SLOWLY RECURRENT COLLET-ECKMANN MAPS ON THE RIEMANN SPHERE

MAGNUS ASPENBERG

Abstract. In this paper we study perturbations of rational Collet-Eckmann maps for which the Julia set is the whole sphere, and for which the critical set is allowed to be slowly recurrent. Generically, if each critical point is simple, we show that each such Collet-Eckmann map is a Lebesgue point of Collet-Eckmann maps in the space of rational maps of the same degree $d \geq 2$. The same result holds in each subspace, where we fix the multiplicities of the critical points.

1. Introduction

The Collet-Eckmann condition stems from J-P. Eckmann and P. Collet in the 1980s [9, 11], and was used to show abundance of chaotic behaviour for certain maps on an interval. Chaotic behaviour of a system is usually associated to the property of sensitive dependence on initial conditions, meaning that two points $x, y$ sufficiently close to each other repel each other under iteration up to some large scale. Hence it is natural that such maps possess some kind of expanding property. A map satisfying the Collet-Eckmann condition is expansive along the forward critical orbit(s), and it turned out to be sufficient for chaotic behaviour in many situations, not only the pioneering case studied by J-P. Eckmann and P. Collet. Shortly after their works, M. Jakobson proved in [18] that the set of parameters $a \in (0, 2)$ for which $f_a(x) = 1 - ax^2$ admits an invariant absolutely continuous measure (acim) has positive Lebesgue measure. A corresponding celebrated result for complex rational maps was obtained by M. Rees in [32]. These maps also exhibit chaotic behaviour. The existence of an acim describes the typical orbits of a map in a probabilistic way. It does not immediately imply chaotic behaviour, but it is often very closely related to it and with some additional properties (such as expansion, ergodicity, positive entropy etc) this is usually the case.

It was quite early realised that the Collet-Eckmann condition, or even weaker conditions, are sufficient for the existence of an (ergodic) acim, see e.g. [10], [4, 5], [19], [27], [7], [28], [6], [14], [29]. In the fundamental papers [4, 5], M. Benedicks and L. Carleson showed that the Collet-Eckmann condition is satisfied for a set of positive Lebesgue measure in the quadratic family. Despite of the fact that the Collet-Eckmann condition in general is stronger than the existence of an acim, the two conditions are metrically the same in the real quadratic family. This was a deep result by M. Lyubich and A. Avila and C. G. Moreira, see [3], [26]. Conjecturally it holds more generally. In contrast to the chaotic, non-regular (sometimes called stochastic) parameters stands the regular parameters, for which the map has an attracting orbit. These maps were proven to be open and dense in the real case (the famous real Fatou conjecture), [25], [16], [20]. The complex Fatou conjecture is still open.
In the complex rational setting, not as much is known. A similar result to the papers \[4, 5\] was obtained by the author \[1\], improving the result by M. Rees \[32\]. Apart from implying the existence of an ergodic acim, the Collet-Eckmann condition induces more nice properties, see e.g. \[31\], \[14\]. It has geometric implications, and there are several papers studying perturbations of Collet-Eckmann maps (or similar expanding maps); see \[30\], \[13\], \[4\], \[5\], \[33\], \[12\] for real maps on an interval and families of Hénon maps, and \[32\], \[1\], \[2\], \[17\], \[15\] in the complex setting.

The result in this paper is related to \[33\] (see also \[12\]) in the complex setting. We study perturbations of complex rational Collet-Eckmann maps which have their Julia set equal to the whole sphere, and where the starting map is allowed to be critically slowly recurrent (see \[33\] and \[24\]).

Let \(\text{Crit}\) be the set of critical points for \(f\) and let \(J(f)\) and \(F(f)\) be the Julia set and Fatou set of \(f\) respectively. Let \(\text{Crit}'\) be the set of critical points \(c\) such that there are no other critical points in the forward orbit of \(c\). Derivatives are always in the spherical metric unless otherwise stated. We write \(Df(z) = f'(z)\) as the space derivative throughout the paper.

**Definition 1.1.** Let \(f\) be a non-hyperbolic rational map without parabolic periodic points. Then \(f\) satisfies the Collet-Eckmann condition (CE), if there exist constants \(C > 0\) and \(\gamma > 0\) such that, for each critical point \(c \in \text{Crit}' \cap J(f)\), we have

\[
|Df^n(fc)| \geq Ce^{\gamma n}, \quad \text{for all } n \geq 0.
\]

Let us define the upper and lower Lyapunov exponents for the critical point \(c\) respectively as

\[
\gamma(c) = \liminf_{n \to \infty} \frac{\log |Df^n(fc)|}{n}, \quad \text{and} \quad \gamma(c) = \limsup_{n \to \infty} \frac{\log |Df^n(fc)|}{n}.
\]

Then the CE-condition can be reformulated as the condition that the lower Lyapunov exponent is strictly positive for all critical points \(c \in \text{Crit}' \cap J(f)\). We write \(\gamma = \min \gamma(c)\) where the minimum is taken over all critical points \(c \in \text{Crit}' \cap J(f)\). In this paper we are going to study perturbations of rational CE maps for which the Julia set is the whole sphere, but we expect that the techniques can be used in other situations as well. A critical point \(c \in J(f)\) is slowly recurrent, cf. \[24\], if for each \(\alpha > 0\) there is some \(C > 0\) such that

\[
(1.1) \quad \text{dist}(f^n(J\text{rit}), J\text{rit}) \geq Ce^{-\alpha n}, \quad \text{for all } n \geq 0.
\]

We say that \(f\) is critically slowly recurrent if all critical points in the Julia set are slowly recurrent. Collet-Eckmann maps possess a (unique) conformal measure \(\nu\) supported on the Julia set and a unique ergodic invariant measure \(\mu\), which is absolutely continuous w.r.t. \(\nu\) (e.g. \[29\], \[31\], \[13\]). If the map \(f\) satisfies \(J(f) = \hat{C}\), then \(\nu\) is the standard Lebesgue measure and hence for such maps there exists an invariant absolutely continuous measure w.r.t. Lebesgue measure. We say that the critical points are typical with respect to this measure if the Birkhoff means converges for all critical points \(c \in J\text{rit}\), i.e.,

\[
\frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(c)) \to \int \varphi \, d\mu, \quad \text{as } n \to \infty,
\]

for \(\varphi \in L^1(\mu)\). Setting \(\varphi(z) = \log |Df(z)|\), which belongs to \(L^1(\mu)\) by \[31\], we see that if the critical points are typical, then \(\gamma = \overline{\gamma}\). It follows that the map is slowly
recurrent. The condition $\gamma = \gamma$ implies that $f$ is slowly recurrent but it is not clear if the converse holds. Conjecturally almost all CE-maps have the slow recurrence property. At least it is true in the real quadratic family (see [3]).

The space of rational maps of degree $d$ is denoted by $\mathcal{R}_d$. We will explain later what is meant by a non-degenerate real analytic family in Section 4.

**Theorem A.** Let $f$ be a critically slowly recurrent rational Collet-Eckmann map in $\mathcal{R}_d$, of degree $d \geq 2$, such that the Julia set is the whole sphere and let $f_\alpha, \alpha \in (-\epsilon, \epsilon)$ be a non-degenerate real analytic family of maps around $f = f_0$ for some $\epsilon > 0$. Then $f_0$ is a Lebesgue density point of Collet-Eckmann maps in $(-\epsilon, \epsilon)$.

We use a normalisation of the space of rational maps following Levin [22, 21]. We say that two maps $f$ and $g$ are equivalent if they are conjugate by a Möbius transformation. Then we can consider the space $\Lambda_{d,p'} \subset \mathcal{R}_d$, (see [22]) up to equivalence, as the set of rational maps $f$ of degree $d \geq 2$ with precisely $p'$ critical points, i.e. $\text{Crit} = \{c_1, \ldots, c_{p'}\}$, with corresponding multiplicities $\overrightarrow{p'} = \{m_1, \ldots, m_{p'}\}$ (in the same order). Inside such a set, we may state the following theorem as a direct consequence of Theorem A, by Fubini’s Theorem, since the set of non-degenerate directions has measure zero (this also follows from the results of G. Levin, see Section 4).

**Theorem B.** Let $f$ be a critically slowly recurrent rational Collet-Eckmann map in $\Lambda_{d,p'} \subset \mathcal{R}_d$ of degree $d \geq 2$ such that the Julia set is the whole sphere. Then $f$ is a Lebesgue density point of Collet-Eckmann maps in $\Lambda_{d,p'}$.

Generically, all critical points are simple and then $f$ is a Lebesgue density point of CE-maps in $\mathcal{R}_d$ in Theorem B. It is likely that this is true even if the critical points are not simple, using a suitable reparameterisation of the parameter space so that all critical points move analytically if higher order critical points split. However, we will not consider that case in this paper.

The proof of the main theorem is mainly based on a combination of strong results on transversality by G. Levin and developed classical Benedicks-Carleson parameter exclusion techniques. In particular, it is a generalisation of [1], which was the (revised) thesis of the author. Apart from proving Theorem A, the aim of this paper is partially to make the arguments in the Benedicks-Carleson parameter exclusion techniques more transparent.

**Remark 1.2.** It will be clear from the proof that the slow recurrence condition in Theorems A and B is a little superfluous; one only needs to have slow recurrence (1.1) for some sufficiently small $\alpha > 0$, depending on $f = f_0$. The CE-maps constructed in [1] have this property close to the starting (Misiurewicz-Thurston) map. It follows that the set of maps satisfying this weaker assumption has positive Lebesgue measure.
Acknowledgements. The author wants to thank G. Levin for many fruitful conversations on transversality. The author thanks M. Bylund and W. Cui for many interesting discussions, remarks and corrections, and finally expresses his gratitude to the Department of Mathematics at Lund University.

2. Some definitions

Let \( f = f_0 \) be a slowly recurrent Collet-Eckmann map, and \( f_a, a \in (-\varepsilon, \varepsilon) \) a real analytic family around \( f_0 \). We assume that the family is non-degenerate, which in particular means that every critical point \( c_l(a) \) of \( f_a \) moves analytically with the parameter \( a \). Another condition on the family is given in Section 4. We put \( f_a(z) = f(z, a) \) and \( Df_a(z) = f'_a(z) \). Let \( v_l(a) = f_a(c_l(a)) \) be the critical value, and suppose that \( v_l = v_l(0) \) does not contain any critical points in its forward orbit under \( f_0 \), for all \( l \). Put

\[
\xi_n(l)(a) = f^n_a(c_l(a)).
\]

We will study the evolution of \( \xi_n(l)(\omega) \) for a small interval \( \omega = (-\varepsilon, \varepsilon) \) around the starting map \( f_0 \). In the beginning this curve will grow rapidly from the expansive properties of the starting map, but later on we have to delete parameters that come too close to the set of critical points, denoted by \( \text{Crit}_a \), of \( f_a \). Now, \( \text{Crit}_a \) moves analytically, but it turns out that \( \xi_n(l)(\omega) \) and \( \text{Crit}_\omega \) are very different in diameter, due to the expansion of \( \xi_n(l)(\omega) \); it will be much bigger than \( \text{diam}(\text{Crit}_\omega) \). Let \( U \) be a neighbourhood of the critical points for the unperturbed map. Choose \( \varepsilon > 0 \) so that \( U \) is a neighbourhood around \( \text{Crit}_a \), for all \( a \in (-\varepsilon, \varepsilon) \). Moreover, if we let \( U_l \) be a component of \( U \) which contains the critical point \( c_l \) then we impose the condition \( \text{dist}(c_l(\omega), \partial U_l) \geq \text{diam}(c_l(\omega)) \) for all \( l \). To make \( U \) more precise, we choose \( \delta = e^{-\Delta} > 0 \) so that \( U = \bigcup B(c_l, \delta) \). Hence \( \varepsilon \) depends on \( \delta \).

The approach rate at which the distance \( \text{dist}(\xi_n(l)(a), \text{Crit}_a) \) may go to zero is controlled by the so called basic approach rate assumption which is inherited from the slow recurrent condition.

**Definition 2.1.** Let \( \alpha > 0 \). We say that the critical point \( c_l(a) \), (or parameter \( a \) with critical point \( l \)) satisfies the basic assumption up to time \( n \) with exponent \( \alpha \), if

\[
\text{dist}(\xi_n(a), \text{Crit}_a) \geq K_b e^{-2\alpha k}, \quad \text{for all } k \leq n,
\]

where \( K_b > 0 \) is the same constant which appears in the slow recurrent condition.

Obviously the starting map \( f_0 \) satisfies the basic assumption for all times for any \( \alpha > 0 \). From now on, fix \( \alpha > 0 \) to be much smaller than \( \min(\gamma_0, \gamma_H)(1-\tau)/(400K\Gamma) \), where \( \gamma_0 = \gamma = \gamma \) is the Lyapunov exponent appearing in the Collet-Eckmann condition for \( f_0 \), \( \Gamma = \sup_{x \in (-\varepsilon, \varepsilon)} z \in \mathbb{C} \log |Df_a(z)| \), \( K \) is the maximal degree of the critical points, \( \gamma_H \) is the exponent from Lemma 3.1, and where \( 0 < \tau < 1 \) is fixed (this is used in Section 8). We assume for the starting map \( f_0 \), that there is some constant \( C_0 > 0 \) such that,

\[
|Df^n(v_l(0))| \geq C_0 e^{\gamma_0 n}, \quad \text{for all } n \geq 0.
\]

We will construct a set of parameters around \( a = 0 \) which also satisfies both this basic assumption for this specific \( \alpha \) and the Collet-Eckmann condition for possibly slightly smaller Lyapunov exponents \( \gamma \). Since we fix \( \alpha > 0 \) we only speak of the basic assumption, without mentioning the exponent in the future.

We will make an induction argument based on the fact that we have some “basic” Lyapunov exponent \( \gamma_B > 0 \). This is typically smaller than the original Lyapunov
exponent $\gamma_0$ for $f_0$. During the induction arguments we also have to allow the Lyapunov exponent to decrease down to some certain value, a fraction of $\gamma_B$, a so called "intermediate exponent" $\gamma_I < \gamma_B$, which is required for most lemmas to work. We will also define the number $\gamma_B$ later, but it is slightly smaller than the minimum of $\gamma_I$ in Lemma 3.1 and $\gamma_0$ from the starting function $f_0$.

We write $A \sim_\kappa B$, where $\kappa \geq 1$ if
\[
\frac{1}{\kappa} A \leq B \leq \kappa B.
\]

We write $A \sim B$ to say $A \sim_\kappa B$ for some constant $\kappa \geq 1$. In several inequalities we use $C$ several times for possibly different constants, when it is clear that these constants do not depend on the dynamics, i.e. the number of iterations.

2.1. Bound and free periods. In this section we define some fundamental concepts which will be used throughout the paper. Many of them are direct analogues of corresponding definitions in [4, 5], see also [1]. We speak of a return of the sequence $\xi_{n,l}(\omega)$ into $U$ or $U'$, when we mean that $\xi_{n,l}(\omega) \cap U \neq \emptyset$ or $\xi_{n,l}(\omega) \cap U' \neq \emptyset$ respectively. We also speak of returns into $U$ or $U'$ of the sequence $\xi_{n,l}(a)$ for a single parameter $a$, and this means simply that $\xi_{n,l}(a) \in U$ or $\xi_{n,l}(a) \in U'$ respectively. Returns into the annular neighbourhoods $U' \setminus U$, i.e. when $\xi_{n,l}(a) \cap U' \neq \emptyset$ but $\xi_{n,l}(\omega) \cap U = \emptyset$, are called pseudo-returns. Sometimes we drop the index $l$ and write only $\xi_n(a) = \xi_{n,l}(a)$ for some critical point $a_1(a)$. We will also consider so called deep returns, which are returns into a smaller neighbourhood $U^2 = \cup_l B(c_l, \delta^2) \subset U$ of the critical points. These deep returns will be used only in the end of the paper, in Section 8.

The point is that when a return occurs, so that for example $\xi_n(a) \in U$, then the orbit follows the original orbit, i.e. $\xi_{n+1}(a)$ stays close to $\xi_j(a)$ for the first $j$. This is the so called bound period, which can be defined both for points $\xi_n(a)$ and curves $\xi_n(\omega)$ (precise definitions below). After the bound period ends, the free period starts until the next return, and so on. During the bound period, due to the expansion of the derivative of the original early orbit, we can show expansion of the derivative also during the bound period (with a certain loss due to the actual return, which is close to the critical set). Because of this, we will not consider returns during the bound period (bound returns), but only consider returns after the free period (free returns). When we speak of a return, we mean a free return unless otherwise stated. This is very similar to earlier constructions in [1] and [4, 5]. During the free period we will show a uniform expansion of the derivative. The result is the same as in the old traditions but the techniques stem from quite different sources in this new situation of a more general CE-map. The number $\beta > 0$ below is related to $\alpha$ in the basic approach rate condition. There is lots of freedom, but let us set $\beta = \alpha$, so that we can use the same exponent.

Definition 2.2 (Pointwise bound period). Let $\beta > 0$. Let $\xi_{n,l}(a) \in U_k' \subset U'$ be a return. Then we define the bound period for this return as the indices $j > 0$ for which the inequality
\[
|\xi_{n+j,l}(a) - \xi_{j,k}(a)| \leq e^{-\beta j} \text{dist}(\xi_{j,k}(a), \text{Crit}_{\omega}(a))
\]
holds. The largest number $p > 0$ for which the inequality holds is called the length of the bound period.
To define the bound period for an interval, we consider a return $\xi_{n,l}(\omega)$ into $U$. If
\begin{equation}
\text{diam}(\xi_{n,l}(\omega)) \geq \frac{1}{2} \frac{\text{dist}(\xi_{n,l}(\omega), \text{Crit}_\omega)}{\left(\log(\text{dist}(\xi_{n,l}(\omega), \text{Crit}_\omega))\right)^{2}},
\end{equation}
then we say that the return is essential. Otherwise it is inessential. With $\tilde{r} = -\log(\text{dist}(\xi_{n,l}(\omega), \text{Crit}_\omega))$, then the return is essential if $\text{diam}(\xi_{n,l}(\omega)) \geq (1/2)e^{-\tilde{r}/\tilde{r}^{2}}$, a bit more convenient notation. Actually we will partition the parameter intervals (explained later) so that they become so called partition elements, defined as follows.

**Definition 2.3.** For a given $S > 0$, we call parameter intervals $\omega$ satisfying the inequality
\[
\text{diam}(\xi_{k,l}(\omega)) \leq \begin{cases} \frac{\text{dist}(\xi_{k,l}(\omega), \text{Crit}_\omega)}{(\log(\text{dist}(\xi_{k,l}(\omega), \text{Crit}_\omega)))^2}, & \text{if } \xi_{k,l}(\omega) \cap U \neq \emptyset, \\ S, & \text{if } \xi_{k,l}(\omega) \cap U = \emptyset, \end{cases}
\]
for all $k \leq n$, partition elements at time $n$.

We do not speak of essential or inessential returns for pseudo-returns.

**Definition 2.4 (Bound period for an interval, essential returns or pseudo-returns).** Let $\xi_{n,l}(\omega) \cap U_k \neq \emptyset$, $(U_k' \subset U')$ be an essential return or a pseudo-return. Then we define the bound period for this return as the indices $j > 0$ for which the inequality (recall $f_a(z) = f(z, a)$),
\[
\text{dist}(f^j(z, a), \xi_{j,k}(b)) \leq e^{-\beta j} \text{dist}(\xi_{j,k}(b), \text{Crit}_k),
\]
holds for all $a, b \in \omega$, and all $z \in \xi_{n,l}(\omega)$.

If the return $\xi_{n,l}(\omega)$ into $U_k$ is inessential we will consider a host-curve as follows. Draw a straight line segment $L'$ through the end points of $\xi_{n,l}(\omega)$ with length equal to $e^{-r}/r^2$ where
\[
r = [-\log(\text{dist}(\xi_{n,l}(\omega), \text{Crit}_\omega)) - 1/2].
\]
To make it well defined, let us say that the line segment $L'$ shall be symmetric with respect to the end points of $\omega$. Let $L$ be the part of $L'$ with the central part between the end points deleted. The host curve for this return is then $L \cup \xi_{n,l}(\omega)$.

**Definition 2.5 (Bound period for an interval, inessential returns).** Let $\xi_{n,l}(\omega) \cap U_k \neq \emptyset$ be an inessential return. Then we define the bound period for this return as the indices $j > 0$ for which the inequality
\[
\text{dist}(f^j(z, a), \xi_{j,k}(b)) \leq e^{-\beta j} \text{dist}(\xi_{j,k}(b), \text{Crit}_k),
\]
holds for all $a, b \in \omega$, and all $z \in L \cup \xi_{n,l}(\omega)$.

It will be clear later that the dependence on the parameter in these definitions is inessential.

3. Expansion during the free period.

During the free period we want to show that the derivative of $f^n(z)$ grows exponentially as long as $f^j(z)$ stays outside $U$ for $j = 0, \ldots, n - 1$. In earlier papers, this was settled via the orbifold metric for postcritically finite (rational) maps, given that the postcritical set consists of at least 3 points. Here we have to use different techniques to build a uniform expansion using the second Collet-Eckmann condition discussed in [13]. In Proposition 1 of that paper, it is stated that the second Collet-Eckmann condition is satisfied for all critical points of maximal multiplicity.
However, with the slow recurrence condition, this statement holds for every critical point in the Julia set [8].

Without going through the whole construction, we refer to [13] and [14] for the details. The main idea is based on three types of iterated preimages of shrinking neighbourhoods of a given point \( z \), which in our case is a critical point \( c \) in the Julia set (actually we assume that \( J(f) = \mathbb{C} \)). This critical point is assumed not to have any critical points in its backward orbit. The type 2 and type 3 orbit have a uniform expansion (see p. 83 in [13]).

The type 1 preimages connects two critical points in the backward orbit in a way that one has a ball \( B(c, r) \) and considers preimages \( U_k \) which are sequences of components of \( f^{-k}(B(c, r_k)) \) of shrinking neighbourhoods \( B(c, r_k) \), where \( r_k \leq r \) is decreasing and \( \lim_{k \to \infty} r_k \geq r/2 \). For a type 1 orbit one has a critical point \( c_1 \in \partial U_n \) for some \( n \), and no critical points in \( U_k \) for \( 0 < k < n \). The length of this type 1 orbit is \( n \). Due to the difference in multiplicity of the critical points, type 1 orbits do not ensure immediate uniform expansion. This is resolved by looking at preimages of the type \( \ldots 111113 \), i.e. a sequence of 1s followed by a type 3 orbit. Such iterated preimages have uniform expansion (see p. 83 in [13]).

What can happen is that the induction starts (from the right) with a sequence of 1s only. Then it may happen that we do not have the desired expansion. Looking at such a block of 1s in the beginning of the sequence, we see from the calculations on p. 83 [13] that at a preimage \( y = f^{-k}(c) \) can estimate the growth of the derivative as follows. Let \( \mu_{\text{max}} \) be the maximal multiplicity of the critical points and \( \mu \) the multiplicity of \( c \). The number \( d = 0 \) below because that is the distance from the centre of the ball \( B(c, r) \) to the critical point \( c \). For some \( Q > 1 \) we then have, verbatim,

\[
|Df^k(y)|^{\mu_{\text{max}}} \geq Q^k \frac{r^{\mu_{\text{max}} - 1}}{(r + d)^{\mu - 1}} = Q^k r^{\mu_{\text{max}} - \mu}.
\]

So if \( \mu < \mu_{\text{max}} \) then this expansion is not uniform. By assumption there is a critical point \( c_1 \) on the boundary of the shrinking neighbourhood of \( f^{-k}(B(c, r)) \), for some \( k \), i.e. \( c_1 \in \partial U \), where \( U = \text{comp}(f^{-k}(B(c, r_k))) \) where \( r/2 \leq r_k \leq r \). However, by the slow recurrence condition, we have \( \text{dist}(f^k(c_1), c) \geq e^{-\alpha k} \) for some small \( \alpha > 0 \). This means that \( e^{-k\alpha} \leq r \). Since \( \alpha > 0 \) can be chosen as small as we like, (3.1) becomes, for some \( Q_1 > 1 \) possibly slightly smaller than \( Q \),

\[
|Df^k(y)|^{\mu_{\text{max}}} \geq Q^k e^{-\alpha k} = Q_1^k.
\]

**Lemma 3.1.** There exist a neighbourhood \( U' \) of the critical points such that the following holds. Let \( U \subset U' \). There exist \( \lambda > 1 \) and \( C' > 0 \), where \( C' \) depends on \( \delta' \) but not on \( \delta \) such that, if \( F^k(z) \notin U \) for \( k = 0, \ldots, n - 1 \), then

\[
|Df^k(z)| \geq C e^{-K \Delta \lambda^n}.
\]

For each \( 0 < q \leq 1 \) there exists a neighbourhood of the critical points \( \hat{U} \subset U' \) such that for any neighbourhood of the critical points \( U_1 \subset U \subset \hat{U} \) satisfying \( \text{diam}(U_{1,j}) \geq q \text{diam}(U_j) \), where \( U_{1,j} \subset U_j \) are components of \( U_1 \) and \( U \) respectively, we have the following. If \( z \notin U_1 \), \( F^k(z) \notin U \) for \( k = 1, \ldots, n - 1 \), and \( f^n(z) \in U \) then

\[
|Df^n(z)| \geq C' \lambda^n,
\]

(were \( C' \) only depends on \( U' \)). If \( q = 1 \) we can set \( U_1 = U \) and \( \hat{U} = U' \).
Proof. Let us first consider the unperturbed map $f_0$. By the argument before the lemma, it follows from [13] that the Collet-Eckmann condition implies the second Collet-Eckmann condition, for all critical points. Looking at any iterated preimage $z = f_{-n}(c)$ to a critical point $c$, the second Collet-Eckmann condition implies

$$|Df^n(z)| \geq C_2 \lambda_n^n,$$

for some $\lambda_2 > 1$ and a constant $C_2 > 0$. Let $0 < \kappa < 1$ and $N > 0$ (we give conditions of these constants below). We follow partially the idea of [31] (pp. 40–41). Let $U'$ to be a union of disks $U_j'$ around the critical points with radius $\delta'$, so that for any iterated preimage $f^{-k}(U_j')$ of a component of $U'$, we have

$$\text{(3.2)} \quad \text{diam}(f^{-k}(U_j')) \leq \kappa \cdot \text{dist}(f^{-k}(U_j'), \text{Crit}_0), \quad \text{for all } k \leq N.$$  

This implies that we have distortion inside $f^{-k}(U_j')$, that is, for any choice of $z, w$ in the same component of $f^{-k}(U_j')$ we have

$$\text{(3.3)} \quad \frac{|Df(z)|}{|Df(w)|} \leq C_3,$$

where $C_3 = C_3(\kappa) \to 1$, as $\kappa \to 0$.

If (3.2) is not valid, then we can use another estimate as follows. For any disk $D$ of radius at most $\delta' > 0$ there is a constant $C_4$ such that

$$|Df(z)| \text{diam}(D') \leq C_4 \text{diam}(D), \quad \text{for all } z \in D',$$

where $D'$ is a component of $f^{-1}(D)$. Here $C_4$ only depends on $\delta'$.

Suppose now that $N_0 > 0$ is the largest time where (3.2) is valid. If we put $W_k' = f^{-k}(U_j')$ and $z_k \in W_k'$ the corresponding preimage of $c_j \in U_j'$, and if $N$ is large enough, then

$$\text{diam}(U_j') \geq C_3^{-(N_0 - 1)} |Df^{N_0}(z_{N_0 - 1})| \text{diam}(W'_{N_0 - 1})$$

$$\geq C_4^{-1} C_3^{-(N_0 - 1)} |Df^{N_0}(z_{N_0})| \text{diam}(W'_{N_0}) \geq \lambda_1^{-N_0} \text{diam}(W'_{N_0}),$$  

for some $\lambda_1 > 0$. Now let $N$ be so large so that this holds and also that $\lambda_1^{-N_0} \geq 10/\kappa$. So from now on $U'$ and $N$ are fixed.

Now suppose that $U_1 \subseteq U \subseteq \hat{U} \subseteq U'$, and let $U_1 = \cup_j B(c_j, \delta_1)$, $U = \cup_j B(c_j, \delta)$, $\hat{U} = \cup_j B(c_j, \hat{\delta})$ i.e., $\delta_1 \leq \delta \leq \hat{\delta} \leq \delta'$. Suppose that $z \notin U_1$, $f^k(z) \notin U$ for all $k = 1, \ldots, n - 1$ and $f^n(z) \in U$. Let now $n_0 > 0$ be the first time for which (3.2) is not valid with $U_j'$ replaced by $U_j$, the components of $U$. Let $W_k = f^{-k}(U_j)$ be the corresponding preimages of $c_j \in U_j$. By the definition of $n_0$, 

$$\text{dist}(W_{n_0}, \text{Crit}) \leq \frac{1}{(1/\kappa)} \text{diam}(W_{n_0}) \leq (1/\kappa) \lambda_1^{-n_0} \text{diam}(U_j).$$

Let us now consider the condition

$$\text{(3.6)} \quad \frac{1}{(1/\kappa)} \lambda_1^{-n_0} \leq \frac{q}{10} \leq \frac{\text{diam}(U_{1,j})}{10 \text{diam}(U_j)},$$

where $U_{1,j} \subseteq U_j$ is the corresponding component of $U_1$ inside $U_j$ and the second inequality is valid by assumption. We discuss this condition soon. It implies that

$$\text{dist}(W_{n_0}, \text{Crit}) \leq \frac{\text{diam}(U_{1,j})}{10 \text{diam}(U_j)} \text{diam}(U_j) = \frac{1}{10} \text{diam}(U_{1,j}), \quad \text{and}$$

$$\text{(3.8)} \quad \text{diam}(W_{n_0}) \leq \lambda_1^{-n_0} \text{diam}(U_j) \leq \frac{\kappa}{10} \text{diam}(U_{1,j}).$$
Clearly, this implies that \( W_{n_0} \subset U_{1,j} \subset U_1 \). If \( n_0 \leq n \), this was not allowed, since \( z_k \notin U_1 \), \( 1 \leq k \leq n \). Hence (3.2) is valid all the time up until \( n \). Therefore, if \( w \in f_0^n(U) \), we have, by the distortion estimate (3.3),

\[
|Df_0^n(w)| \geq |Df_0^n(z)|c_3^{-n} \geq C_2 \lambda_1^n,
\]

where \( z \) is the preimage of the corresponding critical point and \( C_2 \) is the constant form the second Collet-Eckmann condition, and hence does not depend on \( U \).

Let us now discuss the condition (3.6). Then the condition implies that

\[
n_0 \log \lambda_1 \geq \log(1/q) - \log \kappa + \log 10 \geq \Delta_1 - \Delta - \log \kappa + \log 10.
\]

Hence this basically forces \( n_0 - 1 \), the time when (3.2) is valid, to be bounded below by the difference \( \Delta_1 - \Delta \). Let now \( \hat{N} \) be the largest integer such that (3.2) is valid with \( U_j \) replaced by \( \hat{U}_j \), the components of \( \hat{U} \). Then we can say that \( \hat{U} \) depends on \( q \) in the following sense. For a fixed \( q \) we choose \( \hat{N} \) so that the corresponding \( \hat{N} \) satisfies (3.10), with \( n_0 \) replaced by \( \hat{N} \). Clearly, if \( q = 1 \) we can put \( \hat{U} = U' \).

From a classical result by R. Mañé, we have an estimate as follows. If \( f_0^k(z) \notin U \) for \( k = 0, \ldots, n \), then

\[
|Df_0^n(z)| \geq C \lambda_3^n,
\]

for some constant \( C > 0 \) that depends on \( U \) and \( \lambda_3 > 1 \) which does not depend on \( U \). We may assume that \( \lambda_3 > \lambda_1 \), otherwise diminish \( \lambda_1 \) so that this holds. This proves the first statement of the lemma.

Choose \( N_1 > 0 \) so that outside \( U_1 \) the orbits \( f_a^k(z) \) and \( f^k(z) \) follow each other up to \( N_1 \), i.e. for \( k \leq N_1 \), and so that \( |Df_a^{N_1}(z)| \geq C \lambda_3^{N_1} \geq \lambda_1^{N_1} \), for all \( a \in (-\varepsilon, \varepsilon) \). Here \( \lambda_3 > 1 \) comes from a perturbed version of (3.11). Since also (3.9) is valid for small perturbations if we bound the number of iterations by \( N_1 \), we let \( \lambda_1 > 1 \) be the corresponding perturbed version of \( \lambda_1 > 1 \). Let us write \( n = qN_1 + r \), where \( r < N_1 \). Then, if we assume that \( z \notin U_1, f_a^k(z) \notin U \) for \( k = 1, \ldots, n - 1 \) and \( f_a^n(z) \in U \) we get

\[
|Df_a^n(z)| = |Df_a^n(f^{qN_1}(z))||Df_a^{N_1}(f^{(q-1)N_1}(z))| \ldots |Df_a^N(z)| \geq C_2 \lambda_1^n,
\]

where we used (3.9) for \( |Df_a^n(f^{qN_1}(z))| \geq C_2 \lambda_1^n \), so that \( |Df_a^n(f^{qN_1}(z))| \geq C_2 \lambda_1^n \). The second statement of the lemma follows with \( \lambda = \bar{\lambda}_1 \).

The classical outside expansion lemma is obtained by setting \( U_1 = U \) in the above lemma, i.e. \( q = 1 \). From [14], it can be seen that the Lyapunov exponent from the second Collet-Eckmann condition is inherited from the exponent from the ordinary Collet-Eckmann condition. Hence the uniform “outside exponent” \( \log \bar{\lambda}_1 \), is close to the Lyapunov exponent for the starting map \( f_0 \) (but likely lower than it), depending on the neighbourhood \( U' \). Let us set \( \gamma_H = \log \bar{\lambda}_1 \).

4. Parameter-phase distortion

One fundamental result we need is the comparison between space and parameter-derivatives. This has been proved in [4, 5] and many other papers. But for our purposes we need a stronger form of this result due to G. Levin. We use a normalised space, described in [22], of maps in \( \mathcal{R}_d \) as follows. We consider the set \( \Lambda_{d,p'} \subset \mathcal{R}_d \) of all rational maps of degree \( d \) with exactly \( p' \) distinct critical points \( c_j \) with
corresponding multiplicities \( m_j, 1 \leq j \leq p' \), where \( \overline{p} = \{ m_1, \ldots, m_{p'} \} \), normalised to that every map \( f \in \Lambda_{d, \overline{p}} \) has the form

\[
f(z) = az + b + \frac{P(z)}{Q(z)},
\]

where \( a \neq 0 \), and \( \deg(P) \leq d - 2, \deg(Q) \leq d - 1 \) and where \( P \) and \( Q \) have no common zeros. By Proposition 8 in [22], every \( f \in \mathcal{R}_d \) is conjugate by a Möbius transformation to some \( \tilde{f} \in \Lambda_{d, \overline{p}} \). So we can view \( \mathcal{R}_d \) as a union of sets of the type \( \Lambda_{d, \overline{p}} \) up to equivalence by Möbius transformations. Note that in every such set, critical points do not split.

We assume that the real analytic family \( f_a \in \Lambda_{d, \overline{p}}, a \in (-\varepsilon, \varepsilon) \) around \( f_0 \) has a tangent vector \( \overline{a} \neq \overline{0} \). Hence \( f_a(z) = f_0(z) + au(z) + O(a^2) \) for some \( u \neq 0 \). Let us write \( \xi_{n,l}(\overline{a}) = f^n_a(c_l(\overline{a})) \), where \( \overline{a} = (a_1, a_2, \ldots, a_{p'}) \) is a parameterisation of the parameter space \( \Lambda_{d, \overline{p}} \) around \( \tilde{f} = f_0 \), where \( f_0 \) corresponds to \( (a_1, a_2, \ldots, a_{p'}) = (0, 0, \ldots, 0) \) and where \( c_j = c_j(a_1, a_2, \ldots, a_{p'}) \). In [22] and [23], it is proven that the matrix \( L \) formed by the numbers

\[
L(c_l, a_k) = \lim_{n \to \infty} \frac{\partial \xi_{n,l}(\overline{a})}{\partial a_k} \frac{\partial \xi_{n,l}(\overline{a})}{\partial (f^0_{c_l}(f^0_{c_l}))}
\]

is non-degenerate. Let \( \overline{a} = (u_1, u_2, \ldots, u_{p'}) \) be a tangent vector of unit length, i.e., a vector in \( \mathbb{R}^{\Lambda_{d, \overline{p}}} \), and suppose that this is tangent to the family \( f_a \) at \( a = 0 \). Then for almost all directions, i.e. tangent vectors, we have that all entries of \( L \cdot \overline{a} \) are non-zero, since the set of directions when this is not true is a finite union of sets of co-dimension 1 in \( \mathbb{R}^{\Lambda_{d, \overline{p}}} \). This means precisely that, for almost all directions, the limits

\[
\lim_{n \to \infty} \frac{\xi'_{n,l}(0)}{Df^n_{\overline{a}}(f^0_{c_l})} = \sum_k a_k L(c_l, a_k),
\]

is non-zero for every \( l \), where we mean \( \xi'_{n,l}(0) = \frac{\partial \xi_{n,l}(\overline{a})}{\partial c_l} \bigg|_{a=0} \). We thus say that the real analytic family \( f_a \) around \( f_0 \) is non-degenerate, if its tangent vector satisfies this condition, (in addition to the fact that the critical points move analytically).

We summarise this result as a proposition below, which is a direct consequence of Theorem 1 combined with Corollary 2.1, part (8), in [23]. It is a generalisation of a corresponding result in [22] Theorem 1.1.

**Proposition 4.1 (Levin).** Suppose that \( f \) is a rational map with summable critical points without parabolic cycles such that \( J(f) = \hat{\mathbb{C}} \), and suppose that \( f_a \) is a non-degenerate real analytic family around \( f_0, a \in (-\varepsilon, \varepsilon) \). Then for each critical point \( c_l(a) \), the limit

\[
\lim_{n \to \infty} \frac{\xi'_{n,l}(0)}{Df^n_{\overline{a}}(v_l)} = L_l
\]

exists and is different from 0 and \( \infty \).

Indeed, a CE-map has all its critical values summable so the above proposition can be used. We also note that by [14] any Collet-Eckmann map different from a flexible Lattés map carries no invariant line field on its Julia set. We now use this result, to make small perturbations.
Lemma 4.2. Assume that $f_0$ satisfies the CE-condition with exponent $\gamma$. For any $0 < \gamma_1 < \gamma$ and $0 < q < 1$, there exists $N > 0$ and $\varepsilon > 0$ such that if $f_a$, $a \in (-\varepsilon, \varepsilon)$ satisfies the CE-condition up to time $m \geq N$ with exponent $\gamma_1$, we have

$$\left| \frac{\xi_{m,l}^1(a)}{Df_a^{m-1}(v_l(a))} - L_l \right| \leq q |L_l|,$$

for every $l$.

Proof. According to Theorem 1 in [23], we have for $a = 0$,

$$\lim_{n \to \infty} \frac{\xi_{n,l}^1(0)}{Df^{n-1}(v_l(0))} = \sum_{n=0}^{\infty} \frac{\partial_a f_0(\xi_{n,l}(0))}{Df^n(v_l(0))} = L_l.$$

Let us put $\xi_{m,l}(a) = \xi_m(a)$ and $L_l = L$. The reader may verify that for small perturbations $a$ close to 0,

$$\left| \frac{\xi^1_m(a)}{Df_a^{m-1}(v_l(a))} - \sum_{n=0}^{m-1} \frac{\partial_a f_a(\xi_n(a))}{Df^n_a(v_l(a))} \right| = \sum_{n=0}^{m-1} \frac{\partial_a f_a(\xi_n(a))}{Df^n_a(v_l(a))}.$$

We have that $|\partial_a f| = |\partial_a f(z, a)|$ is bounded by some constant $B > 0$. We choose $N > 0$ so that the series

$$(4.2) \quad \sum_{n=N+1}^{\infty} \frac{B}{C e^{\gamma_1 n}} \leq \min_l (q |L_l|/4).$$

By continuity, there exists some $\varepsilon > 0$ such that if $a \in (-\varepsilon, \varepsilon)$ then

$$\left| \sum_{n=0}^{N} \frac{\partial_a f_a(\xi_{n,l}(a))}{Df_a^{n}(v_l(a))} - L_l \right| \leq q |L_l|/2.$$

Since $f_a$ is assumed to satisfy the CE-condition with exponent $\gamma_1$, by (4.2) we get that the tail satisfies

$$\left| \sum_{n=N+1}^{m} \frac{\partial_a f_a(\xi_{n,l}(a))}{Df_a^{n}(v_l(a))} \right| \leq q |L_l|/2,$$

for all $a \in (-\varepsilon, \varepsilon)$ and all $m \geq N$. This finishes the lemma. \qed

When we use this lemma we want to choose $N$ and $\varepsilon$ so that the above lemma is valid for $\gamma_L = (1/6) \min(\gamma_0, \gamma_H)(1 - \tau)$, where $0 < q < 1$ is small, and where $\tau$ is a constant, $0 < \tau < 1$.

5. Weak parameter dependence and weak distortion

We will later see that the expansion of the space derivative induces a great deal of parameter independence. This follows a posteriori from the Main Distortion Lemma 7.3 and the Starting Lemma 7.4, but to start we now prove a weaker statement.

Lemma 5.1. Let $N$ be as in Lemma 4.2 and let $\gamma_1 \geq (1/4) \min(\gamma_H, \gamma_0)(1 - \tau) > 0$. Suppose that $a, b \in (-\varepsilon, \varepsilon)$, where $(-\varepsilon, \varepsilon)$ is a parameter interval around $f_0$ for which $f_a$ is the real analytic family we are considering. If $\varepsilon > 0$ is sufficiently small we have the following. Suppose that we have;

i) for all $n \leq N$, that $|Df^n_a(v_l(a))| \geq C_1 e^{\gamma_1 n}$, and for $k \leq k_1$ for some $k_1 \geq 0$, that $|Df^k_a(\xi_{N,l}(a))| \geq C_2 e^{\gamma_1 k}$,
ii) for all $n \leq N + k_1$, if $\xi_{n,l}(a), \xi_{n,l}(b) \notin U$, then $|\xi_{n,l}(a) - \xi_{n,l}(b)| \leq S$ and, if $\xi_{n,l}(a) \in U$ or $\xi_{n,l}(b) \in U$ (or both), then

$$|\xi_{n,l}(a) - \xi_{n,l}(b)| \leq \text{dist}(\xi_{n,l}(c), \text{Crit}_c)/(\log(\text{dist}(\xi_{n,l}(c), \text{Crit}_c)))^2,$$

where $c \in \{a, b\}$ is such that $\text{dist}(\xi_{n,l}(c), \text{Crit}_c)$ is minimal,

iii) that $f_a$ is slowly recurrent up to time $N + k_1$.

Then there exists $Q > 1$ (arbitrarily close to 1, if $N$ is large enough), and $\gamma_2 > 0$ (arbitrarily close to but slightly smaller than $\gamma_1$), such that

$$|\xi_{N+k,l}(a) - \xi_{N+k,l}(b)| \sim_{Q^k} |Df_a^k(\xi_{N,l}(a))| |\xi_{N,l}(a) - \xi_{N,l}(b)|$$

and

$$|\xi_{N+k,l}(a) - \xi_{N+k,l}(b)| \geq |\xi_{N,l}(a) - \xi_{N,l}(b)| C_2 \gamma_2^k,$$

for any $k \leq k_1$, where $\gamma_2 > 0$ is slightly smaller than $\gamma_1 > 0$.

**Proof.** Since we assume that the critical points $c_l(a)$ move analytically in $a$ we have

$$c_l(a) = K_l a^{k_l} + O(a^{k_l+1}).$$

Let us fix $l$ and consider $\xi_{n,l}(a) = \xi_{l}(a)$. If we consider a sufficiently small parameter interval $(\varepsilon, \varepsilon)$ centred at $a = 0$ corresponding to $f_0$, then, by bounded distortion, we can make $\varepsilon$ so small that we have, for any two points $a, b \in (-\varepsilon, \varepsilon)$,

$$|\xi_{N}(a) - \xi_{N}(b)| \sim_2 |\xi_{N}(c)||a - b|,$$

for any $c \in (-\varepsilon, \varepsilon)$. From the assumption $|Df_a^N(v_l(a))| \geq C_1 \varepsilon_1^N$ we see that we may choose $\varepsilon > 0$ small enough to get $|(f_a^N)'(v_l(b))| \geq C_1 \varepsilon_1^{2N}$ for $b \in \omega = (-\varepsilon, \varepsilon)$ for some $\gamma_2 > 0$ slightly smaller than $\gamma_1$. From Lemma 4.2 we now get, with $q \leq 1/2$ and $L = |L_i|$, for any $c \in (a, b)$,

$$|\xi_{N}(a) - \xi_{N}(b)| \sim |\xi_{N}(c)||a - b| \geq qL|a - b||Df_a^{N-1}(v_l(c))|$$

$$\geq C_1 qL \varepsilon_1^{2(N-1)}|a - b| = C'_1 \varepsilon_1^{2N} |a - b|,$$

where $C'_1 = C_1 qL \varepsilon_1^{-2}$. During the first $N$ iterates, (5.2) implies that for $a$ and $b$ close to 0 we have

$$|c_l(a) - c_l(b)| \leq 2K_l k_l a^{k_l-1} |a - b| \leq C_1 a^{k_l-1} |\xi_{N,l}(a) - \xi_{N,l}(b)| e^{-\gamma_1 N},$$

for some constant $C$. It follows that

$$|\xi_{N,l}(a) - \xi_{N,l}(b)| \gg |c_l(a) - c_l(b)|$$

for all critical points.

Suppose that, for all $0 \leq j \leq k \leq k_1 - 1$, we have

$$|\xi_{N+j}(a) - \xi_{N+j}(b)| \geq C_2 \gamma_2^j |\xi_{N}(a) - \xi_{N}(b)|.$$

We may assume that $\gamma_2 < \gamma_1 < \gamma_0$. Combining (5.4) and (5.2) we conclude that (5.3) holds for $N$ replaced by $N + j$.

For the proof, put $\xi_0(a) = \xi_{0,l}(a)$. Since we assume that the orbit of $w = \xi_0(a)$ stays close to $z = \xi_0(b)$ we have a distortion estimate

$$\frac{1}{C} \leq \frac{|Df_a(w)|}{|Df_b(z)|} \leq C,$$
for some constant \( C \geq 1 \). This constant can be arbitrarily close to 1 if \(|z - w| \leq S \) and \( S \) is small enough (for \( z, w \not\in U \) and \(|z - w| \leq e^{-r}/r^2 \) (if \( \text{dist}(z, \text{Crit}) \in (e^{-r-1}, e^{-r}) \)). Hence iterating this we get

\[
|Df^K_0(\xi_N(b))| \geq \frac{1}{C^k}|Df^K_0(\xi_N(a))| \geq C_2 e^{\gamma_2 k},
\]

for some \( \gamma_2 > 0 \) slightly smaller than \( \gamma_1 \) (at least we may assume, for instance, that \( \gamma_2 \geq (9/10)\gamma_1 \)).

With \( B = \sup |\partial_a f_a| \), using (5.2) and (5.4), there is some \( Q_0 > 1 \) such that

\[
|\xi_{N+k+1}(a) - \xi_{N+k+1}(b)| \\
\geq |f_a(\xi_{N+k}(a)) - f_a(\xi_{N+k}(b))| - |f_a(\xi_{N+k}(b)) - f_b(\xi_{N+k}(b))| \\
\sim_{Q_0} |Df_a(\xi_{N+k}(a))||\xi_{N+k}(a) - \xi_{N+k}(b)| - |\partial_a f_a(\xi_{N+k}(a))||a - b| \geq |Df_a(\xi_{N+k}(a))||\xi_{N+k}(a) - \xi_{N+k}(b)| \\
= \left(|Df_a(\xi_{N+k}(a))| - \frac{B}{C_1 C_2} e^{-\gamma_2(N+k)}||\xi_{N+k}(a) - \xi_{N+k}(b)|. \right.
\]

It is easy to check that a reverse inequality also holds. Note that \( Q_0 \) can be chosen arbitrarily close to 1 if \( N \) is large enough and \( S = \delta e_1 \) is small enough (i.e. \( e_1 \) small enough). Repeating this \( k \) more times we get

\[
(5.5) \quad |\xi_{N+k+1}(a) - \xi_{N+k+1}(b)| \\
\sim_{Q_0} |Df^{k+1}_a(\xi_N(a))| \prod_{j=0}^{k} \left( 1 - \frac{B e^{-\gamma_2(N+j)}}{C_1 C_2 |f_j^{(k)}(\xi_{N+j}(a))|} \right)|\xi_N(a) - \xi_N(b)|.
\]

Now we use that \( f_a \) is slowly recurrent (actually this implies that \( f_a \) is also slowly recurrent since \( \xi_a(a) \) stays close to \( \xi_a(b) \)). But we have assumed from the beginning that the exponent in basic assumption (which is inherited from the slow recurrence condition) satisfies \( \alpha \leq (1/(400KT)) \min(\gamma_H, \gamma_0)(1 - \tau) \ll \gamma_1 \). Then \( \alpha \) is also much smaller than \( \gamma_2 \), and, if \( N \) is sufficiently large, for some constant \( C > 0 \),

\[
|Df_a(\xi_N+j(a))| \geq C e^{-K\alpha(N+j)} \gg e^{-\gamma_2(N+j)}, \text{ for all } 0 \leq j \leq k_1.
\]

We see that the sum \( \sum_{j=0}^{\infty} e^{-(\gamma_2 - \alpha)(N+j)} \) can be made as small as we like. Therefore, the product

\[
\prod_{j=0}^{k} \left( 1 - \frac{B e^{-\gamma_2(N+j)}}{C_1 C_2 |f_j^{(k)}(\xi_{N+j}(a))|} \right) \geq \prod_{j=0}^{\infty} \left( 1 - C e^{-(\gamma_2 + \alpha)(N+j)} \right) > \frac{1}{Q_1}
\]

for some \( Q_1 > 1 \) (independent of \( k \)). Therefore,

\[
(5.6) \quad |\xi_{N+k+1}(a) - \xi_{N+k+1}(b)| \sim_{Q_0} |Df^{k+1}_a(\xi_N(a))||\xi_N(a) - \xi_N(b)|
\]

Since \( |Df^{k+1}_a(\xi_N(a))| \geq C_2 e^{\gamma_1(k+1)} \) we have \( Q_0^k Q_1 |Df^{k+1}_a(\xi_N(a))| \geq C_2 e^{\gamma_2(k+1)}, \) for some \( \gamma_2 > 0 \) slightly smaller than \( \gamma_1 \), given that \( Q_0 \) and \( Q_1 \) are sufficiently close to 1. Hence we have (5.4) satisfied with \( k \) replaced by \( k + 1 \) and we can continue the same argument and obtain (5.4) up until \( k_1 \). This settles both claims with \( Q^k = Q_0^k Q_1 \). 

\[\square\]
We can easily get a little more general statement. If \( f_a \) satisfies the CE-condition and is slowly recurrent up until time \( N + k_1 \), we can use (5.5), and the following arguments to obtain

\[
|\xi_n(a) - \xi_n(b)| \sim Q \cdot |Df^n_a(\xi_{n-j}(a))||\xi_{n-j}(a) - \xi_{n-j}(b)|,
\]

if \( n - j \geq N \) and \( n \leq N + k_1 \). The details are left to the reader.

Remark 5.2. We have seen that the parameter dependence is inessential as long as the derivative of \( |f^n_a(v_l(a))| \) grow with a certain Lyapunov exponent \( \gamma_I \). We call this the \textit{weak parameter dependence property}. We will also require that the Lyapunov exponent never goes below a certain “critical” level, which is related to the “intermediate level” \( \gamma_I = (1/3) \min(\gamma_H, \gamma_0)(1 - \tau) \). Since \( \gamma_L = (1/6) \min(\gamma_H, \gamma_0)(1 - \tau) \) \( < (1/4) \min(\gamma_H, \gamma_0)(1 - \tau) \), then \( \gamma_C = (1/4) \min(\gamma_H, \gamma_0)(1 - \tau) \) (the critical exponent) as a lower bound for \( \gamma_I \) will do. We also let \( \gamma_B = (3/4) \min(\gamma_H, \gamma_0)(1 - \tau) \). This \( \gamma_B \) is the Lyapunov exponent that we want to keep at the end.

6. Distortion and expansion during the bound period

We use the following notations. Below \( \omega \) is assumed to be an interval and a partition element according to Definition 2.3.

Definition 6.1. We say that \( a \in \mathcal{E}_{n,l}(\gamma) \) if

\[
|Df^n_a(v_l(a))| \geq C_0 e^{\gamma_k}, \text{ for all } k \leq n - 1, \text{ and}
\]

(6.2) \[|Df^n_a(v_j(a))| \geq C_0 e^{\gamma_k}, \text{ for all } k \leq (2K\alpha/\gamma_I)n, \text{ and } j \neq l.\]

We say that \( a \in \mathcal{B}_{n,l} \) if

(6.3) \[\text{dist}(\xi_{k,l}(a), \text{Crit}_a) \geq K_0 e^{-2\alpha k}, \text{ for all } k \leq n \text{ and}\]

(6.4) \[\text{dist}(\xi_{k,l}(a), \text{Crit}_a) \geq K_0 e^{-2\alpha k}, \text{ for all } k \leq (2K\alpha/\gamma_I)n \text{ and } j \neq l.\]

We say that \( \omega \in \mathcal{E}_{n,l,*}(\gamma) \) if (6.1) holds and (6.2) holds with \((2K\alpha/\gamma_I)\) replaced by \(4(K\alpha/\gamma_I)\). We say that \( \omega \in \mathcal{B}_{n,l,*} \) if (6.3) holds and (6.4) holds with \((2K\alpha/\gamma_I)\) replaced by \(4(K\alpha/\gamma_I)\).

Note that \( \omega_0 \subset \mathcal{E}_{n,l}(\gamma) \cap \mathcal{B}_{n,l} \) for all \( l \) for some \( \gamma \) close to \( \gamma_0 \). The definitions above is tailored so that if an interval belongs to \( \mathcal{E}_{n,l}(\gamma) \) or \( \mathcal{B}_{n,l} \) then we can use the binding information for the other critical points up until some fraction \( 2K\alpha/\gamma_I \) of the time \( n \). The star is added to be able to use the binding information longer and continue the parameter-exclusion construction up until \( 2n \).

To prove bounded distortion, we will frequently make use of the following lemma, which is standard.

Lemma 6.2. Given complex numbers \( z_1, \ldots, z_n \) we have

\[
\prod_{j=1}^{n} |z_j - 1| \leq -1 + \exp \sum_{j=1}^{n} |z_j - 1|.
\]

Expanding \( f \) in Taylor series near a critical point \( c \) gives

\[
f_a(z) = A(z - c)^k + O((z - c)^{k+1}), \quad Df_a(z) = Ak(z - c)^{k-1} + O((z - c)^k),
\]
where \( A \) is analytic in the parameter \( a \). If \( z \) and \( w \) are close to \( c \) and \( |z - c| \sim |w - c| \), we get,

\[
(6.5) \quad Df_a(z) - Df_a(w) = Ak(z - w)((z - c)^{k-2} + (z - c)^{k-3}(w - c) + \ldots
\]

\[
+ (w - c)^{k-2} + O((z - c)^{k-1})
\]

Hence,

\[
\sum_{j=1}^{n} \frac{|Df_a(\xi_j(a)) - Df_a(\xi_j(b))|}{|Df_a(\xi_j(b))|} \sim 2k \sum_{j=1}^{n} \frac{|\xi_j(a) - \xi_j(b)|}{\text{dist}(\xi_j(b), \text{Crit}_a)}
\]

if \( z \) and \( w \) are sufficiently close to \( \text{Crit}(f) \). In Section 7 we will allow the parameter to vary as well, and have to go a bit further.

**Lemma 6.3** (Distortion during the bound period). Let \( \varepsilon' > 0 \). Then if \( \delta' = e^{-\Delta'} \) is sufficiently small and \( N \) sufficiently large, the following holds. Let \( z = \xi_{v,i}(a) \) be a free return into \( U' \), \( \nu \geq N \), where \( a \in E_{\nu,i}(\gamma) \cap B_{\alpha,i} \) for some \( \gamma \geq \gamma_1 \). Then we have, for all \( w \) on the line segment between \( f_a(z) \) and \( \xi_{v,i}(a) = v_i(a) \),

\[
\left| \frac{Df_a(w)}{Df_a(v_i(a))} - 1 \right| \leq \varepsilon',
\]

for \( j \leq p \), where \( p \) is the length of the bound period for \( z \).

**Proof.** We first prove the lemma for \( w = f_a(z) \). Let \( \text{dist}(\xi_{v,i}(a), \text{Crit}_a) \sim \sqrt{e} e^{-r} \) where \( \xi_{v,i}(a) = z \) and put \( z_j = f_a^j(z) \) and \( \xi_{v,i}(a) = \xi_j(a) \). Following the discussion preceding the lemma, we estimate, for \( \nu \geq N \), the sum

\[
\sum_{j=1}^{p} \frac{|Df_a(z_j) - Df_a(\xi_j(a))|}{|Df_a(\xi_j(a))|} \leq C \sum_{j=1}^{p} \frac{|z_j - \xi_j(a)|}{\text{dist}(\xi_j(a), \text{Crit}_a)}
\]

The last sum can be divided into two subsums \([1, J]\cup[J+1, p]\) where \( J = \text{dr}/(10(2\alpha + \Gamma)) \), where \( d \) the degree of \( f_0 \) at \( c_a \), and \( \Gamma = \sup_{a \in (-\varepsilon, \varepsilon), z \in \mathbb{C}} \log |f_a'(z)| \). Assuming that the basic approach rate assumption holds, the first sum can be estimated as

\[
\sum_{j=1}^{J} \frac{|z_1 - \xi_1(a)| e^{\varepsilon j}}{K b e^{-2\Delta}} \leq \sum_{j=1}^{J} CK^{-1} e^{-(\Gamma + 2\alpha)j} \leq \sum_{j=1}^{J} Ce^{-(9/10 \delta) j} \leq Ce^{-9\Delta'/10}
\]

The second sum can be estimated using the definition of the bound period (remember \( \beta = \alpha \)),

\[
\sum_{j=J+1}^{p} \frac{|z_j - \xi_j(a)|}{\text{dist}(\xi_j(a), \text{Crit}_a)} \leq C \sum_{j=J+1}^{p} e^{-\alpha j} \leq Ce^{-\alpha \text{dr}/(10(2\alpha + \Gamma))}
\]

We see that both sums can be made arbitrarily small if \( \Delta' \) is large enough. This finishes the case \( w = f_a(z) \).

It is easy to see that the same must hold one the line segment between \( f_a(z) \) and \( v_i(a) \). Let \( p' \leq p \) be the least bound period for all such \( w \) on this line. That means that up until \( j = p' \) the distortion estimate holds for all \( w \) on the line. But since \( \varepsilon' \) may be chosen very small this means that the image of the line under \( f_a^{p'} \) is an almost straight line too. It follows that the corresponding \( w \) for \( p' \) has to be \( z \) in fact, so \( p = p' \).

Note that the condition on \( U' \) in the above lemma is only depending on the starting family of functions \( f_a, a \in (-\varepsilon, \varepsilon) \). There is also a condition on \( U'' \) in Lemma 3.1.
Lemma 6.4. Suppose that $\xi_{\nu,l}(a)$ is a return into $U'_i$ and that $a \in E_{\nu,l}(\gamma) \cap B_{\nu,l}$ for some $\gamma \geq \gamma_1$. Then if $N$ is large enough and $p$ is the length of the following bound period we have,

$$|Df_a^p(\xi_{\nu,l}(a))| \geq e^{\frac{2\nu}{\gamma}p},$$

where $d_i$ is the degree of $f$ at $c_i$.

Moreover, if $\text{dist}(\xi_{\nu,l}(a), \text{Crit}_a) \sim \sqrt{r} e^{-r}$, then

$$\frac{d_i r}{2\nu} \leq p \leq \frac{2d_i r}{\gamma}.$$

In particular, $p \leq 2\alpha d_i \nu / \gamma$, where $\alpha$ is the exponent in the basic assumption and $\Gamma = \sup_{a \in (-\varepsilon, \varepsilon), x \in \mathbb{C}} \log |Df_a(z)|$.

Proof. Put $D_j = |Df_a^j(\xi_{\nu,l}(a))|$ and $E_j = |Df_a^j(\xi_{\nu+1,l}(a))|$ for some $a \in \omega$. We have

$$D_j \geq CK_b e^{-2\alpha K_j},$$

where $\gamma' \geq \gamma - 4\alpha K \geq \gamma_C$, provided $N$ is large enough (recall $\nu \geq N$). We can also use Lemma 6.3 to get the following distortion estimate, for some $C > 1$ (close to $1$),

$$|\xi_{\nu+j,l}(a) - \xi_{j,l}(a)| \sim_C |Df_a^j(\xi_{\nu,l}(a))||\xi_{\nu,l}(a) - \xi_{0,l}(a)|,$$

for $j \leq p + 1$. Suppose that $|\xi_{\nu,l}(a) - \xi_{0,l}(a)| \sim 2^{-r}$. We know from the definition of the bound period and the basic assumption, that

$$D_{p+1} e^{-r} \geq \frac{1}{4C} \text{dist}(\xi_{p+1,l}(a), \text{Crit}_a) e^{-\alpha(p+1)} \geq \frac{1}{4C} K_b e^{-2\alpha(p+1) - \beta(p+1)}.$$

Also we have, for some $\kappa_1 \geq 1$,

$$D_{p+1} e^{-r} \sim_{\kappa_1} E_p e^{-\rho d_i},$$

and so

$$e^{-r(d_i - 1)} \sim_{\kappa_1} \left( D_{p+1} e^{-r} \right)^{\frac{d_i - 1}{d_i}} E_p^{\frac{d_i - 1}{d_i}}.$$

(6.7)

Now we can use that $2\alpha + \beta$ is very small compared to $\gamma \geq \gamma_1$. We get,

$$D_{p+1} \sim_{\kappa_1} e^{-r(d_i - 1)} E_p$$

$$\geq \left( \frac{K_b}{4C} \right)^{\frac{d_i - 1}{d_i}} E_p^{\frac{d_i - 1}{d_i}} e^{-3\alpha(p+1) \frac{d_i - 1}{d_i}}$$

(6.8)

if $\nu$ is sufficiently large. Since $D_{p} = D_{p+1}/|Df_a(\xi_{\nu+p}(a))|$, with minor modifications it is easy to see that the same estimate holds for $D_p$. 


To prove the second claim, we note that from (6.6), the slow recurrent condition and the fact that \( |Df^p(v_l(a))| \leq e^{\nu \Gamma} \), we get that, for some very small \( \alpha > 0 \) in comparison to \( \gamma \),

\[
\left| Df^p_{\nu}(v_l(a)) \right| \leq \frac{K_b}{4C_1} e^{-3\alpha(p+1)},
\]

which gives the left inequality if \( \nu \geq N \) is large enough. To prove the right inequality, we note that the spherical distance \( \text{dist} (\xi_{\nu, l}(a), \text{Crit}_a) \) is bounded from above. By the definition of the bound period (now we are considering the time \( p \) iterates from the return into \( U \)), and the fact that we also have \( E_{p-1}e^{-\delta r} \sim \kappa_1 Dp e^{-r} \),

\[
(C_0/2)e^{\gamma(p-1)}e^{-\delta r} \leq E_{p-1}e^{-\delta r} \leq 4C_1 e^{-\alpha p} \text{dist}(\xi_{p, l}(a), \text{Crit}_a).
\]

and the right inequality follows. □

The above lemma gives a quite substantial amount of increase of the derivative during the bound period, even if there is a loss in the first iterate. We can also see that under all circumstances,

\[
|\xi_{\nu+p}(a) - \xi_{\nu+p}(b)| \geq |\xi_{\nu}(a) - \xi_{\nu}(b)|.
\]

### 7. Strong distortion

Our aim now is to use weak distortion and prove that we actually have something stronger, namely, for some small \( \varepsilon' > 0 \),

\[
\left| \frac{Df^n_a(v_l(a))}{Df^n_b(v_l(b))} - 1 \right| \leq \varepsilon',
\]

for all \( a, b \in \omega \) where \( \omega \) is a partition element according to Definition 2.3. We will also make use of the preliminary discussion in Section 6. Let us first see a geometrical consequence of (7.1). By Lemma 4.2 we have, for some small \( 0 < q < 1 \),

\[
\left| \frac{\xi_{n, l}^\prime(a)}{Df^n_{a}}(v_l(a)) - L_l \right| \leq q|L_l|
\]

for \( n \geq N \) as long as \( f_a \) satisfies the CE-condition with some exponent at least \( \gamma_L \) and where \( N > 0 \) is as in Lemma 4.2. So combining (7.1) and (7.2) we get

\[
\left| \frac{\xi_{n, l}^\prime(a)}{\xi_{n, l}^\prime(b)} - 1 \right| \leq \tilde{\varepsilon}, \quad \text{for all } a, b \in \omega,
\]

where \( \tilde{\varepsilon} > 0 \) is arbitrarily small given that \( \varepsilon' \) and \( q \) are small enough. This means that the curve \( \xi_{n, l}(\omega) \) is almost straight, which will be important when we make partitions at returns.

From now on, let us fix \( l \) and write \( \xi_{n, l}(a) = \xi_n(a) \). In the beginning we are going to follow orbits close to the original orbit \( \xi_n(0) \), and then it is rather easy to see that nearby orbits also satisfy the CE-condition, but when considering nearby parameters \( a \) close to \( 0 \), after a long time we have to keep track of the derivative \( Df^n_{a}(v_l(a)) \), since the orbit of \( \xi_n(a) \) and \( \xi_n(0) \) become more or less independent.

Choose some small \( \varepsilon > 0 \) and suppose that \( \omega \subset (-\varepsilon, \varepsilon) \). For \( a, b \in \omega \), consider (7.1). The distortion during the first \( N \) iterates can be made arbitrarily small if the
perturbation $\varepsilon$ is small enough. So we only need to consider iterates after $N$, and hence focus on proving:

\begin{equation}
|Df_a^{n-N}(\xi_N(a)) - Df_b^{n-N}(\xi_N(b))| - 1 | \leq \varepsilon'.
\end{equation}

The main task is to prove this stronger form of the space distortion. By Lemma 6.2, the distortion estimate (7.4) follows if we prove that

\begin{equation}
\sum_{j=0}^{n-N-1} \frac{|Df_a(\xi_{N+j}(a)) - Df_b(\xi_{N+j}(b))|}{Df_b(\xi_{N+j}(b))} \leq \varepsilon'',
\end{equation}

where $\varepsilon' \to 0$ as $\varepsilon'' \to 0$.

With $z = \xi_j(a)$ and $w = \xi_j(b)$, we have

$$|Df_a(\xi_j(a)) - Df_b(\xi_j(b))| \leq |Df_a(z) - Df_a(w)| + |Df_a(w) - Df_b(w)|.$$  

We also see that, for some $a^* \in [a, b]$,

\begin{equation}
|Df_a(w) - Df_b(w)| \leq |a - b| |\partial_a Df_{a^*}(w)| \leq C|\xi_j(a) - \xi_j(b)| e^{-\gamma_{2j}},
\end{equation}

for some constant $C > 0$ since $\partial_a Df(z)$ is bounded.

If $c$ is a critical point, using that $|z - c| \sim |w - c|$, for $z = \xi_j(a)$, $w = \xi_j(b)$, we get, using the Taylor expansion of $f$ near $c$, see (6.5), that

\begin{equation}
\sum_{j=1}^n \frac{|Df_a(\xi_j(a)) - Df_b(\xi_j(b))|}{|Df_b(\xi_j(b))|} \sim 2k \sum_{j=1}^n \frac{|\xi_j(a) - \xi_j(b)|}{\text{dist}(\xi_j(b), \text{Crit}_b)}
\end{equation}

if $z$ and $w$ are sufficiently close to $\text{Crit}(f)$. We will therefore estimate the sum

\begin{equation}
\hat{S} = \sum_{j=N}^n \frac{|\xi_j(a) - \xi_j(b)|}{\text{dist}(\xi_j(b), \text{Crit}_b)}
\end{equation}

Lemma 7.1. If $N$ is large enough we have the following. Suppose that $\nu_k \geq N$ is a return time and that $\xi_{\nu_k}(a)$ is a free return into $U'$ (essential or inessential or a pseudo return), $a \in (-\varepsilon, \varepsilon)$. Moreover, we suppose that $a \in E_{\nu_k}(\gamma) \cap B_{\nu_k}$, where $\gamma \geq \gamma_1$. Then until the next free return, we have,

$$|Df_a^{\nu_k+1}(v_a)| \geq e^{\gamma_1 \nu_k+1},$$

where $\gamma_1 \geq (9/10) \min(\gamma, \gamma_H)$.

Proof. During the bound period $p_k$ starting directly after the return $\nu_k$, we see from Lemma 6.4 that

$$|Df^{\nu_k+p_k}(v_l(a))| \geq C_0 e^{\gamma_0 \nu_k} e^{\bar{\gamma} p_k},$$

for each $a \in \omega$. Moreover, note that $p_k \leq (2K \alpha / \gamma) \nu_k$ from Lemma 6.4. After that the free period starts, and by the outside expansion Lemma 3.1 we get

$$|Df^{\nu_k+1}(v_l(a))| \geq C_0 C'e^{\gamma_0 \nu_k} e^{\bar{\gamma} p_k} e^{\gamma_H(\nu_k+1-\nu_k+p_k)} \geq e^{\gamma_1 \nu_k+1},$$

for some $\gamma_1 \geq (9/10) \min(\gamma, \gamma_H)$ if $N$ is large enough.

If we consider a return of $\xi_n(\omega)$, where $\omega$ is a partition element, we have seen by Lemma 5.1, that we may disregard from the parameter dependence inside $\omega$ as long as the space derivative grows exponentially. The above lemma ensures that $\gamma_1 \geq \gamma_C$.

So, by the weak parameter dependence property, we have

$$|\xi_{\nu'}(a) - \xi_{\nu'}(b)| \sim_{\nu' - \nu} |Df_a^{\nu' - \nu}(\xi_{\nu'}(a))| |\xi_{\nu'}(a) - \xi_{\nu'}(b)|$$
for all $a, b \in \omega$. Since $|Df^{\nu - \nu'}(\xi_\nu(a))|Q^{-(\nu - \nu')}$ is much greater than 1 ($\log Q < \alpha \ll \gamma_1$), it follows that two orbits $\xi_{n,i}(a)$ and $\xi_{n,l}(b)$ repel each other up to some large scale or until the next return takes place. We get the following lemma.

**Lemma 7.2.** Suppose that $\xi_{\nu,i}(\omega)$ is a return (inessential or essential or a pseudo return) and that $a, b \in \omega$, $\omega \subset E_{\nu,i}(\gamma) \cap B_{\nu,i}$ is a partition element, and $\gamma \geq \gamma_I$. Then if $\nu'$ is the next free return time.

$$|\xi_{\nu'}(a) - \xi_{\nu'}(b)| \geq 2|\xi_{\nu}(a) - \xi_{\nu}(b)|.$$  

Next, we prove the Main Distortion Lemma, which is our main object in this section.

**Lemma 7.3** (Main Distortion Lemma). Let $\varepsilon' > 0$. Then if $N$ is sufficiently large we have the following. Let $\omega \subset E_{\nu}(\gamma) \cap B_{\nu,l}$ be a partition element for some $\gamma \geq \gamma_I$ and suppose that $\nu \geq \gamma$ is a return time or does not belong to a bound period. Then we have, until the next free return $\xi_{\nu',l}(\omega)$, a bound on the distortion if $\omega$ is still a partition element at time $n$, namely,

$$\left| \frac{Df_n^{\nu}(v_l(a))}{Df_n^{\nu}(v_l(b))} - 1 \right| \leq \varepsilon', \quad \text{for all } a, b \in \omega$$

if $\nu \leq n \leq \nu'$.

**Proof.** By Lemma 7.1 the CE-condition is fulfilled with exponent $\gamma_1 \geq (9/10) \min(\gamma, \gamma_H)$ up until the next free return. Now, $\gamma_1 \geq \gamma_C$ so we can repeatedly use the weak parameter dependence property. Let us assume that $\nu$ is a return time. If not, replace $\nu$ with the latest return time before $\nu$.

Put $\xi_{n,l}(a) = \xi_n(a)$. We want to estimate the sum

$$\sum_{j=1}^{n} \frac{|\xi_j(a) - \xi_j(b)|}{\dist(\xi_j(b), \Crit_b)}$$

(7.8)

First we look at the contribution from the bound periods. We want to estimate the sum

$$\sum_{j=0}^{p-1} \frac{|\xi_{\nu+j}(a) - \xi_{\nu+j}(b)|}{\dist(\xi_{\nu+j}(b), \Crit_b)}.$$  

Since $|\xi_{\nu}(a) - \xi_{\nu}(b)| \sim_2 e^{-r}/r^2$ and $\dist(\xi_{\nu}(b), \Crit_b) \sim \sqrt{r} e^{-r}$, the first term $(j = 0)$ contributes $\sim 1/r^2$.

To estimate the other terms $(j > 0)$, we use the weak parameter dependence property to get

$$|\xi_{\nu+j}(a) - \xi_{\nu+j}(b)| \sim_2 Df^j(\xi_{\nu}(a))|\xi_{\nu}(a) - \xi_{\nu}(b)| \sim_2 Df^j(\xi_{\nu}(a))|e^{-r}/r^2|.$$  

By the definition of the bound period we have, for $j > 0$, using Lemma 6.3, if $\xi_{\nu}(a) \in U_{\nu}'$,  

$$Df^j(\xi_{\nu}(a))|e^{-r} \sim |\xi_{j,i}(a) - \xi_{\nu+j}(a)| \leq e^{-\alpha j} \dist(\xi_{j,i}(a), \Crit_a).$$  

So we get

$$|\xi_{\nu+j}(a) - \xi_{\nu+j}(b)| \leq CQ^j \frac{e^{-\alpha j}}{r^2} \dist(\xi_{j,i}(a), \Crit_a),$$
and therefore, since \( \text{dist}(\xi_{j,i}(a), \text{Crit}_a) \) is virtually the same for all \( a \in \omega \),

\[
\sum_{j=0}^{p} \frac{|\xi_{j,i}(a) - \xi_{j,i}(b)|}{\text{dist}(\xi_{j,i}(b), \text{Crit}_b)} \leq \frac{C}{r^2} + C \sum_{j=1}^{p} \frac{Q^j e^{-\alpha_j}}{r^2} \leq \frac{2C}{r^2},
\]

where the term \( C/r^2 \) corresponds to \( j = 0 \), and \( \log Q < \alpha \).

Between each adjacent pair of free returns there is a growth of the interval \( \xi_{n,i}(\omega) \)

as follows. Lemma 7.2 implies

\[
2 \text{diam}(\xi_{n,i}(\omega)) \leq \text{diam}(\xi_{n,i+1}(\omega)), \quad \text{for all } a, b \in \omega.
\]

Let \( (r) \) be those indices \( k \) for which \( \text{dist}(\xi_{n,i}(\omega), \text{Crit}_\omega) \sim \sqrt{r} e^{-r} \), and let \( \tilde{k}(r) \) be the largest integer in \( (r) \). Hence going backwards in time, inside each \( (r) \), the contribution from the bound periods is a constant times the last contribution, i.e.

\[
\sum_{k \in (r)} \frac{|\xi_{n,i}(a) - \xi_{n,i}(b)|}{\text{dist}(\xi_{n,i}(b), \text{Crit}_b)} \leq C \frac{|\xi_{n,i}(a) - \xi_{n,i}(b)|}{\text{dist}(\xi_{n,i}(b), \text{Crit}_b)} \leq \frac{C}{r^2}.
\]

Summing over all such possible returns we get

\[
\sum_{r=\Delta}^{\infty} \frac{C}{r^2} \leq \frac{2C}{\Delta}.
\]

Let us now look for the contribution from the free periods. Let us assume that \( \nu_k \)

are the returns up until \( \nu_k = \nu' \), and \( p_k \) their bound periods. By Lemma 3.1 we get

that, for every \( a, b \in \omega \), now assuming that \( \xi_j(\omega) \cap U = \emptyset \) for all \( \nu_k + p_k + 1 \leq \nu_k + 1 - 1 \),

\[
|\xi_{\nu_k+1-1}(a) - \xi_{\nu_k+1-1}(b)| \geq C' \lambda^{\nu_k+1-1-j} |\xi_j(a) - \xi_j(b)|.
\]

Hence,

\[
\sum_{j=\nu_k-\Delta+1}^{\nu_k-1} \frac{|\xi_j(a) - \xi_j(b)|}{\text{dist}(\xi_j(b), \text{Crit}_b)} \leq C' \sum_{j=\nu_k-\Delta+1}^{\nu_k-1} \lambda^{-\nu_k+j} |\xi_{\nu_k-1}(a) - \xi_{\nu_k-1}(b)| \delta
\]

\[
\leq C \frac{|\xi_{\nu_k-1}(a) - \xi_{\nu_k-1}(b)|}{\delta}.
\]

We have, for some \( \kappa_2 \geq 1 \),

\[
|\xi_{\nu_k-1}(a) - \xi_{\nu_k-1}(b)| \sim \kappa_2 |\xi_{\nu_k}(a) - \xi_{\nu_k}(b)| \sim_2 e^{-\nu_k} r^2 / r_k^2,
\]

if \( k < s \), where we have put \( \text{dist}(\xi_{n,i}(\omega), \text{Crit}_\omega) \sim \sqrt{r} e^{-r} \). So for those returns the contribution to the sum (7.8) is going to be very small. Recalling that \( |\xi_j(a) - \xi_j(b)| \leq S \), where \( S = \varepsilon_1 \delta \) is the large scale, \( \delta = e^{-\Delta} \), we get, for the last return,

\[
\sum_{j=\nu_k-\Delta+1}^{\nu_k-1} \frac{|\xi_j(a) - \xi_j(b)|}{\text{dist}(\xi_j(b), \text{Crit}_b)} \leq C S \frac{1}{\delta} \leq C \varepsilon_1,
\]

where \( C \) depends only on \( C' \) and \( \lambda \) (hence not on \( \delta \)). So \( C \varepsilon_1 \) can be made arbitrarily small if \( \varepsilon_1 \) is small enough. We let \( (r) \) be those indices \( k \) such that \( \text{dist}(\xi_{n,i}(\omega), \text{Crit}_\omega) \sim \sqrt{r} e^{-r} \), and \( \tilde{k}(r) \) the maximum index \( k \) for which this happens. Then using Lemma 7.2, we have (7.9), and therefore we conclude that

\[
\sum_{k \in (r)} |\xi_{\nu_k}(a) - \xi_{\nu_k}(b)| \leq C |\xi_{\nu_k}(a) - \xi_{\nu_k}(b)|.
\]
Summing up, we get, excluding the last return,

\[
\sum_{k=1}^{s-1} \sum_{j=\nu_{k-1}+p_{k-1}+1}^{\nu_k-1} \frac{|\xi_j(a) - \xi_j(b)|}{\text{dist}(\xi_j(b), \text{Crit}_b)} = \sum_{r \geq \Delta} \sum_{k \in (r)} \sum_{j=\nu_{k-1}+p_{k-1}+1}^{\nu_k-1} \frac{|\xi_j(a) - \xi_j(b)|}{\text{dist}(\xi_j(b), \text{Crit}_b)} \\
\leq C \sum_{r \geq \Delta} \sum_{k \in (r)} \frac{|\xi_{\nu_k-1}(a) - \xi_{\nu_k-1}(b)|}{\delta} \\
\leq C \sum_{r \geq \Delta} \frac{|\xi_{\nu_{k(r)-1}}(a) - \xi_{\nu_{k(r)-1}}(b)|}{\delta} \\
(7.12) \\
\leq C \sum_{r \geq \Delta} \frac{\Delta-r}{r^2} \leq \frac{C}{\Delta}.
\]

Including the last return we get

\[
(7.13) \sum_{k=1}^{s} \sum_{j=\nu_{k-1}+p_{k-1}+1}^{\nu_k-1} \frac{|\xi_j(a) - \xi_j(b)|}{\text{dist}(\xi_j(b), \text{Crit}_b)} \leq \frac{C}{\Delta} + C\varepsilon_1.
\]

If we now pick some \( n \) such that \( \nu + p \leq n < \nu' \), then letting \( q_1 < \ldots < q_l \) be consecutive, so called pseudo-returns, we get

\[
(7.14) \sum_{j=\nu+p+1}^{q_l} \frac{|\xi_j(a) - \xi_j(b)|}{\text{dist}(\xi_j(b), \text{Crit}_b)} = \sum_{j=\nu+p+1}^{q_1} \frac{|\xi_j(a) - \xi_j(b)|}{\text{dist}(\xi_j(b), \text{Crit}_b)} \\
+ \sum_{r=\Delta'} \sum_{l \geq (r), l > 0} \sum_{j=q_{l-1}+1}^{q_l} \frac{|\xi_j(a) - \xi_j(b)|}{\text{dist}(\xi_j(b), \text{Crit}_b)} \\
\leq C \sum_{r=\Delta'} \sum_{l \geq (r)} \frac{|\xi_{q_l}(a) - \xi_{q_l}(b)|}{\text{dist}(\xi_{q_l}(b), \text{Crit}_b)} \\
\leq C \sum_{r=\Delta'} \frac{|\xi_{q_l(r)}(a) - \xi_{q_l(r)}(b)|}{\text{dist}(\xi_{q_l(r)}(b), \text{Crit}_b)}.
\]

Moreover, we have the assumption that \( \text{diam}(\xi_k(\omega)) \leq S = \varepsilon_1 \delta \), for all \( k \leq n \). If \( \xi_{q_l}(\omega) \) is a pseudo return with \( \text{dist}(\xi_{q_l}(\omega), \text{Crit}_\omega) \sim \sqrt{e} e^{-r_l} \), for \( \Delta' \leq r_l \leq \Delta \), the contribution will be simply bounded by \( \varepsilon_1 e^{-\Delta'/e^{-r_l}} \). We get

\[
(7.15) C \sum_{r=\Delta'}^{\Delta} \frac{|\xi_{q_l(r)}(a) - \xi_{q_l(r)}(b)|}{\text{dist}(\xi_{q_l(r)}(b), \text{Crit}_b)} \leq C \sum_{r=\Delta'}^{\Delta} \varepsilon_1 e^{r-\Delta} \leq C\varepsilon_1.
\]
The contribution from the very last iterates from \( q_l < j \leq n \) is a constant (depending on the large scale) by the uniform expansion along the early orbit (the bound period) and then outside \( U' \). Summing up,

\[
\sum_{j=1}^{n} \frac{|\xi_j(a) - \xi_j(b)|}{\text{dist}(\xi_j(b), \text{Crit}_b)} \leq \frac{2C}{\Delta} + \frac{C}{\Delta} + 2C\varepsilon_1,
\]

which can be made small if \( \varepsilon_1 \) and \( \delta \) are small enough. This finishes the lemma. \( \square \)

We now get a posteriori that \( \xi_n \) is almost affine on each partition element \( \omega \). Hence also \( |\xi_n(a) - \xi_n(b)| \) expands according to the space derivative for any parameter \( c \in [a, b] \) i.e.

\[
|\xi_n(a) - \xi_n(b)| \sim_C |Df_n^{n-j}(\xi_j(c))||\xi_j(a) - \xi_j(b)|,
\]

for \( C > 1 \) close to 1. This is called strong distortion.

Moreover, we see that as long as \( \gamma \geq \gamma_l \) for returns (in general \( \gamma \geq \gamma_l - 4aK \)), we have good geometry control, i.e. for a partition element \( \omega \), we have (7.3), for all \( a, b \in \omega \).

### 7.1. Initial distortion.

As a direct consequence of the Main distortion lemma, we here state that for any sufficiently small \( \varepsilon \) we can find an interval \( \omega \subset (-\varepsilon, \varepsilon) \) such that \( \xi_n,1(\omega) \) grows to some “large scale” (denoted by \( S \)) or returns into \( U \) as an essential first return.

**Lemma 7.4** (Start-lemma). Let \( f = f_0 \) be as in Theorem A and let \( \varepsilon' > 0 \) and \( N > 0 \) from Lemma 4.2. There is a neighbourhood \( U \) of \( \text{Crit}_0 \) and a number \( S > 0 \) (called the “large scale”), which depends on \( U \) such that the following holds. For every sufficiently small \( \varepsilon > 0 \) and each critical point \( c_l \) there is some \( N_l \geq N > 0 \) such that for every \( a \in \omega = (-\varepsilon, \varepsilon) \) we have:

i) For some \( \gamma_l \geq \gamma_0(1 - \varepsilon') \), it holds that

\[
|Df_n^k(f_a(c_l(a)))| \geq Ce^{\gamma_l k}, \text{ for all } k \leq N_l,
\]

ii) for all \( k \leq N_l - 1 \), it holds that

\[
\text{diam}(\xi_{k,l}(\omega)) \leq \begin{cases} \frac{\text{dist}(\xi_{k,l}(\omega), \text{Crit}_l)}{\text{dist}(\xi_{k,l}(\omega), \text{Crit}_l)}^\alpha, & \text{if } \xi_{k,l}(\omega) \cap U \neq \emptyset, \\ S, & \text{if } \xi_{k,l}(\omega) \cap U = \emptyset, \end{cases}
\]

iii) for \( k = N_l \), it holds that

\[
\text{diam}(\xi_{N_l}(\omega)) \geq \begin{cases} \frac{\text{dist}(\xi_{N_l}(\omega), \text{Crit}_l)}{\text{dist}(\xi_{N_l}(\omega), \text{Crit}_l)}^\alpha, & \text{if } \xi_{N_l}(\omega) \cap U \neq \emptyset, \\ S, & \text{if } \xi_{N_l}(\omega) \cap U = \emptyset, \end{cases}
\]

iv) and finally, for all \( a, b \in \omega \) it holds that

\[
\left|\frac{Df_{a}^{n-N}(\xi_N(a))}{Df_{b}^{n-N}(\xi_N(b))} - 1\right| \leq \varepsilon', \text{ for all } n \leq N_l.
\]

**Remark 7.5.** The Whitney type of condition on the diameter of \( \xi_n(\omega) \) and its distance to the critical points has the following meaning. With \( \text{dist}(\xi_{n,l}(\omega), \text{Crit}_l) \sim e^{-r} \) the diameter becomes \( \sim e^{-r}/r^2 \) and this is sufficient for having control of the distortion of the derivative. The condition is also used in the main distortion lemma later.
Proof. Similar to the proof of Lemma 5.1, we can easily conclude that condition ii) implies that we have very small distortion for $Df_c$, $c \in (-\varepsilon, \varepsilon)$ both in the space and parameter variable. We see that if $\xi_n(a)$ and $\xi_n(b)$ are close in this sense, then for all $a, b \in (-\varepsilon, \varepsilon)$, in particular for $b = 0$, the bounded distortion on $Df_c$ implies

$$|Df^n_{a}(v_l(a))| \geq C^{-n}|Df^n_{a}(v_l(0))| \geq C_0 e^{-\gamma n},$$

for some $\gamma_1$ slightly smaller than $\gamma_0$ (we may assume that $\gamma_1 \geq (1 - \varepsilon')\gamma_0$, if $C$ is close enough to 1 (by choosing $S$ small enough). So $f_a$ also satisfies the CE-condition with exponent slightly smaller than $\gamma_0$. We now let $N_l$ be the maximal integer such that ii) holds. We have shown that also i) holds for this $N_l$. We can hence use the Main distortion lemma for all following returns until time $N_l$. Hence iv) holds. \]

7.2. The partition. If $J(f) = \hat{C}$ then we have $2d - 2$ critical points, counting multiplicity. Inside the subspace $\Lambda_{d,T}$ each critical point moves analytically. So Lemma 7.4 gives at most $2d - 2$ numbers $N_l$, given an interval $\omega_0 = (-\varepsilon, \varepsilon)$, such that $\xi_\omega\hat{-1}((\omega_0))$ has grown to some large scale $S$ (same for all $l$), or has reached size $e^{-r}/r^2$ inside $U$, where $e^{-r}$ is, more or less, the distance to the critical points, i.e. $\text{dist}(\xi\hat{-1}_\omega(\omega_0), J\text{rit}_\omega) \sim e^{-r}$. We now assume that, without loss of generality, $N_1 = \min(N_l)$. Thus we have the CE-condition satisfied for all critical points up until time $N_1$, on $\omega_0$.

If $N_1$ is not a return time, we have $\text{diam}(\xi_{N_1,1}(\omega_0)) \geq S$ by Lemma 7.4. As soon as this happens, we partition the interval $\omega_0$ into the least number of smaller sub-intervals $\omega^0_0 \subset \omega$ of equal length such that $\text{diam}(\xi_{N_1,1}(\omega^0_0)) \leq S$. We call the sets $\omega^0_0$ of this type partition elements. We do this partitioning for every critical point at all times outside $U$ until some parameter returns into $U$. In this way we always have $\text{diam}(\xi\hat{-1}(\omega)) \leq S$ for any partition element $\omega$ and study the evolution of each such $\omega$ separately. We will use $\omega \subset \omega_0 = (-\varepsilon, \varepsilon)$ as a standard notion for partition elements in the future.

Let us go back to the critical point $c_1$ ($l = 1$) and assume that $\omega \subset \omega_0$ is such partition element and that $m_1$ is the smallest integer $m_1 \geq N_1$ such that $\xi_{m_1,1}(\omega) \cap U \neq \emptyset$, i.e. $\xi_{m_1,1}(\omega)$ is a return into $U$. If

$$\frac{1}{2} \frac{\text{dist}(\xi_{m_1,1}(\omega), J\text{rit}_\omega)}{\text{log}(\text{dist}(\xi_{m_1,1}(\omega), J\text{rit}_\omega)))^2} \leq \text{diam}(\xi_{m_1,1}(\omega)),$$

we speak of an essential return. Otherwise the return is inessential. For essential returns we then partition the interval $\omega$ into smaller intervals $\omega^0_{m_1} \subset \omega$ such that

$$\frac{1}{2} \frac{\text{dist}(\xi_{m_1,1}(\omega^0_{m_1}), J\text{rit}_{\omega^0_{m_1})}}{\text{log}(\text{dist}(\xi_{m_1,1}(\omega^0_{m_1}), J\text{rit}_{\omega^0_{m_1}})))^2} \leq \text{diam}(\xi_{m_1,1}(\omega^0_{m_1}))$$

$$\leq \frac{\text{dist}(\xi_{m_1,1}(\omega^0_{m_1}), J\text{rit}_{\omega^0_{m_1})}}{\text{log}(\text{dist}(\xi_{m_1,1}(\omega^0_{m_1}), J\text{rit}_{\omega^0_{m_1}})))^2}.$$ 

These smaller intervals $\omega^0_{m_1}$ are also called partition elements (at time $m_1$). The condition (7.16) implies that we have control of the distortion:

$$\frac{|Df_a(\xi_{m_1,1}(a))|}{|Df_b(\xi_{m_1,1}(b))|} \leq C \hat{r},$$

for all $a, b \in \omega^0_{m_1}$. 


where $C(\hat{r}) > 1$ and tends to 1 as $\hat{r} = -\log(\text{dist}(\xi_{m_1,1}(\omega^i_{m_1}), J_{\text{rit}}\omega^i_{m_1}))$ tends to infinity. We let $r = [\hat{r} - 1/2]$. For $\omega^i_{m_1}$ above, we associate $r = r(\hat{r})$ to $\hat{r}$, and it follows that
\[
\text{dist}(\xi_{m_1,1}(\omega^i_{m_1}), J_{\text{rit}}\omega^i_{m_1}) \sim e^{-r},
\]
(cf. with the annular neighbourhoods in [1]). Moreover, we see that
\[
\text{diam}(\xi_{m_1,1}(\omega^i_{m_1})) \sim 2 e^{-r/\ell^2},
\]
if $r \geq \Delta$, and $\Delta$ sufficiently large. When we write $\text{dist}(A, B) \sim \sqrt{e} e^{-r}$, typically we use it when $A = \xi_{n,l}(\omega)$ and $B = \text{Crit}_{\omega}$, then we mean the unique $r$ such that $r = [\log \text{dist}(A, B) - 1/2]$, i.e. $\text{dist}(A, B) \in [e^{-r-1/2}, e^{-r+1/2})$.

For each return, and in particular this first return, we partition parameter intervals according to the above rule. Moreover, we delete parameters not satisfying the basic assumption and show later that the Lebesgue measure of the set deleted is a small portion of the total interval returning into $U$. It is quite easy to see that this is the case for the first return. Because of the slow recurrence condition, we see that
\[
e^{-r} \geq e^{-\alpha m_1} \gg e^{-2\alpha m_1}.
\]
Hence, the basic assumption possibly forces us to delete a small fraction of parameters at time $m_1$.

After this return the first bound period starts, and the whole idea is that binding the old orbit to the early orbit of possibly another critical point, will, via distortion control, transfer the derivative gain from the early orbit to the old orbit. To do this we need to be able to use the binding time for all critical points in the induction. We continue like this as long as we can use the binding information for all critical points, up until time $N_1$. This procedure creates a Cantor-like set (denoted by $\Omega(\gamma, \text{rit})$) of “good” parameters, for each critical point $c_l$, that do satisfy the basic assumption up until some time $m$, which turns out to be much larger than $N_1$, because the bound periods for a return $\xi_{m,l}(\omega)$ into $U$ are much smaller than $m$ itself (Lemma 6.4).

At this point, we have to delete more parameters such that the binding period can be used longer. A potential problem here is that different critical points $c_l$ may produce different Cantor-like sets up until time $m$, and if we take intersections of these sets, we may destroy the partition elements. But the idea is that the partition elements at time from, say $N_1$ until $2N_1$, are much larger that those partition elements formed around time $m \gg N_1$. We develop this idea, which is due to M. Benedicks, later.

In the construction, the growth of the derivative along critical orbits is never allowed to go below a certain level, in order to have the whole machine working. Recall that $\gamma_B = (3/4) \min(\gamma_0, \gamma_H)(1 - \tau)$, where $0 < \tau < 1$. This exponent $\gamma_B$ should be thought of the desired Lyapunov exponent, which we will get at the end. It will also be used as an induction assumption. The number $\tau$ can be chosen freely but $\delta$ depends on it (see Section 8, Lemma 8.7). The intermediate Lyapunov exponent $\gamma_{\ell} = (1/3) \min(\gamma_0, \gamma_H)(1 - \tau) < \gamma_B/2$ will be an assumption in most lemmas.

8. LARGE DEVIATIONS

We will make an induction over time intervals of the type $[n, 2n]$ and assume from now on that we make partitions as described above. Given a good situation at time $n$ with growth of the derivative, we first delete the parameters not satisfying the basic assumption up until time $2n$. But according to Lemma 7.1, this means that
we may lose some part of the Lyapunov exponent. Therefore we make use of the famous large deviation argument, developed by Benedicks and Carleson, to restore the Lyapunov exponent up until time $2n$.

This section is very similar to older papers [5], [1] et al.

**Lemma 8.1.** Suppose that $\xi_{n,l}(ω)$ is an essential return into $U_i$, and that the Lyapunov exponent $γ ≥ γ_l$ for all critical points, $ω ∈ ξ_{n,l}(γ) ∩ B_{n,l}$. Then if $ν'$ is the next return time, we have that the set $\hat{ω}$ of parameters in $ω$ that satisfies the basic assumption, has Lebesgue measure

$$m(\hat{ω}) ≥ (1 − e^{-αν'})m(ω).$$

**Proof.** This follows quite easily, since the interval $ξ_{ν+p}(ω)$ grows rapidly during the bound period $p$. By Lemma 6.4, Lemma 6.3, Lemma 3.1, and Lemma 7.3, we get, for any $a ∈ ω$,

$$\text{diam}(ξ_{ν+p+1,l}(ω)) ≈ e^{−αν}Df(ξ_{ν+p+1,l}(a)) \sim \frac{||ξ_{ν+p+1,l}(a) − ξ_{p+1,l}(a)||}{p} ≥ Ce^{−(p+1)}\cdot \text{dist}(ξ_{p+1,l}(a), \text{Crit})$$

$$≥ CK_0 e^{−2α(p+1)−2α(p+1)} ≥ e^{−/(2)(p+1)2log r},$$

(8.1)

if $p$ is large. So,

$$\text{diam}(ξ_{ν,l}(ω)) ≥ \text{diam}(ξ_{ν+p,l}(ω)) ≥ C' e^{−(2/2)(p+1)2log r} ≥ e^{−7αdr/γ−2log r} ≥ e^{−αν'},$$

Recalling the distortion control from the main distortion Lemma 7.3 together with Lemma 4.2, we see that the measure of parameters deleted at $ν'$ is

$$\frac{|ω| − |\hat{ω}|}{|ω|} ≤ 2 \frac{e^{−2αν'}}{\text{diam}(ξ_{ν,l}(ω))} ≤ 2e^{−α(2−(ν)/2)} ≤ e^{−αν'},$$

since $αK/γ ≤ 1/100$, $K = \max(d)$ (maximal degree of the critical points). □

We now define escape time and escape situation. Let $U^2$ be a neighbourhood of $\text{Crit}(f_0)$ such that $U^2 = \cup_j B(ξ_j, δ^2) ⊂ U$. We say that a deep return is characterised by $ξ_{n,l}(ω) ∩ U^2 ≠ ∅$ and a shallow return means that $ξ_{n,l}(ω) ∩ U^2 = ∅$ but $ξ_{n,l}(ω) ∩ U ≠ ∅$. We then speak of deep returns into $U^2$ and shallow returns into $U \setminus U^2$ even if the actual curve $ξ_{n,l}(ω)$ does not entirely lay inside $U^2$ or $U \setminus U^2$ respectively. We also let $ω_n(a)$ be the corresponding partition element following the parameter $a$, i.e. the unique $ω$ such that $ξ_n(ω)$ has diameter bounded by $S$ if $ξ_n(ω) ∩ U = ∅$ and bounded by $\text{dist}(ξ_n(ω), \text{Crit}_ω)/(\log \text{dist}(ξ_n(ω), \text{Crit}_ω))$ if $ξ_n(ω) ∩ U ≠ ∅$.

**Definition 8.2.** We say that $ξ_n(ω)$ has escaped, or is in escape position, if $\text{diam}(ξ_n(ω)) ≥ S$, and the bound period has passed.

The escape time for a parameter $a ∈ ω$ for a deep return $ξ_{ν,l}(ω)$ into $U^2$ is defined as the least number $n − (ν + p) ≥ 0$ (where $p$ is the bound period for the return) such that $ξ_{n,l}(ω_n(a))$ has reached escape position. We write $E_l(a, ν) = n − (ν + p)$ for this escape time. We also define the escape time for shallow returns, i.e. if $ξ_{ν,l}(ω) ⊂ U \setminus U^2$, to be equal to zero.
Lemma 8.3. Suppose that $\xi_{\nu,i}(\omega)$ is an essential return into $U_i$, $\omega \in E_{\nu,i}(\gamma) \cap B_{\nu,i}$, $\gamma \geq \gamma_I$ and that $\text{dist}(\xi_{\nu,i}(\omega), \text{Crit}_i) \sim \sqrt{r} e^{-r}$. Put $h = 4K^2/\gamma_I$. Then if $q = n - (\nu + p)$ where $n$ is the next essential return or the time when $\xi_{\nu,i}(\omega)$ is in escape position, which ever comes first, we have the estimate

$$q \leq hr.\]$$

Proof. Let us put $D_j = |Df^j(\xi_{\nu,i}(a))|$, for $a \in \omega$. By the definition of the bound period, the basic assumption, and Lemma 6.4, for all $a \in \omega$,

$$(8.2) \quad D_{p+1} \geq Ce^{-\alpha(p+1)} \text{dist}(\xi_{p+1}(a), \text{Crit}_i) e^p$$

$$\geq e^{-2\alpha(p+1)-\alpha(p+1)+r} \geq e^{(1-2\alpha K/\gamma_I)} \geq e^{(1-2\alpha K/\gamma_I)}.$$  

Let $m_j$ be the inessential returns after $\nu$, i.e. $\nu < m_1 < m_2 < \ldots < m_s < n$. Let $p_j$ and $q_j$ be the bound and free periods respectively following $m_j$. Let $p_0$ and $q_0$ be the bound and free periods following the return $\nu$. It can happen that escape takes place before a return takes place, and then $q_s$ is not a complete free period. It can also happen that $n$ is a time during the bound period for $m_s$. But then we have $m_s + p_s = \nu$ as an upper bound for $q$ and we can assume that $q > m_s + p_s$.

Suppose that $\text{dist}(\xi_{m_j}(\omega), \text{Crit}_i) \sim \sqrt{r} e^{-r}$, and let $r = r_0$. Suppose that $n = \nu'$ is a return. Then, as long as the bound period is bounded by $(2K\alpha/\gamma_I)\nu$, we can use the same estimate as (8.2), and Lemma 3.1, to obtain

$$\text{diam}(\xi_\omega(\omega)) \sim |Df^0(\xi_{\nu}(a))| \text{diam}(\xi_{\nu}(\omega))$$

$$\geq \prod_{j=0}^{s} |Df^{p_j}(\xi_{m_j}(a))| C' e^{\gamma_H q_0} \text{diam}(\xi_{\nu}(\omega))$$

$$\geq e^{r(1-2\alpha K/\gamma_I)} C' e^{\gamma_H q_0} \text{diam}(\xi_{\nu}(\omega)) \prod_{j=1}^{s} e^{r_j(1-2\alpha K/\gamma_I)} \prod_{j=1}^{s} C' e^{q_j(1+\gamma_H)}$$

$$(8.3) \quad \geq e^{-r_{\delta K/\gamma_I} + q_0\gamma_H} \prod_{j=1}^{s} e^{r_j(1-2\alpha K/\gamma_I) + q_j\gamma_H}.$$  

If $n$ was not a return, then let $q_1 < \ldots < q_l$ be the pseudo-returns after $m_s + p_s$. Between each pair of pseudo-returns we have uniform expansion of the derivative according to Lemma 7.1. Between $m_s + p_s$ and $q_1$ we also have uniform expansion according to Lemma 3.1. So we only need to consider the last time period, from $q_l$ to $n$. Since $\xi_{q_l}(\omega)$ may belong to $U' \setminus U$ we have $|Df_a(\xi_{q_l}(a))| \geq e^{-K\Delta}$ for all $a \in \omega$. After time $q_l$ we can use the binding information, Lemma 6.3 and the first statement of Lemma 3.1 with $U = U'$, depending on whether $n$ belongs to the bound period or not. In any case we get uniform expansion: $|Df^{n-q_1}(\xi_{q_1+1}(a))| \geq Ce^{\min(\gamma_H)/(n-q_1-1)}.
In other words, with \( z = \xi_{m_s+p_s}(a) \), for \( a \in \omega \),
\[
|Df_a^{-(m_s+p_s)}(z)| = |Df_a^{q_1-(m_s+p_s)}(z)||Df_a^{q_1-q_2}(f_a^{-(m_s+p_s)}(z))| \\
\ldots |Df_a^{q_1-2q_1-1}(f_a^{-(m_s+p_s)}(z))||Df_a(f_a^{-(m_s+p_s)}(z))| \\
\geq C'e^{\gamma I(q_1-(m_s+p_s))}e^{\gamma I(q_1-q_2)}e^{-K\Delta}Ce^{\min(\gamma I(n-q_1-1)} \\
(8.4)
\]
since \( \gamma_1 \geq (9/10)\gamma_I \geq \gamma_C \). So we may have to replace \( q_s\gamma_I \) with \( \gamma_Cq_s - K\Delta \) in (8.3), where \( q_s = (n - (m_s + p_s)) \) in this case.

Since \( \gamma_C < \gamma_I \), and \( \text{diam}(\xi_n(\omega)) \) is assumed to be at most \( S = \varepsilon_I \delta \leq 1 \), we therefore get
\[
\sum_{j=0}^s r_j(1 - \frac{8\alpha K}{\gamma_I}) + \sum_{j=0}^s q_j\gamma_C \leq r_1 + K\Delta.
\]
Hence, if \( q = \sum_{j=1}^s p_j + \sum_{j=0}^s q_j \), we get
\[
q = \sum_{j=1}^s p_j + \sum_{j=0}^s q_j \leq \sum_{j=1}^s \frac{2K}{\gamma_I}r_j + \sum_{j=0}^s q_j \gamma_C \\
= \sum_{j=0}^s \frac{4K}{\gamma_I}(1 - \frac{8\alpha K}{\gamma_I})r_j \gamma_C + \sum_{j=0}^s q_j\gamma_C \\
\leq \max\left(\frac{4K}{\gamma_I}, \frac{1}{\gamma_C}\right) \left(\sum_{j=1}^s r_j(1 - \frac{8\alpha K}{\gamma_I}) + \sum_{j=0}^s q_j\gamma_C\right) \\
\leq \max\left(\frac{4K}{\gamma_I}, \frac{1}{\gamma_C}\right) \left(\frac{8\alpha K}{\gamma_I} r_1 + K\Delta\right) \leq \frac{4K^2}{\gamma_I}r,
\]
(8.5)
since \( 4K/\gamma_I < 1/\gamma_C \), and where we also used that \( \alpha \leq 3\gamma_I/(400K) \). Since \( Kr/\gamma_I \leq (2K\alpha/\gamma_I)\nu \), we can use the binding information the whole time. □

We will now estimate the measure of the set of parameters having a specific history for the returns in a time window of the form \([n, 2n]\). For simplicity, suppose that \( \xi_\nu(\omega_0) \) is an essential return with \( \text{dist}(\xi_\nu(\omega_0), \text{Crit}_{\omega_0}) \sim \sqrt{e^{\gamma_0}} \) and \( \nu \geq n (\nu \text{ should be thought of as the smallest return time after } n) \). Let us study the evolution of \( \xi_m(\omega_m(a)) \) as \( m \) goes through a sequence of essential returns \( \nu_1, \nu_2, \ldots, \nu_n \leq 2n \). Let us also assume that \( \omega_{\nu_1}(a) \subset \mathcal{E}_{\nu_1} \cap \mathcal{B}_{\nu_1} \), for these returns so that we can use the binding information of all other critical points up to time \( 2n \). This is not a strong assumption, as we now explain. Suppose \( a \in \mathcal{E}_{n,1}(\gamma_B) \cap \mathcal{B}_{2n,1} \), i.e. we assume that the basic approach rate condition is fulfilled up until time \( 2n \). The Lyapunov exponent will not drop too much at each return in the interval \([n, 2n]\), because we can use Lemma 6.4 at each return and get a trivial lower bound for the expansion, namely 1 during the bound period. But this means that the actual Lyapunov exponent is bounded from below, and we get a trivial bound,
\[
|Df^{2n}(\nu_1(a))| \geq e^{\gamma_B n} = e^{2n\gamma_I},
\]
since \( \gamma_1 \leq \gamma_B/2 \). In other words, \( a \in \mathcal{E}_{2n,1}(\gamma_B) \cap \mathcal{B}_{2n,1} \).

By the Main Distortion Lemma 7.3, which then gives good geometry control, the diameter of \( \xi_{\nu_1+p_j}(\omega_{\nu_1+p_j}(a)) \) is more or less equal to the length of the curve (which
is then more or less straight), i.e. \( \sim e^{-\frac{1}{2}(7K\alpha/\gamma)r_j} \), see inequality (8.2). After the free period it may expand further, and to get rid of the constant \( C' \) in Lemma 3.1, we may say that the curve \( \xi_{\nu_{j+1}}(\omega_{\nu_{j+1}}(a)) \) has a diameter at least \( e^{-\frac{1}{2}(8K\alpha/\gamma)r_j} \). We therefore get, with \( \gamma \geq \gamma_I \), that the measure of those parameters \( b \in \omega_{\nu_j}(a) \) entering into \( U \) with \( \text{dist}(\xi_{\nu_j+1}(b), \text{Crit}_b) \sim e^{-r_{j+1}} \) is

\[
(8.6) \quad m(\omega_{\nu_{j+1}}(a)) = m(\{b \in \omega_{\nu_j}(a) : \xi_{\nu_j+1}(b) \sim e^{-r_{j+1}}\}) \leq C e^{-r_{j+1}} e^{-8K\alpha/\gamma r_j} m(\omega_{\nu_j}(a)),
\]

(recall that we do not partition \( \omega_{\nu_j}(a) \) until the next return, so \( \omega_{\nu_j}(a) = \omega_{\nu_j+1}(a) \)).

So suppose now that we have a sequence of \( s \) essential returns \( \nu_1, \nu_2, \ldots, \nu_s \leq 2n \). Let us also assume that we always have a lower bound, \( \gamma_I \), for the Lyapunov exponent, i.e. \( a \in E_{2n,l}(\gamma_I) \cup B_{2n,l} \) for the parameters we are considering. Then the portion from the starting interval, call it \( \omega_0 = \omega_{\nu_j}(a) \) for some \( a \in \omega_0 \), that has this specific history is, with \( \omega_j = \omega_{\nu_j}(a) \),

\[
(8.7) \quad \frac{m(\omega_s)}{m(\omega_0)} = \prod_{j=0}^{s-1} \frac{m(\omega_{j+1})}{m(\omega_j)} \leq C^s \prod_{j=0}^{s-1} e^{-r_{j+1}} e^{-8K\alpha/\gamma r_j}.
\]

We continue to follow [1] and [5] more or less verbatim. Let \( R = r_1 + r_2 + \ldots + r_s \). We now compute the number of combinations of choosing such \( r_j \) given that \( r_j \geq \Delta \geq 0 \). Let us not yet take into account that we are partitioning the intervals into smaller intervals such that

\[
(8.8) \quad \text{diam}(\xi_{\nu_j}(\omega)) \sim e^{-r_j}/r_j^2, \text{ for each } j = 1, \ldots, s,
\]

where \( \omega = \omega_{\nu_j}(a) \). Hence for each such set we have another \( r_j^2 \) possibilities.

By the pigeonhole principle, this number of combinations is, disregarding from these extra \( r_j^2 \) possibilities,

\[
\binom{R + s - 1}{s - 1}.
\]

By Stirling’s formula this can be estimated as follows, using that \( R \geq s\Delta \),

\[
\binom{R + s - 1}{s - 1} \leq C \frac{1}{\sqrt{2\pi}} \frac{\sqrt{R^s}}{\sqrt{R^R}} e^{-R} e^{-s} e^{-s} \frac{R + s - 1}{s^s} \leq e^{R + \frac{R}{\Delta}(1 + \frac{1}{\Delta})^{1+\frac{1}{\Delta}}} e^{-2s} \leq 2(1 + \eta(\Delta))^R
\]

if \( \Delta \) is large enough, where \( \eta(\Delta) = O(1/\Delta) \).

Taking into account now (8.8), we get that the number of combinations is

\[
2(1 + \eta(\Delta))^R \prod_{j=1}^s r_j^2 \leq e^{R/32} (1 + \eta(\Delta))^R.
\]

We can rewrite (8.7) to get, (recall the condition on \( \alpha \),

\[
\frac{m(\omega_s)}{m(\omega_0)} = C^s e^{\alpha r_0(8aK/\gamma) - \sum_{j=1}^s r_j(1 - 8aK/\gamma) - r_s} \leq C^s e^{\alpha r_0(8aK/\gamma) - (15/16)R}.
\]
Given an essential return $\xi_{\nu,l}(\omega)$, let $A_{s,R} \subset \omega$ be the set of those parameters having exactly $s$ essential returns as above before escaping at the $s + 1$:st return, for a fixed $R$. Each pair of sequences $\{v_j\}_{j=1}^{s}$, $\{r_j\}_{j=1}^{s}$ defines a unique history for a parameter $a \in A_{s,R}$. Letting $s$ and $R$ vary, then $\omega$ gets partitioned into a (likely huge) number of smaller intervals having this specific history. But let us fix $s$ and let $\hat{\omega}_s$ be the largest of these partition intervals for this fixed $s$. Then

$$|A_{s,R}| \leq |\hat{\omega}_s|e^{R/32(1 + \eta(\Delta))R}.$$ 

Now we show that the set of those parameters for which $\xi_{\nu,l}(a)$ returns too frequently and too deep into $U$ has very small Lebesgue measure. This is handled via so famous large deviation argument, originally developed in [5], which is an idea from a probabilistic point of view, and although the system we are considering is deterministic.

For an essential return $\xi_{\nu,l}(\omega)$ into $U^2$ where $\text{dist}(\xi_{\nu,l}(\omega), \text{Crit}_\omega) \sim \sqrt{\frac{e}{\pi e}} e^{-r}$, suppose that $a \in \omega$ has $s$ essential returns before it has escaped. Then according to Lemma 8.3, we have

$$E_l(a, \nu) \leq \sum_{j=0}^{s} hr_j = hr + hR,$$

where $R = r_1 + \ldots + r_s$. So the escape time $t \leq hr + hR$, i.e. it is bounded in terms of how deep the returns are. Let us estimate the measure of those parameters that escape at a certain (long) time $t$.

Put $r = r_0$. We get, given that $\Delta$ is large enough,

$$m(\{a \in \omega : E_l(a, \nu) = t\}) \leq \sum_{R \leq t/h - r_0, s \leq R/\Delta} |A_{s,R}|$$

$$\leq \sum_{R \leq t/h - r_0, s \leq R/\Delta} |\hat{\omega}_s|e^{R/32(1 + \eta(\Delta))R}$$

$$\leq |\omega| \sum_{R=t/h - r_0}^{\infty} \sum_{s=1}^{R/\Delta} e^{R/32(1 + \eta(\Delta))R} C^s e^{r_0(8\alpha K/\gamma - (15/16)R)}$$

$$\leq C'|\omega| \sum_{R=t/h - r_0}^{\infty} C^R/\Delta e^{-R(\frac{29}{32} - \eta(\Delta))} + (8\alpha K/\gamma)r_0$$

$$\leq C'|\omega| e^{-\frac{t}{h} + \frac{28}{32} + (8\alpha K/\gamma)r_0}$$

$$\leq C'|\omega| e^{-\frac{t}{2h} + \frac{28}{32} + (8\alpha K/\gamma)r_0}.$$ 

(8.10)

for some constant $C' > 0$.

By the condition on $\alpha$, if $\gamma \geq \gamma_I$, we get an estimate of the measure of parameters for large escape times. Let us suppose that $t > 2hr_0$. Then

$$m(\{a \in \omega : E_l(a, \nu) = t\}) \leq Ce^{-\frac{t}{2h} |\omega|}.$$ 

We now follow a parameter in $a \in \omega$ in a time window $[n, 2n]$, and estimate its total time spent on escaping from essential returns. Recall that given an essential return $\xi_{\nu}(\omega_{\nu}(a))$, the parameter $a$ has to escape first before we can start counting
the next escape time. Let

\[ T_n(a) = T_{n,l}(a) = \sum_{j=1}^{s(a)} E_l(a, \nu_j(a)), \]

where \( \nu_j(a) \) are essential returns after escape situations, and \( s = s(a) \) the total number of such returns in \([n, 2n]\). We include shallow returns above also but then, by definition, the escape time is zero, so one needs only consider deep returns in the sum.

Remark 8.4. A note on the last return \( \nu_s \) in the expression of \( T_{n,l}(a) \). The escape period of the last return \( \nu_s = \nu_s(a) \), by definition, has to transcend into the next time window \([2n, 4n]\). If it is too long it may deteriorate the Lyapunov exponent for that parameter too much. Here we make the following convention, namely that if \( E(a, \nu_s) \geq 6\delta \) (where \( 6\delta \) (\( \delta \) is the escape time), then we delete those parameters. They constitute an exponentially small portion of the parameters in \( \omega \) (put \( t = 6\delta a \) in equation (8.11)), i.e. has measure \( \leq |\omega|Ce^{-qn} \), where \( q = 2a \). We simply disregard those parameters in the above expression for \( T_{n,a} \). They can easily be taken care of in the final proof in the next section.

In order to reach the main conclusion that the set of parameters having too many too deep returns in the time window \([n, 2n]\) has small measure, we want to estimate, for suitable \( \theta > 0 \), the integral

\[ \frac{1}{|\omega|} \int_{\omega} e^{\theta T_{n}(a)} \, da. \]

There is some freedom of how to choose \( \theta \), but let us set \( \theta = 1/(6h) \).

**Lemma 8.5.** Let \( \xi_{n,l}(\omega) \) be a deep essential return with \( \text{dist}(\xi_{n,l},(\omega), \text{Crit}_l) \sim \sqrt{e} e^{-r} \), \( n \leq \nu \leq 2n \), and \( \omega \subset E_{n,l}(\gamma_1) \cap B_{\nu,1,\ast} \) for some \( \gamma_1 \geq \gamma_1 \). Suppose also that all parameters \( a \in \omega \cap B_{2n,1} \) has that \( a \in E_{n,l}(\gamma_1) \). Then

\[ \int_{\{a \in \omega: 2h_\nu \leq E_l(a, \nu) \leq n \}} e^{\theta E_l(a, \nu)} \, da \leq Ce^{-r/3} |\omega|, \]

(8.12)

\[ \int_{\{a \in \omega: E_l(a, \nu) \leq 2h_\nu \}} e^{\theta E_l(a, \nu)} \, da \leq Ce^{r/3} |\omega|. \]

(8.13)

**Proof.** By (8.11) we have,

\[ \int_{\{a \in \omega: E_l(a, \nu) \geq 2h_\nu \}} e^{\theta E_l(a, \nu)} \, da \leq C \sum_{t \geq 2h_\nu} e^{-\frac{3}{2}e^{\theta t}} |\omega| \]

(8.14)

\[ \leq Ce^{-\frac{3}{2}e^{\theta t}} |\omega| \leq Ce^{-r/3} |\omega|. \]

The second inequality follows directly. \( \square \)

**Lemma 8.6.** Let \( \xi_{n,l}(\omega) \) be an essential return with \( \text{dist}(\xi_{n,l}(\omega), \text{Crit}_l) \sim \sqrt{e} e^{-r} \), \( n \leq \nu \leq 2n \), and \( \omega \subset E_{n,l}(\gamma_1) \cap B_{\nu,1,\ast} \) for some \( \gamma_1 \geq \gamma_1 \). Suppose also that all parameters \( a \in \omega \cap B_{2n,1} \) has that \( a \in E_{n,l}(\gamma_1) \). Then for any \( \epsilon > 0 \) there is a \( \Delta_2 \) such that if \( \Delta \geq \Delta_2 \) (recall \( \delta = e^{-\Delta} \)), we have

\[ \int_{\omega} e^{\theta T_{n,l}(a)} \, da \leq e^{\epsilon 2n} |\omega|. \]
Proof. Let \( \hat{\omega} \subset \omega \) be a subset of \( \omega \) such that every parameter \( a \in \hat{\omega} \) has \( s \) number of free returns into \( U \) after escape situations. So \( T_{n,l}(a) \) consists of \( s \) terms of the form \( E_l(a, \nu_j(a)) \), \( j = 1, \ldots, s \), where \( \nu_1 = \nu \). Recall that \( E_l(a, \nu_j(a)) = 0 \) if the return is shallow. Set \( \xi_{n,l}(\omega) = \xi_n(\omega) \). Every parameter \( a \in \hat{\omega} \) has a nested sequence of corresponding intervals so that \( a \in \omega^s_1 \subset \omega^s_2 \subset \ldots \subset \omega^s_1 \subset \hat{\omega} \), such that \( \xi_{\nu_j+1}(\omega^j) \) is in escape position and \( \xi_{\nu_j}(\omega^j) \) is an essential return, \( \omega^j \subset \hat{\omega}^j \). We have \( \hat{\omega}^1 = \omega \), by assumption. We also see that \( E_l(a, \nu_j(a)) \) is constant on \( \omega^j = \omega^j(a) \) but not on \( \omega^j_1 \). We think of \( \omega^1 = \omega^1(a) \subset \hat{\omega} \) as an interval around \( a \) which has escaped at time \( \nu_2 = \nu_2(a) \) (possibly earlier). Then \( \omega^2 \) is another smaller interval around \( a \) which has escaped at time \( \nu_3 \) (possibly earlier) and so on. In the construction one should think of \( \omega \) as contained in some larger interval \( \omega^0, \omega \subset \omega^0 \) where \( \xi_{\nu_1}(\omega^0) \) is in escape position, and where \( \xi_{\nu_1}(\omega) \) is an essential return.

Since \( T_{n,l}(a) = \sum_{j=1}^s E_l(a, \nu_j) \), and \( E_l(a, \nu_j(a)) \) is constant on \( \omega^j \) but not on \( \omega^j_1 \), we get,

\[
\int_{\omega^s_1} e^{\theta T_{n,l}(a)} da = \sum_{j=1}^{s-1} e^{\theta E_l(a, \nu_j)} \int_{\omega^s_1} e^{\theta E_l(a, \nu_j)} da.
\]

Now, \( \xi_{\nu_1}(\omega^s_1) \) is in escape position and therefore each interval \( \omega^s_1 \subset \omega^s_1, r \) where \( \text{diam}(\xi_{\nu_1}(\omega^s_1, r)) \sim e^{-r} \). Also \( \omega^s_1 \) is a union of disjoint intervals \( \omega^s_1, r \), i.e.

\[
\omega^s_1 = \bigcup_{r=\Delta}^\infty \omega^s_1, r.
\]

Recall that the escape time \( E_l(a, \nu_s(a)) = 0 \) for \( a \in \omega^s_1, r \) if \( r \leq 2\Delta \). By Lemma 8.5 we have

\[
\int_{\omega^s_1} e^{\theta E_l(a, \nu_s)} da \leq |\omega^s_1| + \sum_{r=2\Delta}^{\infty} \int_{\omega^s_1, r} e^{\theta E_l(a, \nu_s)} da
\]

\[
\leq |\omega^s_1| + \sum_{r=2\Delta}^{\infty} \left( \int_{\{a \in \omega^s_1: E_l(a, \nu_s(a)) \geq 2hr\}} e^{\theta E_l(a, \nu_s)} da
\]

\[
+ \int_{\{a \in \omega^s_1: E_l(a, \nu_s(a)) \leq 2hr\}} e^{\theta E_l(a, \nu_s)} da \right)
\]

\[
\leq |\omega^s_1| + C \sum_{r=2\Delta}^{\infty} (e^{r/3} + e^{-r/3}) |\omega^s_1, r|.
\]

Since \( \xi_{\nu_1}(\omega^s_1) \) is in escape position, by the Main Distortion Lemma the parameters \( a \) that enter into the set where \( \text{diam}(\xi_{\nu_s}(\omega^s_1, r)) \sim e^{-r} \) has measure \( \sim \frac{e^{-r}}{\delta} |\omega^s_1| \). Therefore,

\[
\int_{\omega^s_1} e^{\theta E_l(a, \nu_s)} da \leq |\omega^s_1| + C \sum_{r=2\Delta}^{\infty} (e^{r/3} + e^{-r/3}) \frac{e^{-r}}{\delta} |\omega^s_1|
\]

\[
= |\omega^s_1| (1 + C e^{-\Delta/3}) = |\omega^s_1| (1 + \eta(\Delta)),
\]

where \( \eta(\Delta) \to 0 \) as \( \Delta \to \infty \).

Next, we want to compute the integral over \( \omega^{s-2} \): Again \( \xi_{\nu_{s-2}}(\omega^{s-2}) \) is in escape position and therefore \( \omega^{s-2} \) is subdivisioned into disjoint intervals of the type
\( \omega^{s-2,r} \subset \omega^2 \) as \( \omega^{s-1} \):

\[
\omega^{s-2} = \bigcup_{r=\Delta}^{\infty} \omega^{s-2,r}.
\]

Since \( E_i(a, \nu_j(a)) \) is constant on \( \omega^j \), we now compute

\[
\int_{\omega^{s-2}} e^{\theta(E_i(a, \nu_s) + E_i(a, \nu_{s-1}))} \, da = \sum_{\omega^{s-1}} \int_{\omega^{s-2} \cap \omega^{s-1}} e^{\theta E_i(a, \nu_s)} \, da \\
\leq \sum_{\omega^{s-1} \subset \omega^{s-2}} e^{\theta E_i(a, \nu_{s-1})} (1 + \eta(\Delta)) |\omega^{s-1}| \\
(8.17) \quad = (1 + \eta(\Delta)) \int_{\omega^{s-2}} e^{\theta E_i(a, \nu_{s-1})} \, da.
\]

Thus,

\[
\int_{\omega^{s-2}} = \sum_{r \geq 2\Delta} \int_{\omega^{s-2}} e^{\theta E_i(a, \nu_{s-1}) + E_i(a, \nu_s)} \, da \\
\leq (1 + \eta(\Delta)) \sum_{r \geq 2\Delta} \int_{\omega^{s-2}} e^{\theta E_i(a, \nu_{s-1})} \, da \\
(8.18) \quad \leq (1 + \eta(\Delta)) \int_{\omega^{s-2}} e^{\theta E_i(a, \nu_{s-1})} \, da \leq (1 + \eta(\Delta))^2 |\omega^{s-2}|.
\]

Repeating this \( s \) times and noting that \( s \leq \nu \) trivially and that \( \eta(\Delta) \to 0 \) as \( \Delta \to \infty \), we get

\[
\int_{\omega^{0}} e^{\theta T_{n,i}(a)} \, da \leq (1 + \eta(\Delta))^s |\omega^0| \leq e^{\epsilon_2 n} |\omega^0|.
\]

Since this holds for every set of the type \( \tilde{\omega} \) (and letting \( s \) vary) the lemma follows. \( \square \)

Finally we can prove the main goal in this section.

**Lemma 8.7.** Let \( \tau > 0 \) be such that \( \tau \theta > \varepsilon_2 \) and suppose that \( \xi_{\nu,l}(\omega) \) is a deep essential return with \( \text{dist}(\xi_{\nu,l}(\omega), \text{Crit}_\omega) \sim \sqrt{\tau} e^{-r} \), \( n \leq \nu \leq 2n \), and \( \omega \subset \mathcal{E}_{\nu,l,\gamma} \cap B_{\nu,l} \) for some \( \gamma \geq \gamma_1 \). Suppose also that all parameters \( a \in \omega \cap B_{2n,l} \) has that \( a \in \mathcal{E}_{2n,l}(\gamma_1) \). Then

\[
m(\{a \in \omega : T_n(a) \geq \tau n\}) \leq e^{n(\varepsilon_2 - \theta r)} |\omega|.
\]

**Proof.** We have by Lemma 8.6,

\[
e^{\theta n} m(\{a \in \omega : T_n(a) \geq \tau n\}) \leq \int_{\{T_n(a) > \tau n\}} e^{\theta T_n(a)} \, da \leq \int_{\omega} e^{\theta T_n(a)} \, da \leq e^{\varepsilon_2 n} |\omega|,
\]

from which we conclude that

\[
m(\{a \in \omega : T_n(a) \geq \tau n\}) \leq e^{n(\varepsilon_2 - \theta r)} |\omega|.
\]

\( \square \)
9. Conclusion and proof of the main theorem

We make induction over time intervals of the type \([n, 2n]\). By Lemma 7.4, for a sufficiently small starting interval \(\omega_0 = (−\varepsilon, \varepsilon)\) around the starting map \(f_0\), there are numbers \(N_l\) such that \(\xi_{N_l,l}(\omega_0)\) has grown to the large scale or returned into \(U\) with \(\text{dist}(\xi_{N_l,l}(\omega_0), \text{rit}(\omega_0))/\log(\text{dist}(\xi_{N_l,l}(\omega_0), \text{rit}(\omega_0))) \leq \text{diam}(\xi_{N_l,l}(\omega_0))\), i.e. in the case of a return, it has to be essential. Suppose, without loss of generality, that the first critical point \((l = 1)\) has that \(N_1 = \min(N_l)\). Let \(v_0 \geq N_1\) be the first return into \(U\). It follows that \(\xi_{v_0,1}(\omega_0)\) is an essential return.

If \(v_0 > 2N_1\) then it means we have no more returns in \([N_1, 2N_1]\) for \(l = 1\) so we go on to the next critical point. To start, put \(n = N_1\). For each critical point, we consider the returns \(\nu_j \in [n, 2n]\) and delete parameters according to the basic approach rate condition. If \(\hat{\omega}_0 \subset \omega_0\) is the set that is left from \(\omega_0\) when we have deleted parameters not satisfying this condition up until time \(2n\), then by Lemma 8.1,

\[|\hat{\omega}_0| \geq (1 - e^{-\alpha n})|\omega_0| .\]

We make this construction for each critical point, and thereby get a set \(\Omega_l(2N_1)\), which corresponds to \(\hat{\omega}_0\) for each \(l\), and which contains parameters in \(\omega_0\) that satisfy the basic assumption for the critical point \(c_l\). Up until time \(n = N_1\) we see that \(\omega_0 \subset E_{N_1,l}(\gamma_B)\), by making \(\varepsilon\) sufficiently small. Actually we have a stronger statement at this early stage according to the Starting Lemma (the Lyapunov exponents are close to \(\gamma_0\)), but we do not need that. Moreover, by definition we have \(\omega_0 \subset \mathcal{B}_{N_1-1,l}\) for all \(l\). Obviously, \(\Omega_l(2N_1) \subset \mathcal{B}_{2N_1,l}\).

If we do not do anything more than keeping the parameters satisfying the basic approach rate condition, the Lyapunov exponent may drop in the time window \([n, 2n]\), and over time we may lose too much. Every return in this time window has a bound period \(p_j \leq \nu_j(2K\alpha/\gamma_1) = \hat{\alpha}\nu_j \leq 2n\), for the returns \(\nu_j \in [n, 2n]\), where we have set \(\hat{\alpha} = 2K\alpha/\gamma_1\). Hence we can use the expansion of the early orbits up until time \(2n\) for all such bound periods. We also note that by Lemma 6.4, the bound periods are bounded from below by \(Kr_j/(2\Gamma) \geq K\Delta_1/(2\Gamma)\). Let \(L_j\) be the corresponding free periods. For every parameter which satisfies the basic approach rate condition, by Lemmas 3.1 and 6.4, using that \(a \in \mathcal{E}_{n,l}(\gamma_B)\), we have, if \(\delta = e^{-\Delta}\) is sufficiently small,

\[|Df^{2n}(\nu_l(a))| \geq C_0 e^{\gamma n} \prod_j \left( e^{p_j(\gamma/(2K))} C_0 e^{L_j(\gamma_0)} \right) \geq e^{\gamma n} e^{\sum_j p_j(\gamma/(4K)) + \gamma_0 L_j} \geq C_0 e^{\gamma(1/2)2n} .\]

Hence up until time \(2n = 2N_1\), we may have lost some part of the starting Lyapunov exponent \((\gamma_B)\), but at each return it does not go below \(\gamma_B/2 > \gamma_1\), where \(\gamma_1\) is a lower bound for most lemmas in the induction process. However, precisely after a return the exponent may drop, but not more than \(4K\alpha\) because of the basic approach rate assumption (the \(2n\) is replaced by \(4\alpha\) to eat up constants), and in general each parameter \(a\) we are considering belongs to \(\mathcal{E}_{n,l}(\gamma)\) for some \(\gamma \geq \gamma_B/2 - 4K\alpha \geq \gamma_1\).

Therefore we may have to delete more parameters, that return too often and too deep, in order to restore the Lyapunov exponent for the remaining parameters. This is handled in the section about large deviations. The large deviation argument estimates the set of those parameters that spend a too large portion of the time in \([n, 2n]\) reaching escape positions. Since the escape period is set to zero for shallow
returns, i.e. for returns into \( U \setminus U^2 \), the orbits \( \xi_{n,l}(a) \) outside \( U^2 \) can be considered as free periods. Using Lemma 3.1 for this neighbourhood \( U^2 \) also gives uniform expansion until the next return (let us use the same exponent \( \gamma_H > 0 \) for those free periods). Since each bound period for a deep return into \( U^2 \) is contained in an escape period, we now consider those bound periods \( \tilde{p}_j \) in \([n,2n]\) and the corresponding free periods \( \tilde{L}_j \) outside \( U^2 \). If the parameter \( a \) is such that \( T_n(a) \leq \tau n \) where \( 0 < \tau < 1 \) then

\[
|Df^{2n}(v_l(a))| \geq C_0 e^{\gamma_B n} e^{\sum_j \tilde{p}_j (\gamma/(4K))} e^{\gamma_H \tilde{L}_j} \geq C_0 e^{\gamma_B n} e^{(1-\tau)n\gamma_H}.
\]

According to the definition, \( \gamma_B = (3/4)(1-\tau) \min(\gamma_H, \gamma_0) \), and hence the Lyapunov exponent is restored:

\[
|Df^{2n}(v_l(a))|C_0 \geq e^{\gamma_B 2n}.
\]

Let us now turn to the general case where we use induction. Assume that we have constructed \( \Omega_l(n) \) for every \( l \) and that the sets \( \Omega_l(n) \) are “good” in the following sense. We assume that each partition element \( \omega \subset \Omega_l(n) \) belongs to \( \mathcal{E}_{n,l}(\gamma_B) \cap \mathcal{B}_{n,l} \), i.e. \( \Omega_l(n) \subset \mathcal{E}_{n,l}(\gamma_B) \cap \mathcal{B}_{n,l} \). The sets \( \Omega_l(n) \) have their own structure and should not be mixed until at the very end, because the partition elements in each such set may differ, and and intersection therefore can destroy these elements.

For simplicity assume that \( \nu = n \) is a return time for \( l \). By definition of \( \mathcal{E}_{n,l}(\gamma) \) and \( \mathcal{B}_{n,l} \), we can use the binding information for all critical point up until time \((2K\alpha/\gamma_I)\nu = \tilde{\alpha} \nu \). First let us from \( \Omega_l(n) \) delete parameters so that we can use the binding information of all other critical points \( j \neq l \) for a longer time, in the next time window \([2n,4n]\), i.e. we want to consider \( \mathcal{E}_{n,l}(\gamma) \cap \mathcal{B}_{n,l} \). The point is now that the partition elements we are deleting, i.e. parameters belonging to \((\mathcal{E}_{n,l}(\gamma) \cap \mathcal{B}_{n,l}) \setminus (\mathcal{E}_{n,l}(\gamma) \cap \mathcal{B}_{n,l})\), by this procedure are much larger than the partition elements in \( \Omega_l(n) \) (this was originally observed by M. Benedicks). Indeed, if we look at the length of \( \xi_{m,j}(\omega_1) \) where \( \omega_1 \) is a partition element that got deleted at some time (return) \( m \leq 2\tilde{\alpha} n \) then, by Lemma 4.2,

\[
\text{diam}(\xi_{m,j}(\omega_1)) \sim |\omega_1| |Df^m(v_l(a))| \leq |\omega_1| e^{\Gamma m}.
\]

By the basic assumption, and since \( \omega_1 \) got deleted at time \( m \), we have

\[
\text{diam}(\xi_{m,j}(\omega_1)) \sim \text{dist}(\xi_{m,j}(\omega_1), \text{Crit}_{\omega_1})/(\log(\text{dist}(\xi_{m,j}(\omega_1), \text{Crit}_{\omega_1})))^2 \geq e^{-3\alpha m},
\]

so

\[
e^{-3\alpha m} \leq C \text{diam}(\xi_{m,j}(\omega_1)) \leq C|\omega_1| e^{\Gamma m}.
\]

On the other hand, the partition elements at time \( n \) or higher, are much smaller. This can be seen as follows. Let \( \omega_2 \) be a partition element at time \( n \). Since \( \text{diam}(\xi_{n,j}(\omega_2)) \leq S \), we have

\[
S \geq \text{diam}(\xi_{n,j}(\omega_2)) \sim |\omega_2| |Df^n(v_l(a))| \geq |\omega_2| C_0 e^{\gamma n}.
\]

Therefore, since \( m \leq 2\tilde{\alpha} n \),

\[
\frac{|\omega_1|}{|\omega_2|} \geq C e^{-\gamma n} e^{(3\alpha + \Gamma)m} \geq C e^{\gamma - 2\tilde{\alpha}(3\alpha + \Gamma)n} \gg 1.
\]

Hence \( \omega_2 \) is much smaller than \( \omega_1 \). This means that when deleting partition elements in \( \Omega_l(n) \) that do not satisfy the basic approach rate condition until time \( 2\tilde{\alpha} n \) for other critical points \( j \neq l \), in the time window \([\tilde{\alpha} n, 2\tilde{\alpha} n]\), we do not destroy the partition elements; we only delete whole partition elements of the type \( \omega_2 \in \Omega_l(n) \).
that intersect partition elements of the type $\omega_1$ that was deleted at the time scale $\sim \hat{\alpha} n$.

Starting from the partition elements in $\Omega_l(n) \subset E_{n,l}(\gamma) \cap B_{n,l}$ and passing to $\Omega_l(n, \star) \subset \mathcal{E}_{n,l}(\gamma) \cap B_{n,l,\star}$ is therefore harmless and the measure deleted is

\begin{equation}
|\Omega_l(n, \star)| \geq (1 - C e^{-\hat{\alpha} n})|\Omega_l(n)|. \tag{9.3}
\end{equation}

We have now constructed $\Omega_l(n, \star)$ and want to pass to $\Omega_l(2n) \subset E_{2n,l}(\gamma) \cap B_{2n,l}$. Passing from $\Omega_l(n, \star)$ to $\Omega_l(2n)$, we have to delete parameters that do not satisfy the basic approach rate condition for critical point $c_l$ and also delete those parameters that have too many too deep returns in $[n, 2n]$. We have seen by equation (9.1), that the Lyapunov exponent can decrease to $(1/2)\gamma B > \gamma_I$ during the period form $n$ to $2n$. We also have to take into account the blind escapes, see Remark 8.4, which constitute a small portion, $\leq C e^{-qn}$, of the original set of parameters. For those parameters whose escape periods transcend into $[2n, 4n]$ (these are the escape periods for the last return $\nu_s$ discussed in Remark 8.4), the Lyapunov exponent may drop slightly below $\gamma_B/2$, but never below $\gamma_I$ (if $\alpha$ is sufficiently small, see the condition on p. 4).

By Lemmas 8.1 and 8.6 we get the estimate

\[ |\Omega_l(2n)| \geq (1 - e^{-\alpha n})(1 - C e^{-(\theta \tau - \varepsilon) n})(1 - C e^{-qn})|\Omega_l(n, \star)|. \]

Together with (9.3), we get, for some $\beta > 0$,

\[ |\Omega_l(2n)| \geq (1 - e^{-\alpha n})(1 - C e^{-(\theta \tau - \varepsilon) n})(1 - C e^{-qn})(1 - C e^{-\hat{\alpha} n})|\Omega_l(n)| \geq (1 - e^{-\beta n})|\Omega_l(n)|. \]

It follows that $\Omega_l(2n) \subset \mathcal{E}_{2n,l}(\gamma) \cap B_{2n,l}$, where $\gamma \geq \gamma_B$ by the choice of $\gamma_B$ (possibly, if $n$ is just after a return time, $\gamma \geq \gamma_B - 4K \alpha$). We are then back to the same situation at time $2n$ as we were for time $n$ and the induction argument goes on forever.

Let $M \geq 2$. Choosing the constants correctly, in this way we construct, for each critical point, a set $\Omega_l(n) \subset \mathcal{E}_{n,l}(\gamma_B - 4K \alpha) \cap B_{n,l}$ with measure at least $(1 - 1/(2Md))|\omega_0|$, that holds for $n > 0$, where $d$ is the degree of $f$. Passing to the limit, as $n \to \infty$, we get that the measure of parameters that satisfies the CE-condition for all $n > 0$ is estimated by

\[ \lim_{n \to \infty} m \left( \bigcap_l \Omega_l(n) \right) \geq (1 - \frac{1}{M})|\omega_0|. \]

Since $M$ can be chosen arbitrarily large, it follows that $f_0$ is a Lebesgue density point of CE-maps.
References

[1] Magnus Aspenberg. The Collet-Eckmann condition for rational functions on the Riemann sphere. Math. Z., 273:935–980, 2013.

[2] Matthieu Astorg, Thomas Gauthier, Nicolae Mihalache, and Gabriel Vigny. Collet, Eckmann and the bifurcation measure. Invent. Math., 217(3):749–797, 2019.

[3] Artur Avila and Carlos Gustavo Moreira. Statistical properties of unimodal maps: the quadratic family. Ann. of Math. (2), 161(2):831–881, 2005.

[4] Michael Benedicks and Lennart Carleson. On iterations of $1 – ax^2$ on $(-1, 1)$. Ann. of Math. (2), 122(1):1–25, 1985.

[5] Michael Benedicks and Lennart Carleson. The dynamics of the Hénon map. Ann. of Math. (3), 133(1):73–169, 1991.

[6] Michael Benedicks and Michał Misiurewicz. Absolutely continuous invariant measures for maps with flat tops. Inst. Hautes Études Sci. Publ. Math., (69):203–213, 1989.

[7] Henk Bruin, Weixiao Shen, and Sebastian van Strien. Invariant measures exist without a growth condition. Comm. Math. Phys., 241(2-3):287–306, 2003.

[8] Mats Bylund. Equivalence of Collet-Eckmann conditions for slowly recurrent rational maps. Manuscript in preparation and personal communication.

[9] P. Collet and J.-P. Eckmann. On the abundance of aperiodic behaviour for maps on the interval. Comm. Math. Phys., 73(2):115–160, 1980.

[10] P. Collet and J.-P. Eckmann. Positive Liapunov exponents and absolute continuity for maps of the interval. Ergodic Theory Dynam. Systems, 3(1):13–46, 1983.

[11] Pierre Collet and Jean-Pierre Eckmann. Iterated maps on the interval as dynamical systems, volume 1 of Progress in Physics. Birkhäuser, Boston, Mass., 1980.

[12] Bing Gao and Weixiao Shen. Summability implies Collet-Eckmann almost surely. Ergodic Theory Dynam. Systems, 34(4):1184–1209, 2014.

[13] Jacek Graczyk and Stanislav Smirnov. Collet-Eckmann and Hölder. Invent. Math., 133(1):69–96, 1998.

[14] Jacek Graczyk and Stanislav Smirnov. Non-uniform hyperbolicity in Complex Dynamics. Invent. Math., 175:335–415, 2009.

[15] Jacek Graczyk and Grzegorz Świątek. Analytic structures and harmonic measure at bifurcation loci. Preprint, arXiv:1904.09434.

[16] Jacek Graczyk and Grzegorz Świątek. Generic hyperbolicity in the logistic family. Ann. of Math. (2), 146(1):1–52, 1997.

[17] Jacek Graczyk and Grzegorz Świątek. Fine structure of connectedness loci. Math. Ann., 369(1-2):49–108, 2017.

[18] Michael V. Jakobson. Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. Comm. Math. Phys., 81(1):39–88, 1981.

[19] Gerhard Keller. Exponents, attractors and Hopf decompositions for interval maps. Ergodic Theory Dynam. Systems, 10(4):717–744, 1990.

[20] Oleg Kozlovski, Weixiao Shen, and Sebastian van Strien. Density of hyperbolicity in dimension one. Ann. of Math. (2), 166(1):145–182, 2007.

[21] Genadi Levin. Multipliers of periodic orbits in spaces of rational maps. Ergodic Theory Dynam. Systems, 31(1):197–243, 2011.

[22] Genadi Levin. Perturbations of weakly expanding critical orbits. In Frontiers in complex dynamics, volume 51 of Princeton Math. Ser., pages 163–196. Princeton Univ. Press, Princeton, NJ, 2014.

[23] Genadi Levin. Fixed points of the Ruelle-Thurston operator and the Cauchy transform. Fund. Math., 254(1):49–67, 2021.

[24] Genadi Levin, Feliks Przytycki, and Weixiao Shen. The Lyapunov exponent of holomorphic maps. Invent. Math., 205(2):363–382, 2016.

[25] Mikhail Lyubich. Dynamics of quadratic polynomials. I, II. Acta Math., 178(2):185–247, 247–297, 1997.

[26] Mikhail Lyubich. Almost every real quadratic map is either regular or stochastic. Ann. of Math. (2), 156(1):1–78, 2002.

[27] Tomasz Nowicki and Sebastian van Strien. Invariant measures exist under a summability condition for unimodal maps. Invent. Math., 105(1):123–136, 1991.
[28] Michal Misiurewicz. Absolutely continuous measures for certain maps of an interval. *Inst. Hautes Études Sci. Publ. Math.*, (53):17–51, 1981.

[29] Feliks Przytycki. Iterations of holomorphic Collet-Eckmann maps: conformal and invariant measures. Appendix: on non-renormalizable quadratic polynomials. *Trans. Amer. Math. Soc.*, 350(2):717–742, 1998.

[30] Feliks Przytycki. Hölder implies Collet-Eckmann. *Astérisque*, (261):xiv, 385–403, 2000. Géométrie complexe et systèmes dynamiques (Orsay, 1995).

[31] Feliks Przytycki, Juan Rivera-Letelier, and Stanislav Smirnov. Equivalence and topological invariance of conditions for non-uniform hyperbolicity in the iteration of rational maps. *Invent. Math.*, 151(1):29–63, 2003.

[32] Mary Rees. Positive measure sets of ergodic rational maps. *Ann. Sci. École Norm. Sup. (4)*, 19(3):383–407, 1986.

[33] Masato Tsujii. Positive Lyapunov exponents in families of one-dimensional dynamical systems. *Invent. Math.*, 111(1):113–137, 1993.

Department of Mathematics, Lund University, Box 118, 221 00 Lund, Sweden