\xi_0 - \xi_0'|}{d_{\mathcal{C}}(\xi_0)}\right\}, \quad \angle(w_p(\xi_0), w_p(\xi_0')) \le Kb^{\frac{1}{2}}\frac{|\xi_0 - \xi_0'|}{d_{\mathcal{C}}(\xi_0)}.$$

Proof: (a) We remark that this is a rough *a priori* bound in which factors of $K^{\alpha i}$ are allowed to accumulate. Let $s \mapsto \xi_0(s)$ be the parametrization of the segment from ξ_0 to ξ_0'. We write $S = S(\tau, \mathbf{v})$, $e_p = e_p(S)$, and decompose τ into $\tau = Ae_p + B\mathbf{v}$. For $\ell < i \le p$, since $|DT^i(e_p)|$ is negligible, we have

$$|DT^i_{\xi_0}\tau| \approx |B||w_i(\xi_0)| \quad \text{where} \quad |B| \approx |\frac{d}{ds}e_1| \ d_{\mathcal{C}}(\xi_0). \tag{53}$$

The combined use of Proposition 5.1, Lemma 5.1 and Lemma 5.2 gives, on the other hand,

$$\frac{|w_i(\xi_0(s_1))|}{|w_i(\xi_0(s_2))|}, \ \frac{|w_i(\xi_0)|}{|w_i(z_0)|} \ \le \ K^{3\alpha i} \tag{54}$$

where s_1, s_2 are any parameters and z_0 is the guiding critical point. Clearly, $|B(s_1)|/|B(s_2)| < K$. We have thus shown that

$$\frac{|DT^i_{\xi_0(s_1)}\tau|}{|DT^i_{\xi_0(s_2)}\tau|} \approx \frac{|B(s_1)|}{|B(s_2)|}\frac{|w_i(\xi_0(s_1))|}{|w_i(\xi_0(s_2))|} < K^{1+3\alpha i}. \tag{55}$$

Using "$\sim$" to denote omitted factors of $K^{3\alpha i}$ so the main terms show up more clearly, we then have for $\ell < i \le p$:

(i) $|\xi_i - \xi_i'| \ \lesssim \ |DT^i_{\xi_0}\tau| \cdot |\xi_0 - \xi_0'|$;

(ii) $|\xi_i - z_i| \ \sim \ (|w_i(z_0)| \ d_{\mathcal{C}}(\xi_0)) \cdot d_{\mathcal{C}}(\xi_0)$.

(i) comes from $|\xi_i - \xi_i'| \le \int |DT^i_{\xi_0(s)}\tau(\xi_0(s))|ds$ together with (55); (ii) is (A5)(iii). The assertions in this sublemma are immediate upon comparing (i) and (ii), substituting in (53), and using the comparison of $|w_i(\xi_0)|$ and $|w_i(z_0)|$ in (54).

(b) Proposition 5.1 can be written as $|w_p(\xi_0)|/|w_p(\xi_0')| < \exp\{\sum_{i=1}^{p-1} K\mathcal{D}_i\}$ where

$$\mathcal{D}_i = 2^{K\alpha\theta i}|\xi_i - \xi_i'| \left(\frac{1}{d_{\mathcal{C}}(z_i)} + \frac{b^{\frac{1}{4}}}{d_{\mathcal{C}}(z_{i+1})} + \cdots + \frac{b^{\frac{p-i}{4}}}{d_{\mathcal{C}}(z_p)} \right) < 2e^{2\alpha i}|\xi_i - \xi_i'|.$$

The upper bound for $|\xi_i - \xi_i'|$ in (a) is used in the estimates below.

Case 1: $i \geq \ell$. Using $|\xi_i - z_i| < e^{-\beta i}$, we obtain $\mathcal{D}_i < Ke^{-(\beta - K\alpha)i} \cdot \frac{|\xi_0 - \xi_0'|}{d_{\mathcal{C}}(\xi_0)}$.

Case 2: $i < \ell$. Using $|\xi_i - z_i| \leq \|DT\|^i \ d_{\mathcal{C}}(\xi_0)$ and $b^{3\ell} < d_{\mathcal{C}}(\xi_0)$, we obtain

$$\frac{d_{\mathcal{C}}(\xi_0)}{|\xi_0 - \xi_0'|} \sum_{i<\ell} \mathcal{D}_i < \sum_{i<\ell} K^i d_{\mathcal{C}}(\xi_0) < K^\ell d_{\mathcal{C}}(\xi_0) < (d_{\mathcal{C}}(\xi_0))^{1-K\theta}.$$

The angle estimate is similar. ◊

Remark: For $i = p$, the inequality sign in (i) in the proof of (a) becomes $\sim$ because $T^p\gamma$ is roughly horizontal. The same argument then gives

$$\frac{|\xi_p - \xi_p'|}{|\xi_p - z_p|} > K^{-6\alpha p - 1} \frac{|\xi_0 - \xi_0'|}{d_{\mathcal{C}}(\xi_0)} \ . \tag{56}$$

Sublemma A.22.2 *Letting τ and τ' be unit tangent vectors to γ at ξ_0 and ξ_0' respectively, we have*

$$\frac{|DT^p_{\xi_0}\tau|}{|DT^p_{\xi_0'}\tau'|} < \exp\left\{ K \frac{|\xi_0 - \xi_0'|}{d_{\mathcal{C}}(\xi_0)} \right\}.$$

Proof: Splitting $\tau = Ae_p + B\mathbf{v}$ where e_p is the most contracted direction of DT^p in $S(\tau, \mathbf{v})$, and letting $V = DT^p_{\xi_0}\mathbf{v}$, $V' = DT^p_{\xi_0'}\mathbf{v}$, $E = DT^p_{\xi_0}e_p$, and $E' = DT^p_{\xi_0'}e_p'$, we obtain

$$\frac{|DT^p_{\xi_0'}\tau'|}{|DT^p_{\xi_0}\tau|} < \frac{|V'|}{|V|} \cdot \frac{|B'|}{|B|} \cdot (1 + 2(I) + 2(II)). \tag{57}$$

where

$$(I) = \left| \frac{V}{|V|} - \frac{V'}{|V'|} \right|, \quad (II) = \left| \frac{A'E'}{B'|V'|} - \frac{AE}{B|V|} \right|.$$

To obtain (57), we have used $|AE| << |BV|$ and $|A'E'| << |B'V'|$. For $\frac{|V'|}{|V|}$, see Sublemma A.22.1(b). Since γ is $C^2(b)$, we have $|\tau - \tau'| \leq \frac{Kb}{\delta^3}|\xi_0 - \xi_0'|$. Lemma 3.1 then gives $|e_p - e_p'| \leq K|\xi_0 - \xi_0'|$. The remaining estimates resemble those in the proof of Proposition 5.1 in Appendix A.14. As in (32), we have

$$|B - B'|, \ \ |A - A'| < K(|\xi_0 - \xi_0'| + |e_p - e_p'| + |\tau' - \tau|) < K|\xi_0 - \xi_0'|. \tag{58}$$

This gives $\frac{|B'|}{|B|} \leq 1 + \frac{|B' - B|}{|B|} < 1 + \frac{K|\xi_0 - \xi_0'|}{d_{\mathcal{C}}(\xi_0)}$. (I) is the angle part of Sublemma A.22.1(b). As in case 3 in the proof of Proposition 5.1, (II) is bounded by the sum of a collection of terms of the form

$$(i) = K\frac{|E - E'|}{d_{\mathcal{C}}(\xi_0)} \quad (ii) = \frac{|E|}{d_{\mathcal{C}}(\xi_0)} \left| \frac{B'}{B} - 1 \right| \quad (iii) = \frac{|E|}{d_{\mathcal{C}}(\xi_0)} \left| \frac{A'}{A} - 1 \right| \quad (iv) = \frac{|E|}{d_{\mathcal{C}}(\xi_0)} \left| \frac{|V'|}{|V|} - 1 \right|.$$

For (i), Lemma 3.1 gives $|E - E'| < (Kb)^p|\xi_0 - \xi_0'|$. Observing that $|E| << d_{\mathcal{C}}(\xi_0)$, we estimate (ii) using the bound on $\frac{|B'|}{|B|}$ above, (iii) is similar, and (iv) is given by Sublemma A.22.1(b). ◊

Proof of (P2'): Extending ω as a $C^2(b)$ curve to $B^{(j)}(\hat{z}_0)$ if necessary, we obtain (P2')(i) from (A5)(i). For (ii), the desired bound follows from (A5)(ii) and (54) above. For (iii), (56) gives

$$|T^p(\omega)| \geq K^{-1}e^{-6\alpha p}|\xi_p - \hat{z}_0| \cdot \frac{|I_{\mu j}|}{e^{-\mu}} > K^{-1}\frac{1}{\mu^2}e^{-(\beta+6\alpha)p} > \frac{1}{\mu^2}e^{-K_1\alpha\mu}.$$

Proof of (P3'): Follow the proof of (P3) in Appendix A.1 and use Sublemma A.22.2. □

A.23 Bounds for $\frac{d}{da}z_0$ (Sects. 10.1 and 10.2)

Proof of Lemma 10.1: (a) To bound the first derivatives of u, it suffices to estimate $\partial_a\psi$ and $\partial_{ax}\psi$; bounds for $\partial_x\psi$ and $\partial_{xx}\psi$ are known since $x \mapsto \gamma(x,a)$ is $C^2(b)$. To pass between γ and $l = T_a^{-k}\gamma$, we use the notation

$$t(x,a) := \pi_x(T_a^{-k}\gamma(x,a)) \quad \text{and} \quad (X(t,a), Y(t,a)) := T_a^k(t, y_0),$$

assuming $l \subset \{y = y_0\}$. Differentiating $\psi(x,a) = Y(t(x,a),a)$, we obtain

$$\partial_a\psi = \partial_t Y(t,a)\partial_a t(x,a) + \partial_a Y(t,a). \tag{59}$$

Here and in the rest of the proof, we use the fact that all first and second partial derivatives of Y are bounded above by $K^k b$, and corresponding partials of X are bounded by K^k. Partials in t, however, are potentially problematic and must be treated with care. To bound $\partial_a t(x,a)$, we write it as

$$\partial_a t(x,a) = -\frac{\partial_a X(t,a)}{\partial_t X(t,a)}. \tag{60}$$

Since $T_a^k|_l$ is controlled, $|\partial_t X(t,a)| > 1$, and so this term is $< K^k$. Thus $|\partial_a\psi| < K^k$.

To estimate $\partial_{ax}\psi$, we take one more derivative to obtain

$$\partial_{ax}\psi = \partial_{tt}Y\partial_x t\partial_a t + \partial_t Y\partial_{ax}t + \partial_{at}Y\partial_x t.$$

Since $t = t(x,a)$ is implicitly defined by $x = X(t,a)$, we have $|\partial_x t(x,a)| = |\partial_t X(t,a)|^{-1} < 1$, and finally

$$|\partial_{xa}t(x,a)| = \frac{1}{|\partial_t X(t,a)|^2}|\partial_{at}X\partial_x t\partial_t X - \partial_{tt}X\partial_x t\partial_a X| < K^{k+1}.$$

This completes the proof of $|\partial_{ax}\psi| < K^{k+1}$.

To bound the second derivatives of u, we need to bound the third derivatives of ψ. These are estimated similarly and are left to the reader. Since $v = (\mathbf{v} - \langle u, \mathbf{v}\rangle u)/|\mathbf{v} - \langle u, \mathbf{v}\rangle u|$, bounds for $\|v\|_{C^2}$ follow from those of u. This completes the proof of (a).

For (b) we cannot appeal simply to Lemma 3.1 because the bound on the C^2-norms of u and v in Lemma 10.1 is not a single number depending on the family T_a; it increases with the generation of the critical point. We go directly instead to the formulas for the most contracted directions in Appendix A.4II. Since η_{k+1} is the quantity β in Appendix A.4II with $S = S(u,v)$ and $M = DT_a^{k+1}(\gamma(x,a))$, we have $\eta_{k+1} = \langle Mu, Mv\rangle$. (b) follows now from (a) and the C^2-norm of M. □

Proof of Lemma 10.3: From Proposition 10.1, $z_0^{(k)}(a)$ is well-defined on J_n with $n = k\theta^{-1}$. Let $k' < k$ be the largest integer such that $Q^{(k')}(a) \supset Q^{(k)}(a)$, and let $z_0^{(k')}(a) = z_0^*(Q^{(k')}(a))$. Then $(1+2\theta)^{-1}k \leq k'$ (see (A1')) and $|z_0^{(k')}(a) - z_0^{(k)}(a)| < Kb^{\frac{k'}{4}}$ (Lemma 4.1).

The following calculus estimate will be used: Let g be a real valued C^2-function defined on an interval of length L, and assume that $|g| \leq M_0$ and $|g''| < M_2$. If $4M_0 < L^2$, then $|g'| \leq \sqrt{M_0}(1 + M_2)$.

To apply this estimate, we write $z_0^{(k)}(a) = (x_0^{(k)}(a), y_0^{(k)}(a))$, and let $g(a) = x_0^{(k)}(a) - x_0^{(k')}(a)$. Then g is defined on J_n, so $L = 2\rho^n = 2\rho^{k\theta^{-1}}$. Here $M_0 = Kb^{\frac{k'}{4}}$, and $M_2 = K^k$ from Corollary 10.1. Assuming $b^\theta < \rho^5$, $4M_0 < L^2$ holds. Therefore

$$\left|\frac{d}{da}(x_0^{(k)}(a) - x_0^{(k')}(a))\right| < b^{\frac{k'}{8}} K^k < b^{\frac{k'}{9}}.$$

A similar estimate holds for $\frac{d}{da}y_0^{(k)}$. □

A.24 Equivalence of τ- and a-derivatives (Sect. 11.1)

Proof of Lemma 11.1: In this proof we fix i_0 and let $(a,b) \to (a^*, 0)$. Recall that if $\tau_0 = (\tau_{0,x}, \tau_{0,y})$, then by Corollary 10.2, $\tau_{0,y} \to 0$ as $b \to 0$. The two terms of V are estimated as follows:

(i) Writing $T^{i_0}_{a^*,0} = (T^1, 0)$, we have, as $b \to 0$,

$$(DT_a^{i_0})_{z_0}\tau_0 \;\to\; \left(\frac{\partial T^1}{\partial x}(x_0,0)\tau_{0,x} + \frac{\partial T^1}{\partial y}(x_0,0)\tau_{0,y},\ 0\right) = (0,0).$$

(ii) Assume z_s stays out of $\mathcal{C}^{(1)}$ for $> i_0$ iterates. Then as $(a,b) \to (a^*, 0)$,

$$\begin{aligned}\frac{\sum_{s=1}^{i_0} DT_{z_s}^{i_0-s}\psi(z_{s-1})}{|w_{i_0}|/|w_1|} \quad &\to \quad \left(\frac{\sum_{s=1}^{i_0}(f^{i_0-s})'(x_s(a^*))\frac{d}{da}(f_a(x_{s-1}))(a^*)}{\pm(f^{i_0-1})'(x_1(a^*))},\ 0\right) \\ &= \quad \left(\pm\sum_{s=1}^{i_0}\frac{\frac{d}{da}(f_a(x_{s-1}))(a^*)}{(f^{s-1})'(x_1(a^*))},\ 0\right).\end{aligned}$$

□

Proof of Lemma 11.2:

$$\begin{aligned}|\angle(w_i,\tau_i)| \;&\approx\; \frac{|w_i\wedge\tau_i|}{|w_i||\tau_i|} \;\le\; \frac{1}{|\tau_i|}\left(\sum_{s=1}^{i}\frac{1}{|w_i|}|w_i\wedge DT_{z_s}^{i-s}\psi(z_{s-1})| + \frac{|w_i\wedge DT_{z_0}^i\tau_0|}{|w_i|}\right) \\ &\le\; \frac{1}{|\tau_i|}\left(\sum_{s=1}^{i}\frac{|w_s|}{|w_i|}\left|\frac{w_s}{|w_s|}\wedge\psi(z_{s-1})\right| b^{i-s} + \frac{|\tau_0|}{|w_i|}\ b^i\right) \;\le\; \frac{K}{|\tau_i|}\sum_{s=0}^{\infty} b^s.\end{aligned}$$

The last inequality is valid if $|w_s| \le |w_i|$ for all $s \le i$, which is the case at free returns. □

A.25 Bound period estimates for parameters (Sect. 11.2)

We begin with some estimates on derivative comparisons. Let $a, a' \in \hat{\omega}$ be as in Lemma 11.5. We let $\xi_0 = \zeta_n(a)$, $\xi_0' = \zeta_n(a')$, $w_i(\xi_0) = (DT_a^i)_{\xi_0}\mathbf{v}$, $w_i(\xi_0') = (DT_{a'}^i)_{\xi_0'}\mathbf{v}$ and $p = p(\hat{\omega})$. We wish to compare $w_i(\xi_0)$ and $w_i(\xi_0')$ for $i \le p$.

Sublemma A.25.1 (Parameter version of Proposition 5.1) *There exists $K > 0$ such that*

$$\frac{|w_i(\xi_0)|}{|w_i(\xi_0')|},\ \frac{|w_i(\xi_0')|}{|w_i(\xi_0)|} \;\le\; K^{3\alpha i}\cdot\exp\left\{\sum_{j=0}^{i-1} Ke^{2\alpha j}|\xi_j - \xi_j'| + K^i|a-a'|\right\} \quad \text{for } i \le p; \qquad (61)$$

$$\angle(w_p(\xi_0), w_p(\xi_0')) < b^{\frac{1}{2}}\sum_{j=0}^{p-1}(Kb)^{\frac{j}{4}}|\xi_{p-j} - \xi_{p-j}'| + K^p|a-a'|. \qquad (62)$$

Remarks (i) The factor $K^{3\alpha i}$ in (61) can be dropped if ξ_i is out of all splitting periods (see the proof below). (ii) We may assume the quantity inside brackets in (61) is $<< 1$ (cf. Lemma 5.1). This is because $p < K\alpha n$ and $|a - a'| < \hat{K}e^{-\lambda' n}$ (Proposition 11.1).

Proof: Let $\eta_j = T_a^j \xi_0'$, $v_j = (DT_a^j)_{\xi_0'}\mathbf{v}$. Then $\frac{|w_i(\xi_0)|}{|w_i(\xi_0')|} = \frac{|w_i(\xi_0)|}{|v_i|}\frac{|v_i|}{|w_i(\xi_0')|}$. Since $|\xi_j - \eta_j| \leq |\xi_j - \xi_j'| + |\xi_j' - \eta_j| < |\xi_j - \xi_j'| + K^j|a - a'|$, we have, by Proposition 5.1,

$$\frac{|w_i(\xi_0)|}{|v_i|} \leq K^{3\alpha i} \exp\left\{\sum_{j=0}^{i-1} Ke^{2\alpha j}|\xi_j - \eta_j|\right\} \leq K^{3\alpha i} \exp\left\{\sum_{j=0}^{i-1} Ke^{2\alpha j}|\xi_j - \xi_j'| + K^i|a - a'|\right\}, \tag{63}$$

the $K^{3\alpha i}$ factor being there to account for the discrepancy between $w_i(\xi_0)$ and $w_i^*(\xi_0)$. Since $|w_i(\xi_0')| > K^{-1}$ and $|v_i - w_i(\xi_0')| < K^i|a - a'|$, we have

$$\frac{|v_i|}{|w_i(\xi_0')|} < 1 + \frac{|v_i - w_i(\xi_0')|}{|w_i(\xi_0')|} < 1 + K^i|a - a'|, \tag{64}$$

completing the proof of (61).

For (62), we write $\angle(w_p(\xi_0), w_p(\xi_0')) \leq \angle(w_p(\xi_0), v_p) + \angle(v_p, w_p(\xi_0'))$. By Proposition 5.1,

$$\angle(w_p(\xi_0), v_p) < b^{\frac{1}{2}} \sum_{j=0}^{p-1}(Kb)^{\frac{j}{4}}|\xi_{p-j} - \eta_{p-j}| < b^{\frac{1}{2}} \left(\sum_{j=0}^{p-1}(Kb)^{\frac{j}{4}}|\xi_{p-j} - \xi_{p-j}'| + K^p|a - a'|\right). \tag{65}$$

To estimate $\angle(v_p, w_p(\xi_0'))$, note that by (64) and Remark (ii) above, $\frac{|v_i|}{|w_i(\xi_0')|} \approx 1$ for all $i \leq p$. Proceeding inductively, we let $i \leq \frac{p}{2}$, $u = \frac{v_i}{|v_i|}$, and $\hat{u} = \frac{w_i(\xi_0')}{|w_i(\xi_0')|}$. Since $|w_{2i}(\xi_0')| > K^{-1}$, we have

$$\begin{aligned}
\angle(v_{2i}, w_{2i}(\xi_0')) &< |(DT_a^i)_{\eta_i}u \wedge (DT_{a'}^i)_{\xi_i'}\hat{u}| \frac{|v_i|}{|v_{2i}|}\frac{|w_i(\xi_0')|}{|w_{2i}(\xi_0')|} \\
&\leq \left(|(DT_a^i)_{\eta_i}u \wedge (DT_a^i)_{\eta_i}\hat{u}| + |(DT_a^i)_{\eta_i}u \wedge ((DT_a^i)_{\eta_i} - (DT_{a'}^i)_{\xi_i'})\hat{u}|\right) K^{2i} \\
&\leq (Kb)^i \angle(v_i, w_i(\xi_0')) + K^{4i}|a - a'|.
\end{aligned}$$

We conclude inductively that $\angle(v_p, w_p(\xi_0')) < K^{2p}|a - a'|$, completing the proof of (62). ◊

Sublemma A.25.2 (Parameter version of Sublemma A.22.1 in Appendix A.22) *Let* $\hat{z}_0 = \phi(\xi_0(a))$. *Then*

(a) for $\ell < i < p$ *where* ℓ *is the splitting period of* ξ_0, *we have*

$$\frac{|\xi_i - \xi_i'|}{|\xi_i - \hat{z}_i|} < K^{6\alpha i+1}\frac{|\xi_0 - \xi_0'|}{|\xi_0 - \hat{z}_0|}, \qquad \frac{|\xi_p - \xi_p'|}{|\xi_p - \hat{z}_p|} > K^{-6\alpha p-1}\frac{|\xi_0 - \xi_0'|}{|\xi_0 - \hat{z}_0|};$$

(b)

$$\frac{|w_p(\xi_0)|}{|w_p(\xi_0')|} < \exp\{K\frac{|\xi_0 - \xi_0'|}{d_{\mathcal{C}(a)}(\xi_0)} + K^p|a - a'|\}.$$

Proof: The proof follows closely that of Sublemma A.22.1 with the following modifications: In part (a), we consider the parametrization of the critical curve ζ_{n_k} from ξ_0 to ξ_0' by arclength, and split its tangent vectors τ. The correctness of this splitting is a consequence of Lemma 11.2 and the fact that w_{n_k} splits correctly. (53) is a statement about individual parameters. To prove (54), we use Sublemma A.25.1 instead of Proposition 5.1. The rest of the proof then proceeds as before. The term $K^p|a - a'|$ in (b) is from the corresponding term in (61). ◊

Proof of Lemma 11.5: Let $\tilde{a} \in \hat{\omega}$ be the parameter at which the minimum in the definition of $p(\hat{\omega})$ is attained. Then (a) is an immediate consequence of (A5)(i) for $T_{\tilde{a}}$.

Let $\hat{z}_0(a) = \phi_a(\zeta_n(a))$. (b) follows from the fact that for all $a \in \hat{\omega}$ and $j < p(\hat{\omega})$, $|\hat{z}_j(a) - \hat{z}_j(\tilde{a})| \le K^j|\hat{\omega}| \le K^{\alpha n}\hat{K}e^{-\lambda' n} << e^{-\beta j}$. In the second inequality we have used $p(\hat{\omega}) \le \alpha n$ and $|\hat{\omega}| \le \hat{K}e^{-\lambda' n}$, which follows from $|\zeta_n(\hat{\omega})| \le 1$ and Proposition 11.1.

(c) is proved via the following string of inequalities:

$$\frac{|\tau_{n+p}(a)|}{|\tau_n(a)|} > \hat{K}^{-2}\frac{|w_{n+p}(a)|}{|w_n(a)|} > K^{-1}\hat{K}^{-2}\frac{|w_{n+p}(\tilde{a})|}{|w_n(\tilde{a})|} > K^{-1}\hat{K}^{-2}e^{\frac{p}{3}\lambda}.$$

The first inequality above is based on Proposition 11.1. For the second inequality, first recall that for both of the maps T_a and $T_{\tilde{a}}$, since w_n splits correctly, we have $|w_{n+p}| \approx \frac{1}{2}\frac{de_1}{ds}d_{\mathcal{C}}(z_n)|DT^p_{z_n}(\mathbf{v})|\cdot|w_n|$. We then use Sublemma A.25.1 and the remarks following it to compare $|(DT^p_a)_{z_n(a)}(\mathbf{v})|$ and $|(DT^p_{\tilde{a}})_{z_n(\tilde{a})}(\mathbf{v})|$, and note that the other factors are comparable up to a fixed constant. The last inequality follows from Proposition 5.2(2) for $T_{\tilde{a}}$.

(d) is a simple consequence of the bound on $\angle(\tau_{n+i}, w_{n+i})$ (Lemma 11.2) and the fact that $\angle(w_{n+i}, DT^i_{z_n}(\mathbf{v})) << 1$ outside of splitting periods.

(e) is an application of the second inequality in Sublemma A.25.2(a). □

A.26 Distortion estimates for parameters (Sect. 12.2)

We prove Lemma 12.2 in this appendix. Let $\omega \in \mathcal{Q}_{n-1}$ be as in Lemma 12.2, and let $a, a' \in \omega$. Where no ambiguity arises, we will omit mention of the parameters and write $z_i = z_i(a)$, $z_i' = z_i(a')$, $w_i = w_i(z_0) = (DT^i_a)_{z_0}(\mathbf{v})$, $w_i' = w_i(z_0') = (DT^i_{a'})_{z_0'}(\mathbf{v})$, and similarly for τ_i and τ_i'.

Plan of proof: Since the formula for the evolution of τ_i is more involved, we again invoke Proposition 11.1 and prove $\frac{|w_n|}{|w_n'|} < K$. Let $0 < n_1 < n_1 + p_1 \le n_2 < n_2 + p_2 \le n_3 < \cdots < n_q + p_q \le n$ be such that n_k is a free return, p_k is the ensuing bound period, and n_{k+1} is the first return following $n_k + p_k$. We write

$$|w_n| \;=\; |w_{n_1}| \cdot \frac{|w_{n_1+p_1}|}{|w_{n_1}|} \cdot \frac{|w_{n_2}|}{|w_{n_1+p_1}|} \cdot \frac{|w_{n_2+p_2}|}{|w_{n_2}|} \cdot \frac{|w_{n_3}|}{|w_{n_2+p_2}|} \cdots \tag{66}$$

and estimate the factors in (66) separately. These factors are of two types, the more complicated of which being $\frac{|w_{n_k+p_k}|}{|w_{n_k}|}$. In the proof of Lemma 11.5(c) in Appendix A.25, we reduced the comparison of $\frac{|w_{n_k+p_k}|}{|w_{n_k}|}$ to that of $|DT^{p_k}_{z_{n_k}}(\mathbf{v})|$ with bounded error. A more refined estimate is needed here to control the cumulative effect of these errors over time intervals that may contain arbitrarily large numbers of bound periods. Such a comparison involves the difference in slopes between w_{n_k} and w'_{n_k}. Let $\theta_i := \angle(w_i, w_i')$.

Sublemma A.26.1 *(i) Let i_0 be as in Proposition 11.1. Then $\theta_{i_0} < K^{i_0}|a - a'|$.*
(ii) For all $k \ge 1$,

(a) $\theta_{n_k} < Kb^{\frac{1}{2}}|z_{n_k} - z'_{n_k}| + 2b^{\frac{1}{2}}|a - a'| + b^{\frac{1}{2}(n_k - (n_{k-1}+p_{k-1}))}\theta_{n_{k-1}+p_{k-1}}$,
with "$n_0 + p_0$" in the inequality above replaced by "i_0" in the case $k = 1$;

(b) $\theta_{n_k+p_k} \;\le\; 2b^{\frac{1}{2}}\sum_{j=0}^{p_k-1}(Kb)^{\frac{j}{4}}|z_{n_k+p_k-j} - z'_{n_k+p_k-j}| + K^{p_k}|a - a'| + b^{\frac{p_k}{4}}\theta_{n_k}$.

Proof: (i) is straightforward using $|z_0 - z_0'| < K|a - a'|$ from Corollary 10.2. (ii) follows from estimates very similar to those in the proof of Proposition 5.1 (Appendix A.14). More precisely:

(a) Let $\Theta_j = \angle((DT^j_a)_{z_{n_k+p_k}}w_{n_k+p_k}, (DT^j_{a'})_{z'_{n_k+p_k}}w'_{n_k+p_k})$. The assertion is proved inductively by showing, as in case 1 of Proposition 5.1,

$$\Theta_j \le b^{\frac{1}{2}}(\Theta_{j-1} + |a - a'| + |z_{n_k+p_k+j-1} - z'_{n_k+p_k+j-1}|). \tag{67}$$

(b) Let $p = p_k$, and write $u = \frac{w_{n_k}(\xi_0)}{|w_{n_k}(\xi_0)|}$. Let $e = e_p(S)$ where $S = S(u, \mathbf{v})$. As usual, we split u into $u = B\mathbf{v} + Ae$. Following the computation in the proof of Proposition 5.1, Case 2, we obtain

$$|B' - B|,\ |A' - A| \ < \ K(\theta_{n_k} + |z_{n_k} - z'_{n_k}| + |a - a'|). \tag{68}$$

Here $|u - u'| = \theta_{n_k}$, and by Lemma 3.1, $|e - e'| \le K(\theta_{n_k} + |a - a'| + |z_{n_k} - z'_{n_k}|)$. The rest of the proof follows Case 3 in the same proof. As usual, we write $V = (DT_a^p)_{z_n}\mathbf{v}$, $V' = (DT_{a'}^p)_{z'_n}\mathbf{v}$, $E = (DT_a^p)_{z_n}e$, $E' = (DT_{a'}^p)_{z'_n}e'$. Then we have $\theta_{n_k+p_k} \le (I) + (II)$ where $(I) = \angle(V, V')$, $(II) = \left|\frac{A'E'}{B'|V'|} - \frac{AE}{B|V|}\right|$. For (I) we use the angle part of Sublemma A.25.1. The other estimates involve the same terms as in the proof of Proposition 5.1, case 3, and are carried out similarly. $\Diamond$

Corollary A.1 *(i)* $\theta_{n_k} \le b^{\frac{1}{2}}|z_{n_k} - z'_{n_k}| + K^{K\alpha n}|a - a'|$.
(ii) Letting $u = \frac{w_{n_k}}{|w_{n_k}|}$ *and* $p = p_k$, *we have*

$$\frac{|(DT_a^p)_{z_{n_k}}u|}{|(DT_{a'}^p)_{z'_{n_k}}u'|} < \exp\{K\frac{|z_{n_k} - z'_{n_k}|}{d_{\mathcal{C}(a)}(z_{n_k})} + K^{K\alpha n}|a - a'|\}.$$

Proof: (i) follows inductively from Sublemma A.26.1(ii). We use $K\alpha n$ to dominate p_k, and assume n is sufficiently large that $b^{\frac{1}{4}(n-i_0)}K^{i_0} < K^{K\alpha n}$. For (ii), we split u as in part (ii) of Sublemma A.26.1, obtaining $\frac{|(DT_a^p)_{z_{n_k}}u|}{|(DT_{a'}^p)_{z'_{n_k}}u'|} < \frac{|B|}{|B'|}\frac{|V|}{|V'|}(1 + 2(I) + 2(II))$. From the estimates in Sublemma A.26.1(ii) and the bound on θ_{n_k} in (i) above, we see that the right side of this inequality is bounded by terms of the form as claimed. $\Diamond$

Proof of Lemma 12.2 This proof follows that of (P3) in Appendix A.1. Letting $u_i = \frac{w_i}{|w_i|}$, we write $\log\frac{|w_n|}{|w'_n|} \le K\sum_{k=1}^q (S'_k + S''_k)$ where

$$S'_k = \log\frac{|(DT_a^{p_k})_{z_{n_k}}u_{n_k}|}{|(DT_{a'}^{p_k})_{z'_{n_k}}u'_{n_k}|} \quad \text{and} \quad S''_k = \log\frac{|(DT_a^{n_{k+1}-(n_k+p_k)})_{z_{n_k+p_k}}u_{n_k+p_k}|}{|(DT_{a'}^{n_{k+1}-(n_k+p_k)})_{z'_{n_k+p_k}}u'_{n_k+p_k}|}$$

except for S''_q which ends at index $n - 1$.

To estimate S'_k, we let $\sigma_k = |z_{n_k} - z'_{n_k}|$. Then Corollary A.1(ii) gives $S'_k < K\frac{|\sigma_k|}{d_{\mathcal{C}}(z_{n_k})} + K^{K\alpha n}|a - a'|$. The sum $\sum_k K\frac{|\sigma_k|}{d_{\mathcal{C}}(z_{n_k})}$ is estimated as in Appendix A.1. The additional term representing parameter contributions sums to $< nK^{K\alpha n}|a - a'| < nK^{K\alpha n}e^{-\lambda' n}$, which is uniformly bounded in n. $\sum S''_k$, which treats iterates outside of $\mathcal{C}^{(1)}$, is easily estimated to be $< K\frac{|\sigma_q|}{\delta} + nK^{K\alpha n}e^{-\lambda' n}$. $\square$

References

[A] D.V. Anosov, *Geodesic flows on compact Riemanian manifolds of negative curvature*, Proc. Steklov Inst. Math. **90** (1967), 1-209.

[B] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Springer (Lecture Notes in Math., Vol. **470**), Berlin, 1975.

[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.

[BY] M. Benedicks and L.-S. Young, *Sinai-Bowen-Ruell measure for certain Hénon maps*, Invent. Math. **112** (1993), 541-576.

[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.

[J] M. Jakobson, *Absolutely continues invariant measures for one-parameter families of one-dimensional maps*, Comm. Math. Phys. **81** (1981), 39-88.

[LY] F. Ledrappier and L.-S. Young, *The metric entropy of diffeomorphisms*, Ann. Math., **122** (1985), 509-574.

[M] M. Misiurewicz, *Absolutely continues invariant measures for certain maps of an interval*, Publ. Math. IHES. **53** (1981), 17-51.

[MV] L. Mora and M. Viana, *Abundance of strange attractors*, Acta. Math. **171** (1993), 1-71.

[O] V. I. Oseledec, *A multiplicative ergodic theorem: Liapunov characteristic numbers for dynamical systems*, Trans. Moscow Math. Soc. **19** (1968), 197-231.

[P] Ja. B. Pesin, *Characteristic Lyapunov exponents and smooth ergodic theory*, Russ. math. Surv. **32.4** (1977), 55-114.

[R] D. Ruelle, *Ergodic theory of differentiable dynamical systems*. Publ. Math., Inst. Hautes Étud. Sci. **50** (1979), 27-58.

[Si1] Y. G. Sinai, *Gibbs measure in ergodic theory*, Russian Math. Surveys **27** (1972), 21-69.

[Si2] Y. G. Sinai, *Dynamical systems with elastic reflections: Ergodic properties of dispersing billiards*, Russ. Math. Surveys **25** (1970), 137-189.

[Sm] S. Smale, *Differentiable dynamical systems*, Bull. Amer. Math. Soc. **73** (1967), 747-817.

[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.

[W] M. Wojtkowski, *Invariant family of cones and Lyapunov exponents*, Ergodic Theory and Dynamical Systems **5** (1983), 145-161.

[WY1] Q. Wang and L.-S. Young, *Strange attractors with one direction of instability*, Commun. Math. Phys. **218** (2001), 1-97.

[WY2] Q. Wang and L.-S. Young, *From invariant curves to strange attractors*, Commun. Math. Phys. **225** (2002), 275-304.

[WY3] Q. Wang and L.-S. Young, *Strange attractors in periodically-kicked limit cycles and Hopf bifurcations*, Commun. Math. Phys. **240** (2002), 509-529.

[Y1] L.-S. Young, *Ergodic theory of differentiable dynamical systems*, Real and Complex Dynamical Systems, (B. Branner and P. Hjorth, ed.), Kluwer Acad. Press, 1995.

[Y2] L.-S. Young, *What are SRB measures, and which dynamical systems have them?*, J. Stat. Phys. **108** (2002), 733-754.