ISRAEL JOURNAL OF MATHEMATICS **137** (2003), 223–263

EXPANSION OF DERIVATIVES IN ONE-DIMENSIONAL DYNAMICS*,†

BY

HENK BRUIN

Department of Mathematics, University of Groningen
P.O. Box 800, 9700 AV Groningen, The Netherlands
e-mail: bruin@math.rug.nl
http://www.math.rug.nl/~bruin

AND

SEBASTIAN VAN STRIEN

Department of Mathematics, University of Warwick
Coventry CV4 7AL, UK
e-mail: strien@maths.warwick.ac.uk
http://www.maths.warwick.ac.uk/~strien

ABSTRACT

We study the expansion of derivatives along orbits of real and complex one-dimensional maps f, whose Julia set J_f attracts a finite set $Crit$ of non-flat critical points. Assuming that for each $c \in Crit$, either $|Df^n(f(c))| \to \infty$ (if f is real) or $b_n \cdot |Df^n(f(c))| \to \infty$ for some summable sequence $\{b_n\}$ (if f is complex; this is equivalent to summability of $|Df^n(f(c))|^{-1}$), we show that for every $x \in J_f \setminus \bigcup_i f^{-i}(Crit)$, there exist $\ell(x) \leq \max_c \ell(c)$ and $K'(x) > 0$ such that

$$|Df^n(x)|^{\ell(x)} \geq K'(x) \prod_{i=0}^{s-1} D_{n_i - n_{i+1}}(c_{i+1})$$

* The original paper used an incorrect version of the Koebe Lemma cited from [21] as was pointed out by the referee and Genadi Levin in the autumn of 2001. The corrected version of November 2001 only uses the classical Koebe Lemma. Apparently, all results in Feliks Przytycki's paper [21] go through using the classical Koebe Lemma instead of his Lemma 1.2.

† Both authors gratefully acknowledge the support of the PRODYN program of the European Science Foundation. HB was partially supported by a fellowship of The Royal Netherlands Academy of Arts and Sciences (KNAW). SvS was partially supported by GR/M82714/01.

Received October 25, 2000 and in revised form July 26, 2002

for infinitely many n. Here $0 = n_s < \cdots < n_1 < n_0 = n$ are so-called critical times, c_i is a point in $Crit$ (or a repelling periodic point in the boundary of the immediate basin of a hyperbolic periodic attractor), which shadows $\mathrm{orb}(x)$ for $n_i - n_{i+1}$ iterates, and

$$D_k(c_i) = \begin{cases} \max(\lambda, K \cdot |Df^k(f(c_i))|) & \text{if } f \text{ is real,} \\ \max(\lambda, K \cdot b_k \cdot |Df^k(f(c_i))|) & \text{if } f \text{ is complex,} \end{cases}$$

for uniform constants $K > 0$ and $\lambda > 1$. If all $c \in Crit$ have the same critical order, then $K'(x)$ is uniformly bounded away from 0. Several corollaries are derived. In the complex case, either $J_f = \hat{\mathbb{C}}$ or J_f has zero Lebesgue measure. Also (assuming all critical points have the same order) there exist $\kappa > 0$ such that if n is the smallest integer such that x enters a certain critical neighbourhood, then $|Df^n(x)| \geq \kappa$.

1. Introduction and statement of results

The behaviour of (real and complex) one-dimensional systems depends essentially on the behaviour of their critical points. Derivatives along orbits tend to grow at least as fast as the derivatives along the critical orbits, and so the growth along critical orbits plays a central role in the question whether invariant densities exist. This goes back to Collet and Eckmann's result [7] on non-flat S-unimodal maps f: if there exist $C > 0$ and $\lambda > 1$ such that

$$\text{(CE)} \qquad |Df^n(f(c))| \geq C\lambda^n \quad \text{for all } n \geq 0,$$

where c is the critical point, and

$$\text{(BCE)} \qquad |Df^n(f(x))| \geq C\lambda^n \quad \text{where } n \text{ is minimal such that } f^n(x) = c,$$

then f admits an invariant probability measure (acip) that is absolutely continuous with respect to Lebesgue. In the S-unimodal case, Nowicki [16] showed that the Collet–Eckmann condition (CE) implies the backward Collet–Eckmann condition (BCE), thus eliminating the need of (BCE), and in [17] it was shown that (CE) and (BCE) are equivalent.

In the (multimodal) real and complex case it is convenient to define

$$J_f = \begin{cases} \overline{\{x \in I; f^n(x) \not\to \text{ attracting periodic orbit}\}} & \text{if } f \text{ is real,} \\ \text{the Julia set} & \text{if } f \text{ is complex.} \end{cases}$$

The set of critical points c such that $\overline{\mathrm{orb}(c)}$ intersects J_f will be denoted by $Crit$. We say that f is a *Collet–Eckmann* map if there exist $C > 0$ and $\lambda > 1$ such that

$$\text{(CE)} \qquad |Df^n(f(c))| \geq C\lambda^n \quad \text{for all } c \in Crit \text{ and } n \geq 0.$$

Here, and throughout the paper, in the complex case Df^n is the derivative in terms of the spherical metric on $\hat{\mathbb{C}}$. The (BCE) condition is extended similarly. Some authors allow some critical points in *Crit* to be mapped onto other critical points. It is not too hard to extend our results to that setting. In the setting of rational maps, Przytycki [21] showed that (CE) together with some technical assumptions imply the existence of an acip (with respect to a non-atomic conformal measure on the Julia set).

Our main aim in this paper is to deal with weaker growth rates, similar to what was done in Nowicki and van Strien [18], who showed that for S-unimodal maps, summability of $|Df^n(f(c))|^{-1/\ell}$ (where ℓ is the order of the critical point c) guarantees the existence of an acip. In [5] we extended this to the multimodal case:

THEOREM 1.1: *Let f be a real multimodal map such that $Sf \leq 0$ and all critical points have the same finite order ℓ. If $\sum_n |Df^n(f(c))|^{-1/\ell} < \infty$ for each critical point c, then f admits an acip.*

It is the aim of this paper, using the ideas of [18], to estimate derivatives along arbitrary orbits, and give alternative approaches to (and in some cases improvements of) results by Graczyk and Smirnov [9] and Przytycki [21]. We assume that

$$\begin{cases} |Df^n(f(c))| \to \infty & \text{if } f \text{ is real with } Sf \leq 0, \\ \sum_n |Df^n(f(c))|^{-1} < \infty & \text{if } f \text{ is rational on } \hat{\mathbb{C}}. \end{cases}$$

Except for finiteness, there are no restrictions on the number of critical points and their orders. The complex assumption is the same as used in [9] or in [21, Theorem B].

We will show that $\mathrm{orb}(x)$ can be decomposed into pieces in which it loosely shadows a critical orbit $\mathrm{orb}(c_i)$, and that the derivatives grow accordingly: There exists $K'(x) > 0$ and $\ell(x) \leq \max_c \ell(c)$ such that

$$|Df^n(x)|^{\ell(x)} \geq K'(x) \cdot \prod_{i=0}^{s-1} (K_i \cdot |Df^{n_i - n_{i+1}}(f(c_i))|),$$

for infinitely many n, and critical times $0 = n_s < \cdots < n_1 < n_0 = n$ (defined in later sections). Here $K_i \geq K$ (if f is real) and $K_i \geq K \cdot b_{n_i - n_{i-1}}$ (if f is complex), and $K > 0$ is a uniform constant. See Theorem 1.2 for the precise statement.

The second purpose of this paper is to strengthen the results in [4]. In that paper we established the existence of an acip μ for real multimodal maps under certain summability conditions and we derived strong statistical properties concerning the mixing rate of μ. However, we assumed a property, called

bounded backward contraction (BBC), which states that $|Df^n(x)|$ is uniformly bounded away from 0 whenever x first enters a critical neighbourhood. Theorem 1.3 below shows that (BBC) holds automatically, both in the real and complex case, provided all critical points have the same order. Examples in Section 6 show that this assumption is essential, and that without it also (CE) no longer implies (BCE).

1.1. Definitions and statements of results. In the remainder of this section we give the precise statements of the results and relate them to known results. Let I denote the interval or circle and $\hat{\mathbb{C}}$ the Riemann sphere with spherical metric. Because $Crit$ is the set of critical points c such that $\overline{\text{orb}(c)}$ intersects J_f, the assumptions (2) and (3) below exclude the existence of parabolic points. In particular, all critical points which are not in $Crit$ are contained in immediate basins of hyperbolic periodic orbits. We assume that $Crit \neq \emptyset$, thus excluding circle diffeomorphisms, as well as maps where J_f is hyperbolic. These maps are completely understood as far as the results presented here are concerned. Each $c_i \in Crit$ has critical order $\ell_i < \infty$ and we write $\ell_{\max} = \max\{\ell_i; c_i \in Crit\}$. Let

$$N_\epsilon(Crit) = \bigcup_i B(c_i, \epsilon^{1/\ell_i}), \tag{1}$$

where $B(x, \delta)$ denotes the open ball of radius δ. We shall consider expansion properties of the following two types of maps:

(i) Smooth real one-dimensional maps $f\colon I \to I$ with non-flat critical points, such that the Schwarzian derivative $Sf \leq 0$, such that ∂I contains no parabolic points of period ≤ 2, and such that

$$|Df^n(f(c))| \to \infty \quad \text{as } n \to \infty \text{ for each } c \in Crit. \tag{2}$$

(ii) Rational maps $f\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ such that there exists a sequence $b_n > 0$ such that

$$\exp\Big(\sum_{n\geq 0} b_n\Big) < 2 \quad \text{and} \quad b_n \cdot |Df^n(f(c))| \to \infty \tag{3}$$

for each $c \in Crit$. The b_n's feature in the notion of shrinking neighbourhoods from [21]. Condition (3) is equivalent to $\sum_{n\geq 0} 1/|Df^n(f(c))| < \infty$; see Lemma 2.7.

Given a point x, define for the moment $\ell(x)$ to be the maximal order of all critical points on which x accumulates (if it does not accumulate on any critical points define $\ell(x) = 1$). A sharper definition of $\ell(x)$ will be given in formulas (8) and (38).

THEOREM 1.2: *Assume that (2) or (3) holds. Then there exist $\lambda > 1$ and $K > 0$ such that for each $x \in J_f$ such that $f^k(x) \notin Crit$ for all $k \geq 0$, there exists $K'(x) > 0$ and arbitrarily large times n such that*

$$|Df^n(x)|^{\ell(x)} \geq K'(x) \prod_{i=0}^{s-1} D_{n_i - n_{i+1}}(c_{i+1}), \tag{4}$$

where

$$D_k(c_i) = \begin{cases} \max(\lambda, K \cdot |Df^k(f(c_i))|) & \text{if } f \text{ is real,} \\ \max(\lambda, K \cdot b_k \cdot |Df^k(f(c_i))|) & \text{if } f \text{ is complex,} \end{cases} \tag{5}$$

and $0 = n_s < \cdots < n_0 := n$ are critical times (defined in later sections), and c_i are critical points in $Crit$ (or repelling periodic points in the boundary of the immediate basin of a hyperbolic periodic attractor). In particular, there exists a sequence of integers n so that $|Df^n(x)| \to \infty$.

Usually, the c_i in this theorem are critical points that shadow the orbit of x for the iterates n_i to n_{i-1}. Only when the shadowing critical point belongs to the (immediate) basin of a periodic attractor, we can relace c_i by a hyperbolic periodic boundary point of this basin, and use its expansion in the estimates.

This allows the following immediate corollary for periodic points:

COROLLARY 1.1: *The Collet–Eckmann condition (CE) implies uniform hyperbolicity on periodic points.*

In the complex setting, this has been shown in [8]. If $Crit$ consists of only one point, then (CE) is equivalent to uniform hyperbolicity on periodic points; see [17] and [8].

It is known (adapt the example in [23, Section 5] and see also [22]) that uniform hyperbolicity on periodic points does not imply (CE). Recall the growth condition on $Crit$ immediately rules out the existence of parabolic points. It is also known that maps as above do not have Cremer points, Siegel disks and Herman rings; see [9] and [24].

The number $K'(x)$ expresses the small derivatives of close visits of $\text{orb}(x)$ to critical points c_r with order $\ell_r > \ell(x)$. On the set $\{x; \ell(x) = \ell_{\max}\}$, $K'(x)$ is uniformly bounded away from 0. Furthermore, (4) holds for the critical times, which includes any n such that $f^n(x) \in Crit$. Therefore we have

COROLLARY 1.2: *If n is minimal such that $f^n(x) = c_i \in Crit$ and $\ell_i = \ell_{\max}$, then $|Df^n(x)|^{\ell(x)} \geq K \cdot \prod_{i=0}^{s-1} D_{n_i - n_{i+1}}(c_{i+1})$. In particular, (CE) implies (BCE) for all c_i with $\ell_i = \ell_{\max}$.*

As mentioned before, the statement that (CE) implies (BCE) was shown in the unimodal case by Nowicki [16]. Graczyk and Smirnov [8, Theorem 1 (i)] proved it for rational maps on the Riemann sphere. The corollary does not hold for critical points c_i with $\ell_i < \ell_{\max}$, as we demonstrate in Section 6.

In the next theorem we prove (BBC). This property was used as assumption in [4].

THEOREM 1.3 (The BBC property): *Assume that (2) or (3) holds and assume that all critical points of f have the same order. Then there exist $\kappa > 0$ so that for any $\epsilon > 0$ and any $x \in J_f$ and $n = \min\{k \geq 0; f^k(x) \in N_\epsilon(Crit)\}$ one has*

$$|Df^n(x)| \geq \kappa. \tag{6}$$

More precisely, there exist uniform constants $\lambda > 1$ and $K, K' > 0$ such that

$$|Df^n(x)|^{\ell(x)} \geq K' \cdot \prod_{i=0}^{s-1} D_{n_i - n_{i+1}}(c_{i+1}),$$

where $D_k(c_i)$ is defined in equation (5).

The assumption that all critical points have the same order is essential; counterexamples are given in Section 6.

A well-known result that goes back to [15] and [12] is that if f is a C^2 multimodal map and $\sigma > 0$, there exists $C > 0$ and $\lambda > 1$ such that if $\mathrm{dist}(\mathrm{orb}(x), Crit \cup \{\text{non-repelling periodic orbits}\}) \geq \sigma$, then

$$|Df^n(x)| \geq C\lambda^n \quad \text{for all } n \geq 0.$$

This determines the hyperbolic subsets of I. In the complex case, Mañé [13] proved a similar result, using the assumption that x does not belong to the omega-limit set of a recurrent critical point. The complex analog of (7) was called **R-expansion** and used as the assumption in [8, Theorem 2].

COROLLARY 1.3: *Let f satisfy (2) or (3). There exists $C > 0$ such that for every $\sigma > 0$ there exists $\lambda > 1$ such that: if $x \in J_f$, then $\mathrm{dist}(\mathrm{orb}(x), Crit) \geq \sigma$ implies $|Df^n(x)| \geq C\sigma^{\ell_{\max}-1}\lambda^n$ for all n.*

In Section 5 we combine the techniques of Theorem 1.2 with a random walk argument and prove that Lebesgue-a.e. point reaches "large scale univalently". This results in

THEOREM 1.4: *If f satisfies (3), then either $J = \hat{\mathbb{C}}$ or the Lebesgue measure of J is 0.*

This result was also shown (mutually independently) in [24].

2. Notation and preliminaries

If a map (real or complex) f has a critical point c, then its order ℓ is the number such that $(\ell/M)|x-c|^{\ell-1} \leq |Df(x)| \leq \ell M|x-c|^{\ell-1}$, for all x and a uniform constant $M > 1$. In particular, we assume that in the real case, the critical order of a critical point is the same on either side of c. Integration gives $(1/M)|x-c|^{\ell} \leq |f(c) - f(x)| \leq M|x-c|^{\ell}$. These properties, plus the fact that $\ell < \infty$, will be referred to as **non-flatness**.

Denote by $\nu(\sigma)$ the minimal integer n so that $|f^n(c) - c'| \leq \sigma$ for some critical points $c, c' \in Crit$. Since no critical point is mapped to another critical point, $\nu(\sigma) \to \infty$ as $\sigma \to 0$.

Throughout the paper we shall need several Koebe-like estimates:

LEMMA 2.1 (Koebe Lemma in the real case): *Let g: $(a,b) \to \mathbb{R}$ be a diffeomorphism and $Sg < 0$. For each $\tau \in (0,1)$ there exists $KL(\tau)$ such that if $x, y \in (a,b)$ is so that both sides of $g((a,b) \setminus (x,y))$ have size $\geq \tau|g(b) - g(a)|$, then $|Dg(x)|/|Dg(y)| \leq KL(\tau)$.*

LEMMA 2.2 (One-sided Koebe Lemma in the real case): *Let g: $(a,b) \to \mathbb{R}$ be a diffeomorphism and $Sg \leq 0$. For each $\tau \in (0,1)$ there exists $KL(\tau)$ such that if $x \in (a,b)$ is so that $|g(x) - g(a)| \leq \tau|g(b) - g(a)|$, then $|Dg(x)| \geq KL(\tau) \cdot |Dg(a)|$.*

LEMMA 2.3 (Expansion of Cross-Ratios): *Let g: $(a,b) \to \mathbb{R}$ be a diffeomorphism and $Sg \leq 0$. Let $j \subset t$ be subintervals of (a,b) and let l, r be the components of $t \setminus l$. Then*

$$\frac{|g(t)| \cdot |g(j)|}{|g(l)| \cdot |g(r)|} \geq \frac{|t| \cdot |j|}{|l| \cdot |r|}.$$

In the limit $j \to x$ this yields

$$g'(x) \geq \frac{|t|}{|g(t)|} \cdot \frac{|g(r)| \cdot |g(l)|}{|l| \cdot |r|}.$$

In the complex case we shall need

LEMMA 2.4 (Koebe Distortion Lemma in the complex case): *Assume that $\mathbb{D}$ is the unit disc in $\mathbb{C}$ and g: $\mathbb{D} \to \mathbb{C}$ is univalent, then for each $z \in \mathbb{D}$,*

$$\frac{1-|z|}{(1+|z|)^3} \leq \frac{|g'(z)|}{|g'(0)|} < \frac{1+|z|}{(1-|z|)^3}$$

and

$$\frac{1-|z|}{1+|z|} \le \frac{|z|\cdot|g'(z)|}{|g(z)-g(0)|} < \frac{1+|z|}{1-|z|}.$$

The proofs of Lemmas 2.1–2.3 can be found in [14]. Lemma 2.4 can be found in [19]: the first statement is formula (15) of Theorem 1.3, page 9, and the second follows by substituting the Koebe transform

$$h(z) = \left(g\left(\frac{z+w}{1+\bar{w}z}\right) - g(w)\right)/(1-|w|^2)g'(w)$$

in formula (14) and then taking $z=-w$; cf. Exercise 3 on page 13 in [19].

In order to deal with large disks, we need in fact a version of Lemma 2.4 applied to the Riemann surface $\hat{\mathbb{C}}$. Let d be spherical metric on $\hat{\mathbb{C}}$. (The results holds for any conformal metric, but we will use it for spherical metric only.) As usual define $B(x,r)=\{z; d(z,x)<r\}$. For an analytic map $f\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$, let $Df(z)$ be the derivative map with respect to the metric d. Using this notation, the previous lemma implies the following

LEMMA 2.5 (Koebe Distortion Lemma on the Riemann sphere): *Let d be the spherical metric on $\hat{\mathbb{C}}$ and let $r_0 < \operatorname{diam}(\hat{\mathbb{C}})$. Then there exists a constant $KL > 0$ such that for each $r < r_0$, each $x_0 \in \hat{\mathbb{C}}$, each univalent map $g\colon B(x_0,r) \to \hat{\mathbb{C}}$ and each $z \in B(x_0,r)$,*

$$\frac{1-d(z,x_0)/r}{KL} \le \frac{|Dg(z)|}{|Dg(x_0)|} \le \frac{KL}{(1-d(z,x_0)/r)^3},$$

and

$$\frac{1-d(z,x_0)/r}{KL} \le \frac{d(z,x_0)\cdot|Dg(z)|}{d(g(z),g(x_0))} \le \frac{KL}{1-d(z,x_0)/r}.$$

An interval $U \subset I$ is called **wandering** if $f^n|U$ is monotone for all $n \ge 0$ and U is not attracted to a periodic orbit. It is well-known that C^2 maps do not have wandering intervals; see [14]. This fact is equivalent to the following statement.

LEMMA 2.6: *For any $\epsilon > 0$ there exist $\delta > 0$ such that if U is a ball of diameter $\ge \epsilon$ which is not attracted to a periodic orbit, then $f^n(U)$ has diameter $\ge \delta$ for all $n \ge 0$.*

For rational maps on $\hat{\mathbb{C}}$ and open sets U, this lemma is a simple consequence of compactness and Montel's Theorem, provided U intersects the Julia set.

Let us finish this section by showing

LEMMA 2.7: *There exists a summable non-negative sequence (b_n) such that $b_n|Df^n(f(c))| \to \infty$ if and only if $\sum_n |Df^n(f(c))|^{-1} < \infty$.*

Proof: For the if-direction, choose $m_1, m_2, \ldots$ so that $\sum_{n \geq m_k} 1/|Df^n(f(c))| \leq 1/k^3$. Then take $d_n = 1$ for $n < m_1$ and $d_n = k$ for $n = m_k, \ldots, m_{k+1} - 1$ and $k = 1, 2, \ldots$. Then

$$\sum_{n \geq 0} d_n/|Df^n(f(c))| \leq \sum_{n \leq n_1} 1/|Df^n(f(c))| + \sum_{k \geq 1} k/k^3 < \infty$$

and so $b_n = \epsilon \cdot d_n/|Df^n(f(c))|$ satisfies the required properties for some small $\epsilon > 0$. Conversely, $b_n \cdot |Df^n(f(c))| \to \infty$ implies $|Df^n(f(c))|^{-1} \leq b_n$ for n sufficiently large. ∎

3. The real case

Throughout this section we shall assume that $f\colon I \to I$ is an interval or circle map satisfying (2). We write $|x - y|$ for the Euclidean distance between x and y; in case of sets, we prefer to use $\mathrm{dist}(X, Y) = \inf\{|x - y|; x \in X,\ y \in Y\}$. As before, let *Crit* be the set of critical points of f which are not in the basin of periodic attractors (because of the growth assumption f does not parabolic periodic points). For the moment fix an integer n, choose $\sigma > 0$ and take a point $x \in J_f$, such that $f^k(x) \notin Crit$ for all $k \leq n$. Let $T_n(x) \ni x$ be the maximal interval such that $f^n|T_n(x)$ is a diffeomorphism; usually we shall simply write T_n. Let

$$r_n(x) = \mathrm{dist}(\partial f^n(T_n), f^n(x)).$$

We say that n is σ-big time if $r_n(x) \geq \sigma$. Furthermore, n (or $f^n(x)$) is of type

$$\begin{cases} \text{(AP) } (\textit{almost pre-critical}) & \text{if } r_n(x) \geq \mathrm{dist}(f^n(x), Crit), \\ \text{(NAP) } (\textit{not almost pre-critical}) & \text{if } r_n(x) < \mathrm{dist}(f^n(x), Crit). \end{cases}$$

If n is a σ-big time, we still distinguish between type (AP) and (NAP). Types (NAP) and (AP) represent the "sliding" and "transfer" case in [18], while σ-big refers to the "regular" case. We call n **critical** if $f^n(T_n)$ contains a critical point. Note that if n is of type (AP), it has to be critical.

Now let $0 = n_s < n_{s-1} < \cdots < n_0 = n$ be the critical times before time $n_0 = n$ defined as follows (we do not need to assume that n itself is a critical time, but in any case define c_0 to be the (a) nearest critical point to $x_0 := f^{n_0}(x)$): Since T_n is the maximal neighbourhood of x on which f^n is diffeomorphic, there exist critical points c_1, c_1' and integers $n_1, n_1' > 0$ such that $f^n(T_n) = (f^{n-n_1}(c_1), f^{n-n_1'}(c_1'))$

and such that $|f^{n-n_1}(c_1) - f^n(x)| \le |f^{n-n_1'}(c_1') - f^n(x)|$. We should emphasize that it is possible that c_1 is a critical point in the basin of a periodic attractor (and so not in $Crit$), but then $f^n(T_n)$ contains a boundary point of the immediate basin of this periodic attractor, and then we take for c_1 this point (which is hyperbolic); in this case either c_1 is periodic or its f-image is periodic. The same modification for the choice of c_1 applies throughout the construction.

Next, for $T_{n_1}(x)$, the maximal neighbourhood of x on which f^{n_1} is diffeomorphic, we take the corresponding critical points $c_2, c_2' \in Crit$ and integers $n_2, n_2' > 0$, such that $|f^{n_1-n_2}(c_2) - f^{n_1}(x)| \le |f^{n_1-n_2'}(c_2') - f^{n_1}(x)|$, etc. See Figure 1. We shall write $x_i = f^{n_i}(x)$ for $i = 0, \dots, s$ and denote the order of the critical point c_i by ℓ_i. (We write x_0 for $f^n(x)$ since we will always 'pull-back'.)

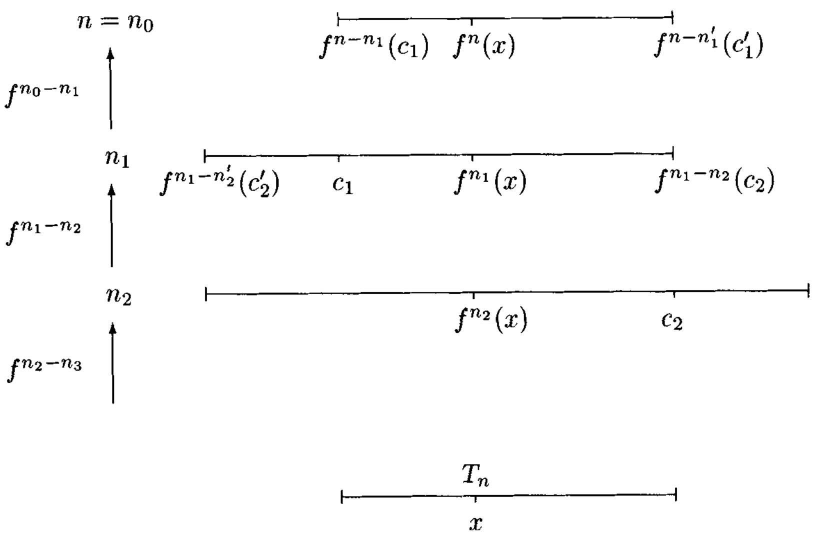

Figure 1. Critical times before n.

3.1. A PROPOSITION GIVING LOWER BOUNDS FOR EXPANSION OF DERIVATIVES. Let us first properly define the exponent $\ell(x)$. Let $Crit'(x)$ be the set of critical points $c \in Crit$ such that there are infinitely many (AP) times n such that $\text{dist}(f^n(x), Crit) = |f^n(x) - c|$. So for other critical points there is a bounded number of such events, and so for later use, we define $\tau(x) < \infty$ so that $n \le \tau(x)$ whenever n is a (AP) time with $\text{dist}(f^n(x), Crit'(x)) > \text{dist}(f^n(x), Crit)$. In Assertion 3 of Proposition 3.1 below, we will show that $Crit'(x) \neq \emptyset$ or there are infinitely many σ_0-big times n for some σ_0.

Now set

$$\ell = \ell(x) = \begin{cases} \max\{\ell_i; c_i \in Crit'(x)\} & \text{if } Crit'(x) \neq \emptyset, \\ 1 & \text{if } Crit'(x) = \emptyset. \end{cases} \tag{8}$$

This choice of ℓ enables the strongest asymptotic results; choosing $\ell = \ell_{\max}$, on the other hand, may give weaker asymptotic estimates, but allows for a x-independent choice of K' in Theorem 1.2.

For $0 \leq i, j \leq s$ define

$$A_i = \begin{cases} |x_i - f^{n_i - n_{i+1}}(c_{i+1})|^{\ell - \ell_i} & x_i \text{ is (AP)}, \\ |x_i - f^{n_i - n_{i+1}}(c_{i+1})|^{\ell - 1} |f^{n_i - n_{i+1}}(c_{i+1}) - c_i|^{1 - \ell_i} & x_i \text{ is (NAP)}, \end{cases}$$

and

$$B_j = \begin{cases} |x_j - c_j|^{\ell - \ell_j} & x_j \text{ is (AP)}, \\ |x_j - f^{n_j - n_{j+1}}(c_{j+1})|^{\ell - 1} \cdot |x_j - c_j|^{1 - \ell_j} & x_j \text{ is (NAP)}. \end{cases}$$

PROPOSITION 3.1: *Given f such that (2) holds, there exist $\sigma_0 > 0$, $\epsilon_0 > 0$, $K > 0$ and a function $E\colon \mathbb{N} \to \mathbb{R}$ with $E(k) \to \infty$ as $k \to \infty$, such that for each $x \in J_f$ and $n \in \mathbb{N}$ for which $x, \ldots, f^{n-1}(x) \notin Crit$ the following properties hold.*

1. *Let $0 = n_s < \cdots < n_1$ be the critical times before time $n_0 = n$. Let $\ell = \ell(x)$ and c_i, A_i, B_i be as above. Then for any $0 \leq i < j \leq s$,*

   $$|Df^{n_i - n_j}(x_j)|^\ell \geq \frac{A_i}{B_j} \prod_{k=i}^{j-1} (K_k \cdot |Df^{n_k - n_{k+1}}(f(c_{k+1}))|), \tag{9}$$

   where $K_0 = K$ and, for each $k = 1, \ldots, s-1$,

   $$K_k \geq \begin{cases} K \cdot [\mathrm{dist}(f^{n_k}(x), Crit)]^{\ell_k - \ell} & \text{if both } \ell_k > \ell \text{ and } n_k \text{ is of type (AP)}, \\ K & \text{otherwise.} \end{cases}$$

   By definition of $Crit'(x)$ the former can happen only finitely (uniformly in n) often: it can only happen when $n_k \leq \tau(x)$.
2. *If $n_{i+1} \geq \tau(x)$ then one of the following holds:*

   $$K \cdot |Df^{n_i - n_{i+1}}(f(c_{i+1}))| \geq 2, \tag{10}$$

   $$K^2 \cdot |Df^{n_i - n_{i+1}}(f(c_{i+1}))| \cdot |Df^{n_{i+1} - n_{i+2}}(f(c_{i+2}))| \geq 4 \tag{11}$$

   or (for any $j > i$)

   $$|Df^{n_i - n_j}(x_j)|^\ell \geq E(n_i - n_j) \cdot \frac{A_i}{B_j}.$$
3. *There exist arbitrarily large times n for which either $f^n(x)$ is in case (AP) or $r_n(x) \geq \sigma_0$ (so n is a σ_0-big time).*

4. *If $r_n(x) \geq \sigma_0$, then $A_0 \geq \sigma_0^{\ell-\ell_0}$. If $f^n(x)$ is in case (AP) and moreover the critical point nearest to $f^n(x)$ has order $\ell(x) = \ell$, then $A_0 = 1$.*
5. *If $\epsilon \in (0, \epsilon_0)$ and $n \geq 0$ is such that $f^n(x) \in N_\epsilon(Crit)$, but for no critical time $k < n$, $f^k(x) \in N_\epsilon(Crit)$, then $f^n(x)$ is in case (AP).*

In the remainder of this subsection we shall prove Proposition 3.1, using a few lemmas.

LEMMA 3.1 (Transfer Expansion): *There exists a universal constant $K > 0$ such that*

$$\frac{|Df^{n_0-n_1}(x_1)|^\ell}{|Df^{n_0-n_1}(f(c_1))|} \geq K \cdot \frac{A_0}{B_1}, \tag{13}$$

where

$$A_0 = \begin{cases} |x_0 - f^{n_0-n_1}(c_1)|^{\ell-\ell_0} & x_0 \textit{ in case (AP)}, \\ |x_0 - f^{n_0-n_1}(c_1)|^{\ell-1}|f^{n_0-n_1}(c_1) - c_0|^{1-\ell_0} & \textit{otherwise}, \end{cases}$$

and

$$B_1 = \begin{cases} |x_1 - c_1|^{\ell-\ell_1} & x_1 \textit{ in case (AP)}, \\ |x_1 - f^{n_1-n_2}(c_2)|^{\ell-1} \cdot |x_1 - c_1|^{1-\ell_1} & \textit{otherwise}. \end{cases}$$

Here, as before, ℓ_i is the order of the critical point c_i. In order to prove Lemma 3.1 we shall need the following

LEMMA 3.2: *There is a Koebe constant KL such that*

$$|Df^{n_0-n_1-1}(f(x_1))| \geq KL \cdot |Df^{n_0-n_1-1}(f(c_1))|, \tag{14}$$

$$|Df^{n_0-n_1-1}(f(x_1))| \geq KL \cdot \frac{|f^{n_0-n_1}(x_1) - f^{n_0-n_1}(c_1)|}{|f(x_1) - f(c_1)|}, \tag{15}$$

and

$$|Df^{n_0-n_1}(x_1)| \geq (1/2) \cdot \frac{|x_0 - f^{n_0-n_1}(c_1)|}{\min\{|x_1 - f^{n_1-n_2}(c_2)|, |x_1 - c_1|\}}. \tag{16}$$

Proof of Lemma 3.2: Since $x_0 = f^n(x)$ is nearer to $f^{n-n_1}(c_1)$ than to the other boundary point of $f^n(T_n)$, the one-sided Koebe lemma implies (14) and (15). Moreover, write $(f^{n_1-n_2}(c_2), f^{n_1-n_2'}(c_2')) = f^{n_1}(T_{n_1}) \ni c_1$. Choose either $t \subset (c_1, f^{n_1-n_2}(c_2))$ or $t \subset (c_1, f^{n_1-n_2'}(c_2'))$ maximal so that $f^{n_0-n_1}|t$ is a diffeomorphism and so that $x_1 \in t$. Let l, r be the components of $t \setminus \{x\}$. Note that if x_1 and $f^{n_1-n_2}(c_2)$ lie on the same side of c_1 then we take the former, and otherwise the latter. In both cases we have $\min(|l|, |r|) \leq \min\{|x_1 - c_1|, |c_1 - f^{n_1-n_2}(c_2)|\}$ because, by construction, $f^{n_1-n_2}(c_2)$ lies nearer to x_1 than $f^{n_1-n_2'}(c_2')$ and in the second case c_1 lies between x_1 and $f^{n_1-n_2}(c_2)$. Now we use Lemma 2.3 to derive the following inequality:

$$|Df^{n_0-n_1}(x_1)| \geq \frac{|f^{n_0-n_1}(l)| \cdot |f^{n_0-n_1}(r)|}{|f^{n_0-n_1}(t)|} \cdot \frac{|t|}{|l| \cdot |r|}.$$

Since at least one of the two intervals $f^{n_0-n_1}(l), f^{n_0-n_1}(r)$ has length $\geq |f^{n_0-n_1}(t)|/2$ and since $|t| \geq |l|, |r|$, one has that this last expression is at least

$$\geq (1/2) \cdot \min\{|f^{n_0-n_1}(l)|, |f^{n_0-n_1}(r)|\} \frac{1}{\min\{|l|, |r|\}}.$$

Because $f^{n_0-n_1}(c_1)$ is the end-point of $f^n(T_n)$ nearest to x_0, and the above estimate, this gives (16). ∎

Proof of Lemma 3.1: The proof follows from some simple algebra, from the non-flatness at critical points and inequalities (14)–(16). Indeed, distinguish two cases:

(AP) CASE: Let us first assume that $x_0 = f^n(x)$ is in case (AP). In this case, we first use the chain rule and (14) to get that $|Df^{n_0-n_1}(x_1)|^\ell$ is at least

$$KL \cdot |Df(x_1)|^\ell \cdot |Df^{n_0-n_1-1}(f(x_1))|^{\ell-1} \cdot |Df^{n_0-n_1-1}(f(c_1))|.$$

Now using non-flatness and applying (15) to the last but one factor, this is at least

$$K_1 \cdot |x_1 - c_1|^{(\ell_1-1)\ell} \cdot \frac{|f^{n_0-n_1}(c_1) - x_0|^{\ell-1}}{|f(c_1) - f(x_1)|^{\ell-1}} \cdot |Df^{n_0-n_1-1}(f(c_1))|.$$

Since we are in case (AP), this is at least

$$\geq \frac{K_1}{2^{\ell_0-1}} \cdot \frac{|f^{n_0-n_1}(c_1) - x_0|^{\ell-\ell_0} \cdot |f^{n_0-n_1}(c_1) - c_0|^{\ell_0-1}}{|x_1 - c_1|^{(1-\ell_1)\ell} \cdot |f(c_1) - f(x_1)|^{\ell-1}} \cdot |Df^{n_0-n_1-1}(f(c_1))|$$

and by non-flatness this gives

$$\frac{|Df^{n_0-n_1}(x_1)|^\ell}{|Df^{n_0-n_1}(f(c_1))|} \geq K \cdot \frac{|f^{n_0-n_1}(c_1) - x_0|^{\ell-\ell_0}}{|x_1 - c_1|^{\ell-\ell_1}}, \tag{17}$$

which is valid if x_1 is in case (AP) or in (NAP). If x_1 is in case (NAP), i.e., x_1 is closer to $f^{n_0-n_1}(c_2)$ than to c_1, then we can sharpen the above inequality: (16) can be written as

$$|Df^{n_0-n_1}(x_1)|^{\ell-1} \geq (1/2)^{\ell-1} \cdot \frac{|f^{n_0-n_1}(c_1) - x_0|^{\ell-\ell_0} \cdot |f^{n_0-n_1}(c_1) - x_0|^{\ell_0-1}}{|x_1 - f^{n_1-n_2}(c_2)|^{\ell-1}}.$$

Hence, if x_1 is in case (NAP),

$$|Df^{n_0-n_1}(x_1)|^\ell = |Df^{n_0-n_1}(x_1)|^{\ell-1} \cdot |Df^{n_0-n_1-1}(x_1)| \cdot |Df(x_1)|$$

can be estimated by using the previous inequality in the first factor, using (14) in the second factor, and by using that $|Df(f^{n_0-n_1}(c_1))| \approx |f^{n_0-n_1}(c_1) - c_0|^{\ell_0-1} \leq 2^{\ell_0-1}|f^{n_0-n_1}(c_1) - x_0|^{\ell_0-1}$ (because of the assumption on x_0). This gives

$$\text{(18)} \qquad \frac{|Df^{n_0-n_1}(x_1)|^\ell}{|Df^{n_0-n_1}(f(c_1))|} \geq K \cdot \frac{|f^{n_0-n_1}(c_1) - x_0|^{\ell-\ell_0}}{|x_1 - f^{n_1-n_2}(c_2)|^{\ell-1}|x_1 - c_1|^{1-\ell_1}}.$$

(NAP) CASE: Let us now assume that $x_0 = f^{n_0}(x)$ is in case (NAP). In this case we first use the chain rule, non-flatness, and (14) to get that $|Df^{n_0-n_1}(x_1)|^\ell$ is at least a universal constant times

$$|Df^{n_0-n_1}(x_1)|^{\ell-1} \cdot |Df^{n_0-n_1}(f(c_1))| \cdot \frac{|x_1 - c_1|^{\ell_1-1}}{|f^{n_0-n_1}(c_1) - c_0|^{\ell_0-1}}.$$

Applying (16) to the first factor gives that this is at least a universal constant times

$$\left(\frac{|x_0 - f^{n_0-n_1}(c_1)|}{\min\{|x_1 - f^{n_1-n_2}(c_2)|, |x_1 - c_1|\}}\right)^{\ell-1} \cdot \frac{|x_1 - c_1|^{\ell_1-1}}{|f^{n_0-n_1}(c_1) - c_0|^{\ell_0-1}} \cdot |Df^{n_0-n_1}(f(c_1))|.$$

That is,

$$\text{(19)} \qquad \frac{|Df^{n_0-n_1}(x_1)|^\ell}{|Df^{n_0-n_1}(f(c_1))|} \geq K \cdot \frac{|x_0 - f^{n_0-n_1}(c_1)|^{\ell-1} \cdot |f^{n_0-n_1}(c_1) - c_0|^{1-\ell_0}}{[\min(|x_1 - f^{n_1-n_2}(c_2)|, |x_1 - c_1|)]^{\ell-1} \cdot |x_1 - c_1|^{1-\ell_1}}.$$

If x_1 is as in case (AP) then this becomes

$$\text{(20)} \qquad \frac{|Df^{n_0-n_1}(x_1)|^\ell}{|Df^{n_0-n_1}(f(c_1))|} \geq K \cdot \frac{|x_0 - f^{n_0-n_1}(c_1)|^{\ell-1} \cdot |f^{n_0-n_1}(c_1) - c_0|^{1-\ell_0}}{|x_1 - c_1|^{\ell-\ell_1}},$$

whereas if x_1 is as in case (NAP) then this becomes

$$\text{(21)} \qquad \frac{|Df^{n_0-n_1}(x_1)|^\ell}{|Df^{n_0-n_1}(f(c_1))|} \geq K \cdot \frac{|x_0 - f^{n_0-n_1}(c_1)|^{\ell-1} \cdot |f^{n_0-n_1}(c_1) - c_0|^{1-\ell_0}}{|x_1 - f^{n_1-n_2}(c_2)|^{\ell-1} \cdot |x_1 - c_1|^{1-\ell_1}}.$$

This completes the proof of Lemma 3.1. ∎

Proof of Assertion 1 of Proposition 3.1: The proof of this part is formal, and is the same for the real and the complex case. For $Df^{n_1-n_2}(x_2)$ the analogous statement to (13) holds, replacing $x_0, c_0, c_1, n_0 - n_1, x_1$ by $x_1, c_1, c_2, n_1 - n_2, x_2$. Let us now analyse what we can say about

$$|Df^{n_0-n_2}(x_2)| = |Df^{n_0-n_1}(x_1)| \cdot |Df^{n_1-n_2}(x_2)| \geq K^2 \frac{A_0}{B_1}\frac{A_1}{B_2}$$

by considering the term A_1/B_1. We distinguish several cases.

$$\text{If } x_1 \text{ in case (AP) and } \ell \geq \ell_1, \text{ then}$$

$$\frac{A_1}{B_1} = \frac{|f^{n_1-n_2}(c_2) - x_1|^{\ell-\ell_1}}{|x_1 - c_1|^{\ell-\ell_1}} \geq 1. \tag{22}$$

Also,

$$x_1 \text{ in case (AP) and } \ell < \ell_1 \text{ implies}$$

$$\frac{A_1}{B_1} = \frac{|f^{n_1-n_2}(c_2) - x_1|^{\ell-\ell_1}}{|x_1 - c_1|^{\ell-\ell_1}} \geq [\text{dist}(x_1, Crit)]^{\ell_1-\ell}. \tag{23}$$

Moreover,

$$x_1 \text{ in case (NAP) implies}$$

$$\frac{A_1}{B_1} = \frac{|x_1 - f^{n_1-n_2}(c_2)|^{\ell-1} \cdot |f^{n_1-n_2}(c_2) - c_1|^{1-\ell_1}}{|x_1 - f^{n_1-n_2}(c_2)|^{\ell-1} \cdot |x_1 - c_1|^{1-\ell_1}} \geq 2^{1-\ell_1}. \tag{24}$$

Because of these inequalities and similar ones for A_k/B_k for $k = 2, \ldots, s-1$,

$$|Df^{n_i-n_j}(x_j)|^\ell \geq \frac{A_i}{B_j} \prod_{k=i}^{j-1} K_k \cdot |Df^{n_k-n_{k+1}}(f(c_k))|, \tag{25}$$

where $K_k := K \cdot A_k/B_k$ are bounded below in the way described in the statement of Assertion 1 of the proposition. (So for all but a finite number, $K_k \geq K2^{1-\ell}$ is uniformly bounded away from 0.) Note that we have not used the assumption about the growth rate of derivatives or any assumption on x to obtain this inequality. ∎

Proof of Assertion 2 of Proposition 3.1: We need to show that (10), (11) or (12) holds. Take $\rho > 0$ so small that $K \cdot |Df^k(f(c))| \geq 2$ for each $k \geq \nu(\rho)$ and each $c \in Crit$. Assume that the first alternative (10) fails, i.e., assume $K \cdot |Df^{n_i-n_{i+1}}(f(c_{i+1}))| < 2$. Then $n_i - n_{i+1} \leq \nu(\rho)$. Hence $|c_i - f^{n_i-n_{i+1}}(c_{i+1})| \geq \rho$ and so either $|x_i - f^{n_i-n_{i+1}}(c_{i+1})| \geq \rho/2$ or $|x_i - c_i| \geq \rho/2$.

If $|x_i - f^{n_i-n_{i+1}}(c_{i+1})| \geq \rho/2$, then, for any $j > i$, $f^{n_i-n_j}(T_{n_i-n_j}(x_j))$ contains a $\rho/2$-neighbourhood of x_i. Let t denote the neighbourhood of x_j for which $f^{n_i-n_j}(t)$ is equal to a $\rho/4$-neighbourhood of x_i. Then using the Koebe Lemma 2.1,

$$|Df^{n_i-n_j}(x_j)|^\ell \geq KL \cdot \frac{(\rho/4)^\ell}{\operatorname{diam}(t)^\ell} \geq KL' \cdot \frac{A_i}{\operatorname{diam}(t)^\ell}. \tag{26}$$

Let $t_\pm$ be the components of $t \setminus \{x_j\}$. Since there is absolute Koebe space of order $\rho/4$ around $f^{n_i-n_j}(t)$, the quotient $|t_+|/t_-|$ is universally bounded from below and above. Because t is contained in $(c_j, f^{n_j-n_{j+1}}(c_{j+1}))$,

$$\operatorname{diam}(t) \leq C \cdot \min(|x_j - c_j|, |x_j - f^{n_j-n_{j+1}}(c_{j+1})|),$$

for a universal constant C. Hence, since $\ell_j > 1$, the right hand side of (26) is much larger than A_i/B_j when t is small. Since f has no wandering intervals, t is small when $n_i - n_j$ is large. It follows that there exists a function $E'_\rho\colon \mathbb{N} \to \mathbb{R}$ with $\lim_{k\to\infty} E'_\rho(k) = \infty$ such that

$$|x_i - f^{n_i-n_{i+1}}(c_{i+1})| \geq \rho/2 \quad \text{implies} \quad |Df^{n_i-n_j}(x_j)|^\ell \geq E'_{\rho/2}(n_i - n_j)\frac{A_i}{B_j} \tag{27}$$

and so (12) holds.

So we may assume that $|c_i - x_i| \geq \rho/2$ and $|f^{n_i-n_{i+1}}(c_{i+1}) - x_i| < \rho/2$. Therefore there exists $\rho' = \rho'(\rho)$ such that $|x_{i+1} - c_{i+1}| < \rho'$ and $\rho' \to 0$ as $\rho \to 0$.

Let us first consider the case that $|f^{n_{i+1}-n_{i+2}}(c_{i+2}) - x_{i+1}| \geq \rho'$ (so in particular x_{i+1} is in case (AP)). Then we can repeat the previous argument (replacing $\rho/2$ by ρ') and obtain that

$$|Df^{n_{i+1}-n_j}(x_j)|^\ell \geq E''_{\rho'}(n_{i+1} - n_j)\frac{A_{i+1}}{B_j},$$

for some appropriate function $E''_{\rho'}$. Since x_{i+1} is in case (AP), we find by (22) and Lemma 3.1 that $|Df^{n_i-n_{i+1}}(x_{i+1})|^\ell \geq KA_i/B_{i+1} \geq K$ because we have assumed that $n_{i+1} \geq \tau(x)$ and so $\ell_{i+1} \leq \ell$. Combined, this gives

$$\begin{aligned} |f^{n_{i+1}-n_{i+2}}(c_{i+2}) - x_{i+1}| \geq{}& \rho' \quad \text{implies} \\ |Df^{n_i-n_j}(x_j)|^\ell \geq{}& K \cdot E''_{\rho'}(n_i - n_j)\frac{A_i}{B_j}, \end{aligned} \tag{28}$$

and so again (12) holds.

The final case is when $|f^{n_{i+1}-n_{i+2}}(c_{i+2}) - x_{i+1}| < \rho'$, which implies that $|f^{n_{i+1}-n_{i+2}}(c_{i+2}) - c_{i+1}| < 2\rho'$. Therefore $n_{i+1} - n_{i+2} \geq \nu(2\rho')$, and if ρ' is sufficiently small,

$$|Df^{n_{i+1}-n_{i+2}}(f(c_{i+2}))| \geq \frac{4}{K^2}|Df^{n_i-n_{i+1}}(f(c_{i+1}))|^{-1}.$$

This gives (11), and completes the proof of Assertion 2 of Proposition 3.1. ∎

Proof of Assertions 3–5 of Proposition 3.1: Let us first prove Assertion 3. Take $\sigma_0 > 0$ so that for each critical point $c \notin Crit$, $\mathrm{dist}(\omega(c), J(f)) > \sigma_0$. Moreover, let $0 < \sigma_0 < \sigma_0' < 1$ be so that

$$\inf\{|Df^n(f(c))|; n \geq 1, c \in Crit\} \cdot {\sigma_0'}^{\ell_{\max}} > K\sigma_0, \tag{29}$$

and so that $n_{i-1} - n_i \geq \nu({\sigma_0'}^{1/\ell_{\max}})$ implies that $|Df^{n_{i-1}-n_i}(f(c_i))| \geq K4^{\ell_{\max}}$. Here $K = KL \cdot M$ is a common Koebe/non-flatness constant from the two inequalities below.

Assume by contradiction that Assertion 3 does not hold, and that there exists N_0 such that for arbitrarily large $n_0 = n$ and all critical times $n_j \geq N_0$ before n_0, both $r_{n_j} \leq \sigma_0$ **and** n_j is of type (NAP). Let us first show that this implies that $|x_i - c_i| \geq \sigma_0'$ whenever $n_i \geq N_0$ and $i > 0$. Indeed, assume by contradiction that $|x_i - c_i| < \sigma_0'$. We apply the One-sided Koebe Lemma 2.2 to the appropriate branch of $f^{-(n_{i-1}-n_i-1)}$ that maps a maximal neighbourhood of x_{i-1} diffeomorphically onto a neighbourhood of $f(x_i)$. Using the non-flatness of f at x_i and since n_{i-1} is of type (NAP), we get

$${\sigma_0'}^{l_i} \leq M\ d(f(c_i), f(x_i)) \leq K \cdot \frac{\sigma_0}{|Df^{n_{i-1}-n_i-1}(f(c))|}, \tag{30}$$

contradicting (29).

Now take a large integer $n = n_0$ and critical times $n_{j_2} < n_{j_1}$ before n_0 so that $n_{j_1} - n_{j_2}$ is large. Let σ_{n_j} be the maximum of the distances between the points $f^{n_j}(x) = x_j$, $f^{n_j - n_{j+1}}(c_{j+1})$ and c_j. By the One-sided Koebe Lemma 2.2, there exists a universal constant KL such that

$$\frac{1}{M}|x_i - c_i|^{\ell_i} \leq KL \cdot \frac{|x_{i-1} - f^{n_{i-1}-n_i}(c_i)|}{|Df^{n_{i-1}-n_i-1}(f(c_i))|} \tag{31}$$

and this can be bounded from above by

$$\begin{aligned} \leq & KL \cdot \frac{|x_{i-1} - f^{n_{i-1}-n_i}(c_i)|}{|Df^{n_{i-1}-n_i-1}(f(c_i))|} \cdot \frac{|f^{n_{i-1}-n_i}(c_i) - c_{i-1}|^{\ell_{i-1}-1}}{|Df(f^{n_{i-1}-n_i}(c_i))|} \\ = & KL \cdot \frac{|x_{i-1} - f^{n_{i-1}-n_i}(c_i)| \cdot |f^{n_{i-1}-n_i}(c_i) - c_{i-1}|^{\ell_{i-1}-1}}{|Df^{n_{i-1}-n_i}(f(c_i))|}. \end{aligned}$$

Hence, for any $0 < i \leq s$ and $K = KL \cdot M$,

$$|x_i - c_i|^{\ell_i} \leq \frac{|x_{i-1} - f^{n_{i-1}-n_i}(c_i)| \cdot |c_{i-1} - f^{n_{i-1}-n_i}(c_i)|^{\ell_{i-1}-1}}{|Df^{n_{i-1}-n_i}(f(c_i))|/K}. \tag{32}$$

Next assume that σ_0 and hence σ_0' is so small that $n_{i-1} - n_i \geq \nu({\sigma_0'}^{1/\ell_{\max}})$ implies that the denominator $|Df^{n_{i-1}-n_i}(f(c_i))|/K \geq 4^{\ell_{\max}}$. Then, provided $|x_i - c_i|^{\ell_i} < \sigma_0'$, and since x_i is in case (NAP), (32) gives

$$\sigma_{n_i} \leq 2d(x_i, c_i) \leq \lambda_i \cdot \sigma_{n_{i-1}}^{\ell_{i-1}/\ell_i}, \tag{33}$$

where $\lambda_i \leq 1/2$ and $\lambda_i \leq \xi(n_i - n_{i-1})$ is a function which tends to zero when $n_{i-1} - n_i$ tends to infinity. If, on the other hand, $|x_i - c_i|^{\ell_i} \geq \sigma_0'$, then since no critical time n_j, $N_0 \leq n_j < n_0$ is σ_0-large, we get $|x_{i-1} - f^{n_{i-1}-n_i}(c_i)| = r_{n_{i-1}} < \sigma_0$. So (32) would give $|Df^{n_{i-1}-n_i}(f(c_i))| \cdot M \cdot \sigma_0' \leq KL \cdot \sigma_0 \cdot \sigma_{i-1}^{\ell_i - 1}$, contradicting (29). So (33) holds in all cases and therefore

$$\sigma_{n(j_1)} \leq \left(\prod_{j=j_1+1}^{j_2} \lambda_j \right) \cdot \sigma_{n(j_2)}^{\ell_{j_2}/\ell_{j_1}}.$$

But since we may assume that $\sigma_{n(j_2)}$ is some given positive number (with $n(j_2) \geq N_0$), we get a contradiction by choosing $n(j_1) - n(j_2)$ arbitrarily large.

The same argument also proves Assertion 5. Indeed, assume that $n = n_0$ is the smallest time with $f^n(x) \in N_\epsilon(Crit)$ and assume by contradiction that $x_0 = f^n(x)$ is in case (NAP). Inequality (32) implies that

$$|x_1 - c_1|^{\ell_1} \cdot \frac{|Df^{n_0-n_1}(f(c_1))|}{KL} \leq |x_0 - f^{n_0-n_1}(c_1)| \cdot |c_0 - f^{n_0-n_1}(c_1)|^{\ell_0 - 1}$$

and since x_0 is assumed to be in case (NAP) this is at most

$$2^{\ell_0-1}|x_0 - c_0|^{\ell_0}.$$

Either $|x_1 - c_1|^{\ell_1} < \sigma_0'$ and then (provided σ_0' is small) $|x_1 - c_1|^{\ell_1} \leq |x_0 - c_0|^{\ell_0}$, contradicting that x_0 is the first entry in $N_\epsilon(Crit)$. Or $|x_1 - c_1|^{\ell_1} > \sigma_0'$, but then $|x_0 - c_0| > \sigma_0$; so if we fix $\epsilon < \sigma_0^{\ell_{\max}}$, then x_0 does not belong to $N_\epsilon(Crit)$.

Assertion 4 of Proposition 3.1 follows from the formulas for A_0. ∎

3.2. The proofs of Theorems 1.2, 1.3 and Corollary 1.3 in the real case. Before proving Theorem 1.2, let us make a remark on the role of the λ in formula (5). It is this part that assures that $\limsup_n Df^n(x) = \infty$ for all $x \in J_f$ that are not (pre)critical. The difficulty in proving it is not in very close visits of x to $Crit$, because then the corresponding factor $K_k \cdot |Df^{n_k - n_{k+1}}|$ (from Proposition 3.1) is large, but rather in visits that are intermediately close. In this case, the time $n_k - n_{k+1}$ in the factors $K_k \cdot |Df^{n_k - n_{k+1}}|$ may be too small to guarantee expansion. To remedy this, we need to string several intermediate visits together with a single distortion constant. Assertion 2 in Proposition 3.1 is essential here.

Proof of Theorem 1.2 in the real case: Let us first show that there are arbitrarily large $n = n_0$ such that the corresponding A_0 is uniformly bounded away from 0. By Assertion 3 of Proposition 3.1, there are infinitely many times n which are of case (AP) or for which $f^n(T_n)$ contains a σ_0-neighbourhood of $f^n(x)$. If the latter holds, taking $n_0 = n$ arbitrarily large, Assertion 4 of Proposition 3.1 gives that $A_0 \geq \sigma_0^{\ell_{\max}-1} > 0$ independently of n. If the former holds, then $\ell = \ell(x)$ is defined to be the largest order of a critical point which is involved with infinitely many critical times of type (AP). So we can take n_0 arbitrarily large such that $\ell_0 = \ell$. This implies again by Assertion 4 of Proposition 3.1 that $A_0 = 1$.

So assume that $n = n_0$ is picked as above. In order to prove Theorem 1.2 we shall use (9) in a suitable way. The reason this can be done is because according to (22)–(25) (from the first part of the proof of Proposition 3.1), $A_j \geq B_j$ unless $\ell_j > \ell$ and x_j is case (AP). So if $n_j \geq \tau(x)$ then, in view of the definitions of $Crit'(x)$ and $\ell(x)$ above the statement of Proposition 3.1, we always have $A_j \geq B_j$. Moreover, in Assertion 1 of Proposition 3.1, $K_j \geq K$ whenever $n_j \geq \tau(x)$. If n_j is of type (AP) and $n_j < \tau(x)$, then we have to include a factor $\text{dist}(f^{n_j}(x), Crit)^{\ell_j - \ell}$ to estimate the factor A_j/B_j, but this happens only finitely often. These factors are included in the factor $K'(x)$.

The term B_s in (9) is treated as follows: since n_s is of type (AP) by "default", we have in the worst case that $B_s = |x - c_s|^{\ell - \ell_s}$. This constitutes a single large denominator (if $\ell_s > \ell$), which is represented in the first factor $K'(x)$.

Let us now regroup the critical times n_i as suggested by Assertion 2 of Proposition 3.1. First we define $\lambda > 1$ as follows. Take N_0 such that $E(k) > 2$ for each $k \geq N_0$. Here E is the function from Proposition 3.1 (which does not depend on x). So $\lambda := 2^{1/N_0} > 1$ does not depend on x. Next define

$$k = \begin{cases} 1 & \text{if } K \cdot |Df^{n_0-n_1}(f(c_1))| \geq 2, \\ 2 & \text{if } K^2 \cdot |Df^{n_0-n_1}(f(c_1))| \\ & \quad \cdot |Df^{n_1-n_2}(f(c_2))| \geq 4, \\ \min\{j; n_0 - n_j \geq N_0\} & \text{otherwise.} \end{cases}$$

By Assertion 2 of Proposition 3.1 and the choice of N_0, we have $k \leq N_0$ and

$$|Df^{n_0-n_k}(x_k)|^{\ell(x)} \geq 2A_0/B_k,$$

and by Assertion 1 of Proposition 3.1,

$$|Df^{n_0-n_k}(x_k)|^{\ell(x)} \geq \frac{A_0}{B_k} \prod_{i=0}^{k-1} K \cdot D_{n_i - n_{i+1}}(c_{i+1}).$$

Now we need the following

CLAIM: *Assume that* $\epsilon \in (0,1)$, $a_1, \ldots, a_k > \epsilon$ *and* $A \geq \max\{2, \prod_{i=1}^k a_i\}$. *Then* $A \geq \prod_{i=1}^k \max\{2^{1/k}, \epsilon^k a_i/2\}$.

Proof of the Claim: We may assume that $a_1 > \cdots > a_k$ and that $j \leq k$ is maximal so that $\epsilon^k a_j/2 > 2^{1/k}$. If there is no such j, then the maximum above is always equal to $2^{1/k}$ and the required inequality reduces to $A \geq 2$. Otherwise we have

$$\begin{aligned} A \geq a_1 \cdots a_k \geq & a_1 \cdots a_j \cdot \epsilon^{k-j} \\ \geq & (\epsilon^k a_1/2) \cdots (\epsilon^k a_j/2) \cdot 2^{(k-j)/k} \\ = & \prod_{i=1}^k \max(2^{1/k}, \epsilon^k a_i/2) \end{aligned}$$

(where the last equality holds by the choice of j), completing the proof of the claim.

Now let

$$A = |Df^{n_0 - n_k}(x_k)|^{\ell(x)} \cdot \frac{B_k}{A_0}.$$

By (24), (25) and the choice of k we have that $A \geq 2$ and also that $A \geq \prod_{i=0}^{k-1} a_i$ where $a_i = \min\{K, 2^{1-\ell(x)}\} \cdot |Df^{n_i - n_{i+1}}(c_{i+1})|$. Applying the claim gives

$$|Df^{n_0 - n_k}(x_k)|^{\ell(x)} = A \frac{A_0}{B_k} \geq \frac{A_0}{B_k} \prod_{i=0}^{k-1} \max(\lambda, \tilde{K} \cdot D_{n_i - n_{i+1}}(c_{i+1})),$$

where $\tilde{K} = \min\{K, 2^{1-\ell(x)}\} \cdot \epsilon^k/2$ is a new universal constant and λ is as above. Repeat this construction for a new $k' > k$, etc., as long as $n_{k+1} \geq \tau(x)$. Since $\tau(x)$ is bounded, we exhaust all but a finite number of critical times. Combining all this, we obtain

$$|Df^n(x)|^{\ell(x)} \geq K'(x) \cdot \frac{A_0}{B_s} \prod_{i=0}^{s-1} \max(\lambda, \tilde{K} \cdot D_{n_i - n_{i+1}}(c_{i+1})),$$

where the $K'(x)$ takes care of the estimates before time $\tau(x)$. Rename $\tilde{K}$ back to K and Theorem 1.2 is proved. ∎

Proof of Theorem 1.3 in the real case: From Assertion 4 of Proposition 3.1 and the definition of n, it follows that $f^n(x)$ is in case (AP). From Assertion 3 of Proposition 3.1 and the assumption that all critical points have the same order, it follows that $A_0 = 1$. Moreover, B_s is bounded from above. Hence (6) follows from Assertion 1 of Proposition 3.1. Since by assumption all critical points have the same order, $K_k \geq K$ for all k and so Theorem 1.3 follows. ∎

Proof of Corollary 1.3: From Proposition 3.1, with $\ell = \ell_{\max}$, we get

$$|Df^n(x)|^{\ell_{\max}} \geq \frac{A_0}{B_s} \prod_{i=0}^{s-1} K \cdot D_{n_k - n_{k+1}}(c_{k+1}).$$

Using (10)–(12) and the argument in the proof of Theorem 1.2, we obtain

$$|Df^n(x)|^{\ell_{\max}} \geq C \cdot \frac{A_0}{B_s} \cdot \lambda^s, \tag{34}$$

where C is a constant that is only needed if $n \leq m_0$ with m_0 as in the proof of Theorem 1.2. Due to Lemma 2.6, there exists $\delta = \delta(\sigma)$ such that $|f^n(T_n)| \geq \delta$ for all n. It follows that there exists $L = L(\sigma)$ such that two subsequent critical times are no more than L apart.

We have $B_s = |x - c_s|^{\ell_{\max} - \ell_s} \leq 1$ and, assuming the worst case that n_0 is (NAP), $A_0 \geq |x_0 - f^{n_0 - n_1}(c_1)|^{\ell_{\max} - 1}$. Since $f^{n_0 - n_1}(c_1)$ is the end-point of $f^n(T_n)$ closest to x_0, we can use the One-sided Koebe Lemma and non-flatness to obtain

$$\begin{aligned} |x_0 - f^{n_0 - n_1}(c_1)| \geq& \frac{1}{KL} |Df^{n_0 - n_1 - 1}(f(c_1))| \cdot |f(x_1) - f(c_1)| \\ \geq& \frac{1}{KL \cdot M} |Df^{n_0 - n_1 - 1}(f(c_1))| \cdot |x_1 - c_1|^{\ell_1} \\ \geq& C_0 \sigma^{\ell_{\max}}, \end{aligned} \tag{35}$$

for some $C_0 > 0$ independent of x. Using this estimate, inequality (34) becomes

$$|Df^n(x)| \geq C \sigma^{\ell_{\max} - 1} \lambda^{s/\ell_{\max}} \geq C \sigma^{\ell_{\max} - 1} \lambda^{n/L\ell_{\max}}$$

for the appropriate $C = C(C_0)$. This proves the corollary.

Note that if f satisfies (CE), then it is not hard to prove that λ can be chosen independently from σ. ∎

4. The complex case

In this section we assume that $f\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ is a rational map. Let d be spherical metric on $\hat{\mathbb{C}}$, so Lemma 2.5 holds. Also assume that there exists a sequence $b_n > 0$ such that

$$\exp\Big(\sum_{n \geq 0} b_n\Big) < 2 \quad \text{and} \quad b_n \cdot |Df^n(f(c))| \to \infty \tag{36}$$

for each $c \in Crit$. The reason that we need a stronger assumption than the corresponding assumption (2) in the real case is that the One-sided Koebe Lemma 2.2 does not hold in the complex case.

Let again $N_\epsilon(Crit)$ be defined as in (1). Take a point x such that $f^k(x) \notin Crit$ for all $k \geq 0$, fix a large integer n, write $n_0 = n$ and let $x_0 = f^n(x)$. Our aim is to estimate $Df^n(x)$ or at least $Df^m(x)$ for some large $m < n$. To do this, we replace the maximal intervals of monotonicity $T_n(x)$ and $f^n(T_n(x))$ from the previous section by suitable disks (shrinking neighbourhoods in the terminology of [21]) using the sequence b_n from (36). For $m \leq n$, let $s_m > 0$ be maximal so that the inverse branch of $f^{-(n-m)}$ from $B(x_0, s_m)$ to a neighbourhood of $f^m(x)$ is univalent. Maximality of s_m means that for some $k = 1, 2, \dots, n-m$ the component W_{n-k} of $f^{-k}(B(x_0, s_m))$ containing $f^{n-k}(x)$ contains a critical point c in its boundary, and $s_m = d(x_0, f^k(c))$. Of course,

$$s_n > s_{n-1} \geq \cdots \geq s_0.$$

Let $\tilde{n}_t < \tilde{n}_{t-1} < \cdots < \tilde{n}_0 := n$ be the collection of integers $\tilde{n}_j \in \{0, \dots, n\}$ for which $s_{\tilde{n}_j} < s_{\tilde{n}_j+1}$ when $j > 0$ and define $\tilde{n}_{t+1} = 0$. The set $W_{\tilde{n}_j+1}$ contains a critical value of f in its interior, whereas $f(W_{\tilde{n}_j}) \subset W_{\tilde{n}_j+1}$ contains this critical value only in its boundary.

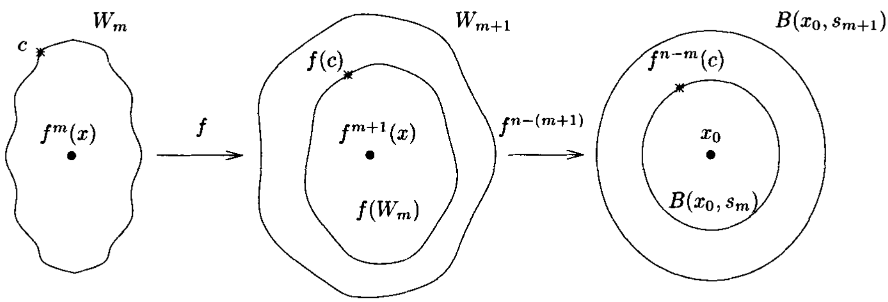

Figure 2. Shrinking neighbourhoods and their pre-images.

Consider all the integers $\tilde{n}_j < n$, $j = 1, \dots, t, t+1$, as above such that

$$\frac{s_{\tilde{n}_{j-1}}}{s_{\tilde{n}_j}} \geq \exp(b_{n-\tilde{n}_j}). \tag{37}$$

Let $n_1 < n$ be the smallest such integer $\tilde{n}_j$ (i.e., with the largest index j) and define $r_{n_0} = s_{n_1}$. By construction f^{n-n_1} maps a neighbourhood W of $f^{n_1}(x)$

univalently onto $B_{n_0} := B(x_0, r_{n_0})$. Moreover, W contains a critical point c_1 in its boundary, and $f(W)$ is contained in a set W' which is mapped univalently by f^{n-n_1-1} onto the disc $B(x_0, \exp(b_{n-n_1}) \cdot r_{n_0})$. Note that it is possible that c_1 is a critical point which is not in $Crit$, i.e., is in the immediate basin of a hyperbolic periodic attractor. Since $x \in J(f)$, in that case W intersects the boundary of this immediate basin and so we can replace c_1 by a repelling periodic point in this boundary. One can choose this repelling periodic point so that its period is comparable to the period of the periodic attractor, i.e., at most some number $C(f)$ times the period of the attractor (one can choose C so that it does not depend on x.) Throughout the remainder of the construction below we will use the same modification, if required. If c_1 is a critical point, then let ℓ_1 be its order, and if it is a repelling periodic orbit, let $\ell_1 = 1$.

CLAIM 1: *$s_0 \geq s_{n_1}/2 = r_{n_0}/2$ and the branch of f^{-n} mapping $B(x_0, r_{n_0})$ to a neighbourhood of x is univalent on $B(x_0, r_{n_0}/2)$.*

Indeed, by minimality of $n_1 = \tilde{n}_j$ one has $s_{\tilde{n}_i}/s_{\tilde{n}_{i+1}} \leq \exp(b_{n-\tilde{n}_{i+1}})$, for all $i \geq j$. But since $s_0 = s_{\tilde{n}_{t+1}}$, this implies that $s_{n_1}/s_0 = s_{\tilde{n}_j}/s_{\tilde{n}_{t+1}} \leq \exp(\sum b_i) \leq 2$. So $s_0 \geq s_{n_1}/2 = r_{n_0}/2$ and since by definition f^{-n} is univalent on $B(x_0, s_0)$ the claim follows.

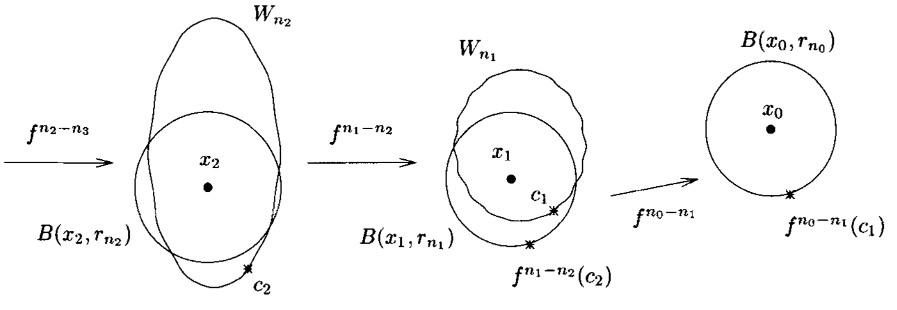

Figure 3. Construction of critical times and the discs $B_{n_j} = B(x_j, r_{n_j})$.

We say that n_1 is a **critical time before time** $n = n_0$. Now let $0 = n_s < n_{s-1} < \cdots < n_1 < n_0 = n$ where for $j = 0, \ldots, s$ the integer n_j is the critical time before n_{j-1}. For each $x_j := f^{n_j}(x)$, let $B_{n_j} = B(x_j, r_{n_j})$ be the disc defined above. By construction we have the following properties:

1. $f^{n_j - n_{j+1}}$ maps a neighbourhood $W_{n_{j+1}}$ of $f^{n_{j+1}}(x)$ univalently onto $B_{n_j} = B(f^{n_j}(x), r_{n_j})$;

2. $W_{n_{j+1}}$ contains a critical point c_{j+1} in its boundary. As mentioned above, we can assume that c_{j+1} either belongs $Crit$ or can be replaced by a repelling periodic point in the boundary of the immediate basin of a hyperbolic periodic attractor.
3. $f(W_{n_{j+1}})$ is contained in a set W' which is mapped univalently by $f^{n_j - n_{j+1} - 1}$ onto the disc $B(f^{n_j}(x), \exp(b_{n_j - n_{j+1}}) \cdot r_{n_j})$;

Note that there are some differences with the real case: $B_{n_{j+1}}$ need not be contained in $W_{n_{j+1}} \subset f^{-(n_j - n_{j+1})}(B_{n_j})$, whereas in the real case $f^{n_{j+1}}(T_{n_{j+1}}) \subset f^{n_{j+1}}(T_{n_j})$. Moreover, B_{n_j} need not contain a critical point, whereas in the real case $f^{n_j}(T_{n_j})$ always contains a critical point (except possibly for n_0). However, neither of these special properties, which only hold in the real case, play an essential role in the proofs as we will show below. The only crucial difference is that (in the complex case), the critical times $n_s < n_{s-1} < \cdots < n_0$ are only defined once the integer $n_0 = n$ is given.

We say that n_j is of type

$$\begin{cases} \text{(AP) } (\textit{almost pre-critical}) & \text{if } r_{n_j}(x) \geq \operatorname{dist}(x_j, Crit), \\ \text{(NAP) } (\textit{not almost pre-critical}) & \text{if } r_{n_j}(x) < \operatorname{dist}(x_j, Crit). \end{cases}$$

Types (NAP) and (AP) represent the "sliding" and "transfer" case in [18]. Let $Crit'(x)$ be the set of critical points for which there are infinitely many integers n_j of type (AP), and let $\tau(x) < \infty$ be so that $n \leq \tau(x)$ whenever n is a (AP) time corresponding to a critical point not in $Crit'(x)$. In Assertion 3 of Proposition 4.1 below, we shall show that there exists $\sigma_0 > 0$ independently of x such that $Crit'(x) \neq \emptyset$ or there are infinitely many σ_0-big times n. Let

$$\ell = \ell(x) = \begin{cases} \max\{\ell_i ; c_i \in Crit'(x)\} & \text{if } Crit'(x) \neq \emptyset, \\ 1 & \text{if } Crit'(x) = \emptyset. \end{cases} \tag{38}$$

LEMMA 4.1 (Transfer Expansion in the Complex Case):

$$\frac{|Df^{n_0 - n_1}(x_1)|^{\ell}}{|Df^{n_0 - n_1}(f(c_1))|} \geq K \cdot b_{n_0 - n_1} \cdot \frac{A_0}{B_1}, \tag{39}$$

where c_1 and x_1 are above and A_0, B_1 are as in Lemma 3.1 (and $|a - b|$ is taken to be the spherical distance between a and b).

Proof: The proof of Lemma 4.1 goes exactly as that of Lemma 3.1. Note that we only need to prove the analogue of Lemma 3.2, since this is the only part of the proof of Lemma 3.1 which is not 'algebraic'. So we merely need to prove the analogues of inequalities (14)–(16). Let us first consider (14). Let g be the inverse branch of $f^{n_0-n_1-1}\colon f(W_{n_1}) \to B_{n_0}$. Because of properties 1–3 just above this lemma, we can apply the lower bound in the first part of Lemma 2.5 and obtain

$$\frac{|Dg(x_0)|}{|Dg(f^{n_0-n_1}(c_1))|} \le \frac{KL}{b_{n_0-n_1}}. \tag{40}$$

This gives that the constant KL in (14) should be replaced by $b_{n_0-n_1} \cdot KL$.

Next we show that the constant KL in (15) is still universal (so it does not depend on n_0 and n_1). Indeed, let B'_{n_0} be the disc centered at x_0 with half the radius of that of B_{n_0} and g the inverse branch from above. By the second statement in Lemma 2.5,

$$|Dg(x_0)| \le K \cdot \frac{\operatorname{diam}(g(B'_{n_0}))}{r/2},$$

where K is universal, and r is the radius of B_{n_0}. Moreover, the boundary of $g(B'_{n_0})$ lies between two discs of comparable radii. Since $g(B_{n_0})$ contains $f(c_1)$ in its boundary, since $g(B'_{n_0}) \subset g(B_{n_0})$, and since $r = d(x_0, f^{n_0-n_1}(c_1))$, it follows that $\operatorname{diam}(g(B'_n)) \le C \cdot d(f(x_1), f(c_1))$ for a universal constant C, and therefore that

$$|Dg(x_0)| \le K \cdot \frac{d(f(x_1), f(c_1))}{d(x_0, f^{n_0-n_1}(c_1))}.$$

It follows that (15) also holds in the complex case for a universal Koebe constant KL.

Let us now prove the analogue (16). Let $\hat{g}$ be the inverse branch of $f^{n_0-n_1}\colon W_{n_1} \to B_{n_0}$ and B'_{n_0} be the disc from above. Note that W_{n_1} contains c_1 in its boundary. Again we have by Lemma 2.5,

$$|D\hat{g}(x_0)| \le K \cdot \frac{\operatorname{diam}(\hat{g}(B'_{n_0}))}{r/2}. \tag{41}$$

By Claim 1, $f^{n_0-n_2}(c_2)$ is outside B'_{n_0}. This implies that $f^{n_1-n_2}(c_2)$ is outside $\hat{g}(B'_{n_0})$. But this, (41) and the definition of r imply that

$$|D\hat{g}(x_0)| \le K \cdot \frac{d(x_1, f^{n_1-n_2}(c_2))}{d(x_0, f^{n_0-n_1}(c_1))}.$$

Again using that $\hat{g}(B'_{n_0})$ is almost round and contained in the set $\hat{g}(B_{n_0}) = W_{n_1}$ which contains c_1 in its boundary also,

$$|D\hat{g}(x_0)| \le K \cdot \frac{d(x_1, c_1)}{d(x_0, f^{n_0-n_1}(c_1))}.$$

Combined with $D\hat{g}(x_0) = [Df^{n_0-n_1}(x_1)]^{-1}$ this shows that (16) holds with $\frac{1}{2}$ replaced by some universal constant. From these three estimates Lemma 4.1 follows. Note that only in one of the ℓ factors of $|Df^{n_0-n_1}(x_1)|^\ell$ one uses (14), and so the spoiling constant $b_{n_0-n_1}$ appears only once. ∎

PROPOSITION 4.1: *Given f such that (3) holds, there exist $\sigma_0 > 0$, $\epsilon_0 > 0$, $K > 0$ and a function $E\colon \mathbb{N} \to \mathbb{R}$ with $E(k) \to \infty$ as $k \to \infty$ such that for each $x \in J_f$ and $n \in \mathbb{N}$, for which $x, \dots, f^{n-1}(x) \notin Crit$, the following properties hold.*

1. *Let $0 = n_s < \cdots < n_1$ be the critical times before time $n_0 = n$. Let $\ell = \ell(x)$ and c_i be as above and let A_i, B_i be as in Proposition 3.1. Then for any $0 \le i < j \le s$,*

$$|Df^{n_i-n_j}(x_j)|^\ell \ge \frac{A_i}{B_j} \prod_{k=i}^{j-1} (K_k \cdot b_{n_k-n_{k+1}} \cdot |Df^{n_k-n_{k+1}}(f(c_{k+1}))|),$$

where $K_0 = K$, and for each $k = 1, \dots, s-1$,

$$K_k \ge \begin{cases} K \cdot [\mathrm{dist}(f^{n_k}(x), Crit)]^{\ell_k - \ell} & \text{if both } \ell_k > \ell \text{ and} \\ & n_k \text{ is of type (AP)}, \\ K & \text{otherwise.} \end{cases}$$

By definition of $Crit'(x)$ the former happens only finitely (uniformly in n) often: it can only happen when $n_k \le \tau(x)$.

2. *If $n_{i+1} \ge \tau(x)$ then one of the following holds:*

$$K \cdot b_{n_i-n_{i+1}} \cdot |Df^{n_i-n_{i+1}}(f(c_{i+1}))| \ge 2, \tag{42}$$

$$\begin{aligned} &K \cdot b_{n_i-n_{i+1}} \cdot |Df^{n_i-n_{i+1}}(f(c_{i+1}))| \cdot \\ &\quad K \cdot b_{n_{i+1}-n_{i+2}} \cdot |Df^{n_{i+1}-n_{i+2}}(f(c_{i+2}))| \ge 4, \end{aligned} \tag{43}$$

or (for any $j > i$)

$$|Df^{n_i-n_j}(x_j)|^\ell \ge E(n_i - n_j) \cdot \frac{A_i}{B_j}. \tag{44}$$

3. *There exist arbitrarily large times $n = n_0$ that are of type (AP) or for which the radius $r_{n_0} \ge \sigma_0$ (so n is a σ_0-big time).*
4. *If $f^n(W_0)$ contains a σ_0-neighbourhood of $f^n(x) = x_0$, then $A_0 \ge \sigma_0^{\ell-\ell_0}$. If $f^n(x)$ is in case (AP) and moreover the critical point nearest to $f^n(x)$ has order $\ell(x) = \ell$, then $A_0 = 1$.*
5. *If $\epsilon \in (0, \epsilon_0)$ and $n \ge 0$ is the smallest time so that $f^n(x) \in N_\epsilon(Crit)$, then $f^n(x)$ is in case (AP).*

Proof of Assertions 1–2 of Proposition 4.1: The proof of Assertion 1 of this proposition is identical to the proof of Proposition 3.1. Let us now prove Assertion 2. The main difference with the real case is that now B_{n_i} does not necessarily contain W_{n_i}. Choose $\rho > 0$ so that $K \cdot b_k \cdot |Df^k(f(c))| \geq 2$ for all $k \geq \nu(\rho)$ and $c \in Crit$. (We may assume that $K \in (0,1)$.) Assume that $K \cdot b_{n_i - n_{i+1}} \cdot |Df^{n_i - n_{i+1}}(f(c))| < 2$. Then $n_i - n_{i+1} \leq \nu(\rho)$ and so $d(c_i, f^{n_i - n_{i+1}}(c_{i+1})) \geq \rho$. Hence either $d(x_i, f^{n_i - n_{i+1}}(c_{i+1})) \geq \rho/2$ or $d(x_i, c_i) \geq \rho/2$. If $d(x_i, f^{n_i - n_{i+1}}(c_{i+1})) \geq \rho/2$, then by Claim 1, $f^{n_i - n_j}(W_{n_i - n_j}(x_j))$ contains a $\rho/4$-neighbourhood of x_i for all $j > i$. Hence, if we denote by V the neighbourhood of x_j for which $f^{n_i - n_j}(V)$ is the $\rho/16$-neighbourhood of x_i, then by Koebe $Df^{n_i - n_j}$ has universally bounded distortion on V and so

$$(45) \qquad |Df^{n_i - n_j}(x_j)|^\ell \geq KL \cdot \frac{(\rho/16)^\ell}{\mathrm{diam}(V)^\ell} \geq KL' \cdot \frac{A_i}{\mathrm{diam}(V)^\ell}.$$

Since again by the Koebe Lemma 2.5, the diameter of V is at most some universal constant times $d(x_j, f^{n_j - n_{j+1}}(c_{j+1}))$, and since V does not contain c_j, and since $\ell_j > 1$, the right hand side of (45) is much larger than A_i/B_j when the diameter of V is small. Since $x \in J_f$, and $f^{n_i - n_j}|V$ is univalent, the set V has small diameter when $n_i - n_j$ is large. It follows that there exists a function $E'_\rho \colon \mathbb{N} \to \mathbb{R}$ with $\lim_{k\to\infty} E'_\rho(k) = \infty$ such that

$$(46) \quad d(x_i, f^{n_i - n_{i+1}}(c_{i+1})) \geq \rho/2 \text{implies} |Df^{n_i - n_j}(x_j)|^\ell \geq E_{\rho/2}(n_i - n_j)\frac{A_i}{B_j}.$$

The remainder of the proof goes verbatim as in the real case, except of course that we need to add the terms $b_{n_i - n_{i+1}}$. This proves Assertion 2 of Proposition 4.1. ∎

Proof of Assertions 3–5 of Proposition 4.1: The proof of Assertion 3 is only slightly different from the real case. In the first formula (29) we need to add the factor b_n. To prove the analogue of (30), instead of the One-sided Koebe Lemma, we apply the second statement of the Koebe Lemma 2.5 to the branch of $f^{-(n_{i-1} - n_i - 1)}$ that maps a $(1 + b_{n_{i-1} - n_i})$-scaled neighbourhood of $B_{n_{i-1}}$ univalently onto a neighbourhood of $f(W_{n_i})$. We get (for $K = KL \cdot M$)

$$(47) \qquad {\sigma'_0}^{l_i} \leq M\, d(f(c_i), f(x_i)) \leq K \cdot \frac{\sigma_0}{b_{n_{i-1} - n_i} \cdot |Df^{n_{i-1} - n_i - 1}(f(c))|},$$

again contradicting (29). Next, instead of inequality (31) we get

$$d(x_i, c_i)^{\ell_i} \leq \frac{K}{b_{n_{i-1} - n_i}} \frac{d(x_{i-1}, f^{n_{i-1} - n_i}(c_i))}{|Df^{n_{i-1} - n_i - 1}(f(c_i))|}.$$

Indeed, consider the inverse branch g of $f^{n_{i-1}-n_i-1}$ mapping x_i to $f(x_i)$. This branch is univalent on a $(1+b_{n_{i-1}-n_i})$-scaled neighbourhood of the ball centered at x_i with $z := f^{n_{i-1}-n_i-1}(f(c_i))$ on its boundary. Hence, by the second statement of the Koebe Lemma 2.5,

$$\begin{aligned} Dg(z) = \frac{1}{|Df^{n_{i-1}-n_i-1}(f(c_i))|} &\geq \frac{b_{n_{1-1}-n_i}}{3} \cdot \frac{d(g(x_{i-1}), g(z))}{d(x_{i-1}, z)} \\ &= \frac{b_{n_{1-1}-n_i}}{3} \cdot \frac{d(x_i, c_i)}{d(x_{i-1}, f^{n_{i-1}-n_i}(c_i))}. \end{aligned}$$

Therefore equation (32) becomes

$$d(x_i, c_i)^{\ell_i} \leq K \cdot \frac{d(x_{i-1}, f^{n_{i-1}-n_i}(c_i)) \cdot d(f^{n_{i-1}-n_i}(c_i), c_{i-1})^{\ell_{i-1}-1}}{b_{n_{i-1}-n_i} \cdot |Df^{n_{i-1}-n_i}(f(c_i))|}. \tag{48}$$

Because we have assumed (3), we obtain (33) exactly as before and the remainder of the proof of Assertion 2 still holds verbatim. The proof of Assertion 4 of Proposition 4.1 is trivial (as before). For the proof of Assertion 5, we use (48) instead of (32) and the rest goes through word for word. ∎

Proof of Theorems 1.2 and 1.3 in the complex case: The proofs of these propositions and lemma go exactly as in the real case, except that one has to use Proposition 4.1. ∎

Proof of Corollary 1.3 in the complex case: This is basically the same as in the real case, except that we cannot use the One-sided Koebe Lemma as in (35). Instead, $f(W_{n_1})$ is contained in a disk W' that is mapped univalently onto $B(x_0, r_{n_0} \cdot \exp(b_{n_0-n_1}))$. It follows again from the second statement in Lemma 2.5 that

$$\begin{aligned} |x_0 - f^{n_0-n_1}(c_1)| &\geq \frac{b_{n_0-n_1}}{KL} \cdot |Df^{n_0-n_1-1}(f(c_1))| \cdot d(f(x_1), f(c_1)) \\ &\geq \frac{b_{n_0-n_1}}{KL \cdot M} \cdot |Df^{n_0-n_1-1}(f(c_1))| \cdot d(x_1, c_1)^{\ell_1} \\ &\geq C_0 \sigma^{\ell_{\max}}, \end{aligned}$$

for some $C_0 > 0$. This gives the missing estimate. ∎

5. Lebesgue measure of Julia sets

In this section we prove that if f satisfies (3), then the Julia set J_f has either Lebesgue measure $m(J_f) = 0$ or J_f is the Riemann sphere. The real analogue is that for m-a.e. point x, $\omega(x)$ is either a cycle of intervals, or a periodic orbit; see Corollary 5.2. We will concentrate on the complex setting, since there the result is most interesting, and occasionally point out the differences in the arguments for the real setting.

The main tool to prove this is a random walk argument, but before explaining this, and stating the precise result, let us give some preliminary lemmas. The first is easy and in fact well-known; we include it for completeness.

LEMMA 5.1: *Let $J' = \{x \in J_f; \overline{\mathrm{orb}(x)} \cap Crit \neq \emptyset\}$. Then $m(J_f \setminus J') = 0$.*

Proof: Let X_k be the set of points x in J_f such that $\mathrm{dist}(\mathrm{orb}(x), Crit) \geq 1/k$. By Proposition 4.1, $x \in X_k$ has infinitely many times that are either σ_0-big or of type (AP). But as $x \in X_k$, the (AP) times are also $(1/k)$-big. Take $\sigma = \min\{\sigma_0, 1/k\}$. Therefore, each $x \in X_k$ has arbitrarily small neighbourhoods U and $n \in \mathbb{N}$ such that f^n maps U onto $B(f^n(x), \sigma/2)$ with distortion depending only on σ. A definite proportion of $B(f^n(x), \sigma/2)$ is mapped into a $1/k$-neighbourhood of $Crit$ and is therefore disjoint from X_k. It follows that X_k cannot have any density points. Since $J' = J_f \setminus (\bigcup_k X_k)$, the lemma follows. ∎

Hence, the orbit of Lebesgue almost every $x \in J_f$ accumulates on $Crit$. The next lemma shows that, if x is close to $Crit$, then for the first "closer" approach $f^m(x)$ to $Crit$, there is a neighbourhood $U \ni x$ which maps univalently to a neighbourhood, whose diameter is "larger" than the original distant $\mathrm{dist}(Crit, x)$. Here "closer" and "larger" are meant in a sense that takes the critical orders into account. Recall that in the real case, the Julia set is defined to be a set of points that do not converge to a stable or neutral periodic orbit. Recall also that $N_\epsilon(c) = B(c; \epsilon^{1/\ell(c)})$ and $N_\epsilon = \bigcup_{c \in Crit} N_\epsilon(c)$.

LEMMA 5.2: *Suppose f is a rational map satisfying (3), resp. (2). Then for all $R > 1$ there exists δ such that for all $x \notin Crit$ the following properties hold: Write $\delta(x) = \min\{\delta, d(x, c_0)^{\ell(c_0)}\}$, where c_0 is the critical point closest to x. If*

$$m := m(x) := \min\{i > 0; f^i(x) \in N_{\delta(x)}\} < \infty, \tag{49}$$

then there exist a neighbourhood $U \ni x$ such that:

1. *$f^m\colon U \to f^m(U)$ is univalent resp. diffeomorphic;*
2. *$\mathrm{diam}(U) < \frac{1}{2} d(c_0, x)$;*
3. *$f^m(U) = B(c', R\delta(x)^{\ell(c_0)/\ell(c_1)})$, where c_1 is the critical point closest to $f^m(x)$.*

We give the proof for the complex version; the real version is analogous.

Proof of Lemma 5.2: Fix $R > 1$. Let K be the constant from (48). By (3), there exists i_0 such that $b_i|Df^i(f(c))|/K > R^{2\ell_{\max}}$ for all $c \in Crit$ and $i \geq i_0$. Take δ so small that $d(c', f^i(c)) > R^2\delta^{1/\ell(c')}$ for all $0 < i < i_0$ and $c, c' \in Crit$.

Next, let x be arbitrary, and assume that m in (49) is finite. By Assertion 4 of Proposition 4.1, m is critical of type (AP). Let m' be the last critical time of x before m. By construction

$$|f^m(x) - c_1|^{\ell(c_1)} < |x - c_0|^{\ell(c_0)} \leq |f^{m'}(x) - c'|^{\ell(c')},$$

where c_1, c' and c_0 are the critical points closest to $f^m(x)$, $f^{m'}(x)$ resp. x. By (48) and the fact that m is of type (AP) we have

$$\begin{aligned}|f^{m'}(x) - c'|^{\ell(c')}.&\frac{b_{m-m'} \cdot |Df^{m-m'}(f(c'))|}{KL} \\ &\leq |f^m(x) - c_1| \cdot |c_1 - f^{m-m'}(c')|^{\ell(c_1)-1} \\ &\leq |c_1 - f^{m-m'}(c')|^{\ell(c_1)}.\end{aligned}$$

Either $m - m' \geq i_0$, and then we find

$$|c_1 - f^{m-m'}(c')|^{\ell(c_1)} \geq R^{\ell_{\max}}|c_0 - x|^{\ell(c_0)},$$

or $m - m' < i_0$ and

$$|c_1 - f^{m-m'}(c')|^{\ell(c_1)} \geq R^{2\ell(c_1)}\delta.$$

In either case $|c_1 - f^{m-m'}(c')| > R^2\delta(x)^{1/\ell(c_1)}$, and we can find neighbourhoods $T \supset U \ni x$ such that its univalent images $f^m(T) \supset f^m(U)$ are round disks with radii $R^2\delta(x)^{1/\ell(c_1)} > R\delta(x)^{1/\ell(c_1)}$. The distortion of $f^m : U \to f^m(U)$ depends only on R, and for R large, U is an almost round disk in T. Since $c \notin T$, $\mathrm{diam}(U)$ is small compared to $d(x, c)$. The three statements follow directly from this. ∎

Suppose that $m(J_f) > 0$ and $X \subset J_f$ is a totally invariant set which supports an ergodic component m_0 of Lebesgue measure m. To be precise, $f^{-1}(X) = X$, $m_0(X) = 1$ and $m_j(X)$ for all other ergodic components. Assume further that $x \in J'$ is a density point of X. Then by applying Lemma 5.2 repeatedly, one can show that at least one critical point in J_f is also a density point of X. (Furthermore, if any critical point c belongs to $\overline{\mathrm{orb}(y)}$ for a set of points $y \in X$ of positive measure, then c is also a density point of X.) This is in a nutshell the argument for the following proposition.

PROPOSITION 5.1: *Suppose f satisfies (2) or (3); then Lebesgue measure m has at most $\#Crit$ ergodic components.*

In the real case, finiteness of ergodic components was shown by Blokh and Lyubich without any summability condition, see [1, 2, 11, 14]. The number of ergodic components is sharp, also if $\#Crit > 1$. More precisely, for every modality there is a one-dimensional map with exactly $\#Crit$ Lebesgue ergodic components.

In the complex case, the first proof was given by Przytycki, [21, Theorem B], and his proof applies to α-conformal measures. Recall that a probability measure m_α is α-conformal if $m_\alpha(f(U)) = \int_U |Df(x)|^\alpha dm_\alpha$ whenever U is measurable and $f\colon U \to f(U)$ is one-to-one. By Sullivan's result [26], J_f always supports a α-conformal probability measure for some $0 < \alpha \leq 2$. Prado [20] proved (not using summability) that a non-atomic α-conformal m_α is ergodic for certain unicritical polynomials on the complex plane. Graczyk and Smirnov [9, Theorem 2] show ergodicity provided the rational map satisfies a slightly stronger summability condition. Whereas the non-ergodicity result of [25] depends on the existence of Julia sets $J \neq \hat{\mathbb{C}}$ with positive Lebesgue measure, no non-ergodic α-conformal non-atomic measures are known.

Let us now start turn to the major result of this section.

THEOREM 5.1: *If f satisfies the expansion assumption (3) resp. (2), then there exists $\epsilon > 0$ and $K > 1$ such that for m-a.e. $x \in J_f$ there exists $t_i \to \infty$ and neighbourhoods $U_i(x)$ of x such that*

$$f^{t_i}\colon U_i(x) \to B(c, \epsilon^{1/\ell(c)}), \tag{50}$$

is univalent, resp. diffeomorphic, and has distortion bounded by K. Here $c \in Crit$ is the critical point closest to $f^{t_i}(x)$.

Points x that have arbitrarily small neighbourhoods that map univalently to large scale are sometimes called **conical** points. Hence Theorem 5.1 says that m-a.e. point is conical.

Whereas Lemma 5.1 and Proposition 5.1 can be extended to non-atomic α-conformal measures, we have met with serious difficulties extending Theorem 5.1 in this way. One of the reasons is the estimate of the m_α-measure of disks D of radius ρ. Whereas $m_\alpha(D) \leq K_2\rho^\alpha$ for some uniform constant K_1, no uniform constant $K_1 > 0$ such that $m_\alpha(D) \geq K_1\rho^\alpha$ may be expected. This uniform lower bound obviously exists for Lebesgue measure.

In [9, Theorem 2 and 3] a certain summability condition is used to show that J_f supports an α-conformal measure m_α, and m_α-a.e. $x \in J_f$ is conical (x even reaches large scale with, in a sense, positive frequency).

Proof of Theorem 5.1: Assume that $m(J_f) > 0$. Lemma 5.2 shows that if x is close to $Crit$, and $f^m(x)$ is even closer, then there exists a neighbourhood $U \ni x$ such that the majority of the points of U move further away from $Crit$ under m iterates. We will turn this phenomenon into a random walk argument.

Let us first fix some constants and notation. Lemma 2.5 implies that if $V \subset U$ are topological disks, and g maps them univalently onto concentric round disks of radii $1 < R$, then $g|V$ has distortion

$$K(R) \le \left(\frac{R+1}{R-1}\right)^4.$$

Take $R > 4^{10}$ so large that $K(\sqrt{R}) < 1.1$ and let $\rho = R^{-1/10} \in (0, 1/4)$. Let δ be as in Lemma 5.2 and define

$$N_i(c) := B(c; \delta^{1/\ell(c)}\rho^i) \quad \text{and} \quad N_i = \bigcup_{c \in Crit} N_i(c),$$

and also $A_i = N_i \setminus N_{i+1}$. By convention set also $N_{-1} = \hat{\mathbb{C}}$.

As before, let X be a totally invariant set supporting an ergodic component. By Proposition 5.1, there are only finitely many disjoint sets X of this type, X contains at least one critical point c in its closure, and if $c \in \overline{\mathrm{orb}(y)}$ for typical points in X, then c is a density point of X. We can assume that δ is so small that for each such critical point c and each $i \ge 0$, $m(A_i \cap X) \ge 0.999 m(A_i)$.

The collections annuli A_i are considered as *states* of a random walk. We write

$$\chi(x) = k \quad \text{if } x \in A_k.$$

If $x \in J' \setminus Crit$, then

$$m(x) := \min\{i > 0; \chi(f^i(x)) > \chi(x)\} < \infty.$$

Take $x \in J'$ close to $Crit$, say $x \in A_k(c)$ with $k \ge 30$ and let c_m be the critical point closest to $f^{m(x)}(x)$. By Lemma 5.2, there exist neighbourhoods $T(x) \supset U(x) \supset V(x) \ni x$ such that $f^{m(x)}$ maps $T(x)$ univalently onto $N_{k-30}(c_m)$, $U(x)$ univalently onto $N_{k-20}(c_m)$, and $V(x)$ onto $N_{k-10}(c_m)$. Since $f^{m(x)}|V(x)$ has distortion bounded by 1.1, we see that indeed most points in $U(x)$ move to a lower state than their original state. More precisely and strongly:

LEMMA 5.3: *Suppose f satisfies (2) or (3). Let X support an ergodic component of m as above. If f^t maps neighbourhoods $V \subset U$ univalently onto $V(x) \subset U(x)$ for some $x \in X$, then*

$$E_m(\chi \circ f^{m(x)+t}|X \cap V) := \frac{1}{m(X \cap V)} \int_{X \cap V} \chi \circ f^{m(x)+t}(y) dm < \max\{k-3, 30\},$$

and

$$E_m(\chi \circ f^{m(x)+t}|X \cap U) := \frac{1}{m(X \cap U)} \int_{X \cap U} \chi \circ f^{m(x)+t}(y) dm < \max\{k-3, 30\}.$$

Proof of Lemma 5.3: This is a straightforward computation. The distortion of $f^{t+m(x)}|V$ depends only on R and hence is less than 1.1. First assume that $\chi(x) = k$ for $k \geq 30$, so X almost completely fills the annuli A_i for $i \geq k - 10$. Then

$$\frac{2}{3}(1-\rho^d)\rho^{jd} \leq \frac{m(\{y \in X \cap V; f^{m(x)+t}(y) \in A_{k-10+j}\})}{m(X \cap V)} \leq \frac{3}{2}(1-\rho^d)\rho^{jd},$$

where $d = 2 = \dim(\hat{\mathbb{C}})$ and the factors $\frac{2}{3}$ and $\frac{3}{2}$ bound the distortion effects. Thus

$$\begin{aligned} E_m(\chi \circ f^{m(x)+t}|X \cap V) &\leq k + \frac{3}{2}(1-\rho^d)\sum_{j \geq 0} j\rho^{(j+10)d} + \frac{2}{3}(1-\rho^d)\sum_{j=0}^{10}(j-10)\rho^{jd} \\ &< k + \rho^{10d}(1-\rho^d)^{-1} - \frac{10}{3}(1-\rho^d) < k-3. \end{aligned}$$

If $k \leq 30$, a similar proof shows that still $E_m(\chi \circ f^{m(x)+t}|U) \leq 30$.

Note that the expectations $E_m(\chi \circ f^{m(x)+t}|U)$ are naturally lower than $E_m(\chi \circ f^{m(x)+t}|V)$, because the annulus $U \setminus V$ is mapped to lower states $k-11$, $k-12, \ldots, k-20$. (Note that no distortion bound is necessary on $U \setminus V$.)

The proof for the real version of this lemma is the same, except that then $d = \dim(\mathbb{R}) = 1$. ∎

If $x \in U(y) \cap X$ satisfies $\chi(x) = k > 30$, then $f^{m(y)}(x)$ is likely to belong to a lower state than x. At the same time, $f^{m(y)}(x)$ may belong to some $U(y')$, and then most likely to the part of $U(y')$ that has an even lower state after $m(y')$ iterate. Basically, we would like to repeat the argument, and consider the numbers $\chi(x), \chi(f^{m(y)}(x)), \chi(f^{m(y)+m(y')}(x)), \ldots$ as random variables, which have bounded expectation.

However, a single point x can belong to different disks $U(y)$; after all, the disks $U(y)$ form a cover of J_f mod 0, not a partition. To make this cover into a partition is an awkward task, because the partition is not likely to have a Markov property. Therefore, refining the partition according to the successive iterates of the maps $f^{m(y)}$ will be extremely complicated. In addition, an assignment that seems favourable at one iterate may prove unfavourable at the next.

We therefore follow another strategy: instead of assigning each x to a unique disk $U(y)$, we allow x to belong to $U(y)$ for several points y, and therefore to have several images under $f^{m(y)}$. Let us call this map Φ; it is multivalued, but all images $\Phi(x)$ belong to $\mathrm{orb}(x)$. For the n-th iterate of Φ, we allow each of the image points in $\Phi^{n-1}(x)$ to have several new images. When computing the expectation of $\chi \circ \Phi^n$, we can minimize over all $z \in \Phi^n(x)$.

To make this precise, define

$$\tau(x) = \tau_1(x) = \left\{ m(y);\ \begin{array}{ll} 1. & x \in V(y) \text{ or} \\ 2. & x \in U(y) \text{ and } U(y) \supset U(y') \\ & \text{for all } y' \text{ such that } x \in V(y') \end{array} \right\}.$$

Note that the definition of $\tau(x)$ is such that if x and x' belong to the same $V(z)$ for some z, then the points y for which the second rule holds are the same for x and x'. Furthermore, the disks $V(y)$ and $U(y)$ that appear in this definition have the following nesting property:

LEMMA 5.4: *If $V(y) \cap V(y') \neq \emptyset$, then $U(y) \subset U(y')$ or $U(y') \subset U(y)$.*

Proof: By the choice of R, the distortion of $f^{m(y)}|U(y)$ is small, and at the same time $V(y)$ is small compared to $U(y)$ and the modulus of $U(y) \setminus V(y)$ is bounded away from 0. Let $V(y)$ and $V(y')$ intersect each other. If $m(y) = m(y')$, then $V(y) \subset V(y')$ and $U(y) \subset U(y')$, or vice versa. In this case there is nothing to prove. Hence assume that $m(y) < m(y')$. If $U(y') \supset V(y)$, then $f^{m(y)}(U(y')) \ni c$ for some critical point, so $f^{m(y')}|U(y')$ cannot be univalent. Therefore $U(y') \not\supset V(y)$. This means that $\mathrm{diam}(U(y')) \approx \mathrm{diam}(V(y)) \ll \mathrm{diam}(U(y))$, and $U(y') \subset U(y)$. Note that this proves at the same time that the transfer time of the larger set is smaller. ∎

Let Φ be the multivalued function:

$$\Phi(x) = \{f^t(x); t \in \tau(x)\}.$$

Now for the second iterate, define

$$\tau_2(x) = \left\{ m(y) + t; t \in \tau_1(x) \text{ and } \begin{array}{ll} 1. & f^t(x) \in V(y) \text{ or} \\ 2. & f^t(x) \in U(y) \text{ and } U(y) \supset U(y') \\ & \text{for all } y' \text{ such that } f^t(x) \in V(y') \end{array} \right\}$$

and $\Phi^2(x) = \{f^t(x); t \in \tau_2(x)\}$. Similarly, we define τ_n and Φ^n for every $x \in J'$ and $n \in \mathbb{N}$.

If $x \in U(y)$, $\chi(x) = k$ for some $t = m(y) \in \tau(x)$, then by the proof of Lemma 5.2, there is a neighbourhood $T(y)$ such that $f^t(U(y)) = N_i(c) \subset N_{i-10}(c) = f^t(T(y))$ for some $i \leq k$ and c the critical point closest to $f^t(y)$. Also, $\mathrm{diam}(U(y)) \leq \frac{1}{2}\,\mathrm{dist}(y, Crit)$. Since $\rho < 1/4$, this implies that $U(y) \subset A_{k+1} \cup A_k \cup A_{k-1}$. The size of $T(y)$ is not given, but the proof of Lemma 5.2 extends to show that there is a neighbourhood $\tilde{T}(y) \subset T(y)$ such that $\tilde{T}(y) \subset A_{k+1} \cup A_k \cup A_{k-1}$ and $f^t(\tilde{T}(y)) = N_{i-2}(c)$.

If $f^t(x) \in U(y')$, then we can repeat the above construction and find that $U(y') \subset \tilde{T}(y') \subset f^t(\tilde{T}(y))$. This implies that $\tilde{T}(y)$ contains a neighbourhood that maps univalently under $f^{t+m(y')}$ onto $f^{m(y')}(\tilde{T}(y'))$. Therefore we can use the expectation estimates of Lemma 5.3 for the branches of Φ_2 as well, and by induction, also for every branch of Φ^n for every $n \geq 1$.

If, for each $t \in \tau_{n-1}(x)$, $f^t(x) \notin U(y)$ for any y, then we simply put $\tau_n(x) = \tau_{n-1} + 1$. For example, this can happen if x belongs to $X \setminus J'$, so $\mathrm{orb}(x)$ does not accumulate on $Crit$.

To complete the ingredients of the random walk, define $\chi_n(x) = \min\{\chi \circ f^t(x); t \in \tau_n(x)\}$. We claim:

$$\text{For all } n \in \mathbb{N}, \quad E_m(\chi_n) \leq 30. \tag{51}$$

Proof of the Claim: Clearly (51) is true for $n = 0$, because N_{30} is only a small part of $\hat{\mathbb{C}}$ and $m(N_i) \to 0$ exponentially fast.

Assume (51) holds for $n-1$. For each $x \in \hat{\mathbb{C}}$, let $s(x) \in \tau_{n-1}(x)$ be such that $\chi \circ f^{s(x)}(x)$ is minimal among all $s \in \tau_{n-1}(x)$. Let $\tilde{s}(x) = \min\{t; t \in \tau(f^{s(x)}(x))\}$, that is (by Lemma 5.4), $\tilde{s}(x)$ is the transfer time of the largest disk $V(y)$ or $U(y)$ that contains $f^{s(x)}(x)$. In particular, the definition of τ is such that if $f^{s(x)} \in V(y)$ for some y, then (irrespective whether $\tilde{s}(x) = m(y)$ or not) $\tilde{s}(x) = \tilde{s}(x')$ for all points x' such that $f^{s(x')} \in V(y)$.

Obviously $E_m(\chi_n) \leq \int \chi \circ f^{s+\tilde{s}} dm$.

We divide $\hat{\mathbb{C}}$ into $Y_0 \cup Y_1 \cup Y_2$, where the Y_i are defined as follows:

- $Y_0 = \{x; f^{s(x)+\tilde{s}(x)}(x) \notin N_{30}\}$. For instance, Y_0 contains points in $X \setminus J'$. Clearly $\tau_n(x) < 30$ for all $x \in Y_0$.
- $Y_1 = \{x \notin Y_0; \tilde{s}(x) = m(y)$ and $f^s(x) \in U(y) \setminus V(y)$for some $y\}$. There is a countable set of points y such that $Y_1 = \bigcup_y Y_{1,y}$ where $Y_{1,y} = \{x \notin Y_0; \tilde{s}(x) = m(y)$ and $f^s(x) \in U(y) \setminus V(y)$ for some $y\}$. If $f^{s(x)} \in U(y) \setminus V(y)$, then $\chi(f^t(x))$ and $\chi(y)$ differ by at most one, and at the next

step, x is mapped outside $N_{\chi(y)-10}$. Therefore, the points in $Y_{1,y}$ satisfy $\chi_n(x) \leq \max\{\chi(f^s(x)) - 9, 30\}$.

- $Y_2 = \{x \notin Y_0; \tilde{s}(x) = m(y)$ and $f^s(x) \in V(y)\}$. There is a countable set of points y such that $Y_2 = \bigcup_y Y_{2,y}$, where

$$Y_{2,y} = \{x \notin Y_0; \tilde{s}(x) = m(y) \text{ and } f^s(x) \in V(y)\}.$$

Due to Lemma 5.4, all the sets $V(y)$ are disjoint. Moreover, the remark below the definition of τ implies that $\tilde{s}(x)$ is constant on each $V_{1,y}$. We can use Lemma 5.3 to compute the expectations on the sets $V(y)$.

Note that the sets Y_0, Y_1 and Y_2 are pairwise disjoint. Recall that E_m is the expectation with respect to normalized Lebesgue measure on X. We obtain

$$\begin{aligned} E_m(\chi_n) \leq & m(Y_0)E_m(\chi_n|Y_0) + m(Y_1)E_m(\chi_n|Y_1) + m(Y_2)E_m(\chi_n|Y_2) \\ \leq & 30m(Y_0) + \sum_y m(Y_{1,y}) \max\{(E_m(\chi_{n-1}|Y_{1,y}) - 9), 30\} \\ & + \sum_y m(Y_{2,y}) \max\{(E_m(\chi_{n-1}|Y_{2,y}) - 3), 30\} \\ \leq & \max\{E_m(\chi_{n-1}) - 3, 30\} \leq 30. \end{aligned}$$

This shows that $E_m(\chi_n) \leq 30$ for all n, proving the claim.

Since the property "$\chi_n(x) \leq 30$ infinitely often" is invariant under f, we get that for m-a.e. $x \in X$, $\liminf_n \inf\{\chi(f^t(x)); t \in \tau_n(x)\} \leq 30$. This proves Theorem 5.1 with $\epsilon = \delta\rho^{30\ell_{\max}}$. ∎

Theorem 1.4 is a direct consequence of Theorem 5.1.

Proof of Theorem 1.4: Suppose x is a density point of J_f that satisfies (50). Then there exists $c \in J_f$ such that f^{t_i} maps $U_i(x)$ with bounded distortion onto $B(c, \epsilon^{1/\ell(c)})$ for arbitrarily small neighbourhoods $U_i(x)$ and corresponding iterates t_i. It follows that $m(B(c, \epsilon^{1/\ell(c)}) \cap J_f) = m(B(c, \epsilon^{1/\ell(c)}))$. Now $J = \hat{\mathbb{C}}$ is immediate. ∎

COROLLARY 5.1: *If f satisfies (3) and $J_f = \hat{\mathbb{C}}$, then m is ergodic, conservative and exact.*

Proof: Suppose X is any forward invariant set of positive measure. As in the previous corollary, there exists c such that $m(B(c, \epsilon^{1/\ell(c)}) \cap X) = m(B(c, \epsilon^{1/\ell(c)}))$, and therefore X has full measure in $\hat{\mathbb{C}}$. Hence m is ergodic and conservative. Moreover, $X \cap f(X)$ has positive measure. It follows that m is exact as well; see, e.g., [3, Proposition 2.1]. ∎

Theorem 5.1 has its real analog, and the proof is basically the same. This allows the following result.

COROLLARY 5.2: *If f satisfies (2) then for m-a.e. $x \in I$, either $\omega(x)$ is a cycle of intervals or $\omega(x)$ is a periodic orbit.*

Hence, there are no solenoidal or wild attractors.

Proof: Suppose that $m(J_f) > 0$, otherwise there is nothing to prove. Since J_f is closed and invariant, J_f contains an (at least one-sided) interval $B(c, \epsilon^{1/\ell(c)})$ for some $c \in J_f$. Consequently, J_f is the union of cycles of intervals. Suppose there is an open interval U in any of these cycles (call this cycle J_0), such that $m(\{x \in J_0; \omega(x) \cap U = \emptyset\}) > 0$. Then some interval of the form $B(c, \epsilon^{1/\ell(c)})$ contains a preimage V of U. This means m-a.e. $x \in J_0$ has arbitrarily small neighbourhoods $U_i(x)$ of which a definite proportion belongs to $\bigcup_n f^{-n}(U)$. Therefore $\{x \in J_0; \omega(x) \cap U = \emptyset\}$ has no density points, a contradiction. ∎

COROLLARY 5.3: *If f satisfies (3), then either $\omega(x)$ is an attracting periodic orbit m-a.e., or $\omega(x) = \hat{\mathbb{C}}$ m-a.e.*

Proof: Similar to the proof of Corollary 5.2 ∎

6. Some counterexamples

In this section we will show that a multimodal Collet–Eckmann map need not satisfy (BCE) or (BBC) if there are critical points with different orders.

THEOREM 6.1: *For each of the following three statements, there is a bimodal Collet–Eckmann polynomial f satisfying it:*

1. *f fails both (BCE) and (BBC).*
2. *f fails (BCE), but satisfies (BBC).*
3. *f satisfies (BCE), but fails (BBC).*

Proof: To prove the first statement, we construct a bimodal map satisfying (CE) on both critical points, but for which one of them, c_1, has a sequence of preimages $y_i \in f^{-n_i}(c_1)$ such that

$$|Df^{n_i}(y_i)| \leq Ce^{-\alpha n_i}$$

for some $C, \alpha > 0$ and all $i \geq 1$. This shows that (BCE) and (BBC) fail.

Let $f\colon [0,1] \to [0,1]$ be a bimodal map, say a fifth order polynomial, with two critical points $0 < c_1 < c_2 < 1$ such that the corresponding critical orders $\ell_1 = 2$

and $\ell_2 = 4$. Furthermore, suppose that $f(0) = 0$, $f(1) = f(c_1) = 1$. The map assumes a local minimum $f(c_2) > 0$ at c_2; its precise value will be determined in the construction below.

Assume that $f^i(c_2)$ stays out of a neighbourhood of the critical set for $1 \leq i < n$, such that $Df^{n-1}(f(c_2)) \sim \lambda^{n-1}$ for some $\lambda > 1$. For example, this holds if c_2 spends most of the iterates $i < n$ very close to the fixed point $p \in (c_1, c_2)$, and this fixed point has multiplier $|Df(p)| = \lambda$. Next assume that $\epsilon := |f^n(c_2) - c_1| \approx \lambda^{-\beta n}$ for some $\beta \in (\frac{1}{3}, 1)$. Then $Df^n(f(c_2)) \approx \lambda^{(1-\beta)n}$, so the Collet–Eckmann condition is not violated here. In addition, assume that f^n maps a neighbourhood of c_2 with one fold onto a neighbourhood of c_1. Then there is a point y, say $y < c_2$, such that $f^n(y) = c_1$ and

$$|y - c_2| \approx |f(y) - f(c_2)|^{1/4} \approx (\epsilon \cdot \lambda^{-n})^{1/4} \approx \lambda^{-(1+\beta)n/4}.$$

(Notice that $|y - c_2| \gg \epsilon$ if $\beta > \frac{1}{3}$.) This gives

$$|Df^n(y)| \approx \lambda^n \cdot |y - c_2|^3 \approx \lambda^{(1-3\beta)n/4},$$

which is exponentially small if $\beta > \frac{1}{3}$. The idea is now to construct a map f exhibiting a cascade of the above events, for which there exists a sequence n_i such that

$$|Df^{n_i-1}(f(c_2))| \approx \lambda^{n_i} \quad \text{and} \quad |f^{n_i}(c_2) - c_1| \approx \lambda^{-\beta n_i}$$

for all i, and such that f^{n_i} maps a neighbourhood of c_2 with one fold onto a neighbourhood of c_1. Then we get a sequence of points $y_i \to c_2$ such that

$$f^{n_i}(y_i) = c_1 \quad \text{and} \quad |Df^{n_i}(y_i)| \approx \lambda^{-\alpha n_i}$$

for $\alpha = (3\beta - 1)/4 > 0$. There are no (combinatorial) restrictions to such a cascade construction.

Let us now change the construction a little to prove the other two statements. For the second statement, choose $\epsilon := |f^n(c_2) - c_1| \approx \lambda^{-\beta n}$ for $\beta = \frac{1}{3}$. The previous computations then give $|Df^{n_i}(y_i)| \approx 1$.

For the third statement we argue as follows: Instead of having f^n map a neighbourhood of c_2 onto a neighbourhood of c_1, let $f^n(c_2)$ be very close to c_1 such that c_2 assumes a local minimum of $|f^n(x) - c_1|$. Therefore, no neighbourhood of c_2 on which f^n has only one fold contains c_1 in its f^n-image. This adjustment is compatible with *long-branchedness* of f, i.e., there exists $\gamma > 0$ such that for every i and every maximal monotonicity interval J of f^i satisfies $|f^i(J)| > \gamma$.

By use of a cascade of the above construction, we arrive at a map for which there are sequences $\{n_i\} \subset \mathbb{N}$ and $\{y_i\} \subset [0,1]$ such that $y_i \to c_2$, $f^{n_i}(y_i) \to c_1$ and $|Df^{n_i}(y_i)| \leq Ce^{-n_i}$. Moreover, $|f^{n_i}(y_i) - c_1| < |f^n(y_i) - c_1|$ for all $n < n_i$. It follows that (BBC) fails. To ensure that (BCE) holds, we prove the following lemma.

LEMMA 6.1: *Let f be a long-branched map with non-flat critical points and $Sf \leq 0$. Then (CE) implies (BCE).*

Proof: Take any critical point c and let U be a neighbourhood of c such that $|U| < \gamma/2$ and also $|f(U')| < \gamma/2$ for each component U' of $U \setminus \{c\}$. Here the γ comes from the definition of long-branchedness. Let $X = \bigcup_{n \geq 0} f^n(U)$. Obviously X is forward invariant.

By Proposition 3.1 applied to X, every critical point $\tilde{c} \in X$ of maximal critical order satisfies (BCE). Assume by contradiction that (BCE) fails in c; say there are sequences $\{n_i\} \subset \mathbb{N}$ and $\{y_i\} \subset [0,1]$ such that $f^{n_i}(y_i) = c$ and $\limsup_i \frac{1}{n_i} \log |Df^{n_i}(y_i)| \leq 0$.

Recall that T_{n_i} is the maximal neighbourhood of y_i on which f^{n_i} is diffeomorphic. By long-branchedness, n_i is a γ-big time of type (NAP). By the Koebe Principle, there exist N such that $\tilde{c} \in f^N(U)$. Take $x_i \in T_{n_i}$ such that $f^{n_i}(x_i) \in U$ and $f^{N+n_i}(x_i) = \tilde{c}$. Then, due to the Koebe Lemma 2.1,

$$|Df^{N+n_i}(x_i)| \leq KL \cdot \sup\{|Df^N(z)|; z \in U\} \cdot |Df^{n_i}(y_i)|.$$

This contradicts that (BCE) holds in $\tilde{c}$. ∎

This proves Theorem 6.1. ∎

References

[1] A. Blokh and M. Lyubich, *Attractors and transformations of an interval*, Banach Center Publications **23** (1986), 427–442.

[2] A. Blokh and M. Lyubich, *On the decomposition of one-dimensional attractors of unimodal maps of the interval*, Algebra and Analysis (Leningrad Mathematical Journal) **1** (1989), 128–145.

[3] H. Bruin and J. Hawkins, *Exactness and maximal automorphic factors of unimodal maps*, Ergodic Theory and Dynamical Systems **21** (2001), 1009–1034.

[4] H. Bruin, S. Luzzatto and S. van Strien, *Decay of correlations in one-dimensional dynamics*, Annales Scientifiques de l'École Normale Supérieure, to appear.

[5] H. Bruin and S. van Strien, *Existence of acips for multimodal maps*, in *Global Analysis of Dynamical Systems*, Festschrift to Floris Takens for his 60'th birthday, 2001, to appear.

[6] L. Carleson and W. Gamelin, *Complex Dynamics*, Springer, Berlin, 1995.

[7] P. Collet and J.-P. Eckmann, *Positive Lyapunov exponents and absolute continuity for maps of the interval*, Ergodic Theory and Dynamical Systems **3** (1983), 13–46.

[8] J. Graczyk and S. Smirnov, *Collet, Eckmann & Hölder*, Inventiones Mathematicae **133** (1998), 69–96.

[9] J. Graczyk and S. Smirnov, *Non-uniform hyperbolicity in complex dynamics I, II*, Preprint (2001).

[10] J. Graczyk and S. Smirnov, *Non-uniform hyperbolicity in complex dynamics. I Poincaré series and induced hyperbolicity*, Manuscript (2000).

[11] M. Lyubich, *Ergodic theory for smooth one-dimensional dynamical systems*, Preprint, Stony Brook **11** (1990).

[12] R. Mañé, *Hyperbolicity, sinks and measure in one dimensional dynamics*, Communications in Mathematical Physics **100** (1985), 495–524.

[13] R. Mañé, *On a theorem of Fatou*, Boletim da Sociedade Brasileira de Matemática (N.S.) **24** (1993), 1–11.

[14] W. de Melo and S. van Strien, *One-dimensional Dynamics*, Springer, Berlin, 1993.

[15] M. Misiurewicz, *Absolutely continuous measures for certain maps of an interval*, Publications Mathématiques de l'Institut des Hautes Études Scientifiques **53** (1981), 17–51.

[16] T. Nowicki, *Symmetric S-unimodal mappings and positive Liapunov exponents*, Ergodic Theory and Dynamical Systems **5** (1985), 611–616.

[17] T. Nowicki and D. Sands, *Non-uniform hyperbolicity and universal bounds for S-unimodal maps*, Inventiones Mathematicae **132** (1998), 633–680.

[18] T. Nowicki and S. van Strien, *Invariant measures under a summability condition for unimodal maps*, Inventiones Mathematicae **105** (1991), 123–136.

[19] Chr. Pommerenke, *Boundary behaviour of conformal maps*, Grundlehren der mathematischen Wissenschaften **299**, Springer-Verlag, Berlin, 1992.

[20] E. Prado, *Ergodicity of conformal measures for unimodal polynomials*, Conformal Geometry and Dynamics **2** (1998), 29–44.

[21] F. Przytycki, *Iteration of holomorphic Collet–Eckmann maps: conformal and invariant measures*, Transactions of the American Mathematical Society **350** (1998), 717–742.

[22] F. Przytycki, J. Rivera-Letelier and S. Smirnov, *Equivalence and topological invariance of conditions for non-uniform hyperbolicity in the iteration of rational maps*, Inventiones Mathematicae **151** (2003), 29–63.

[23] F. Przytycki and S. Rohde, *Rigidity of holomorphic Collet–Eckmann repellers*, Arkiv för Matematik **37** (1999), 357–371.

[24] J. Rivera-Letelier, *Rational maps with decay of geometry: Rigidity, Thurston's algorithm and local connectivity*, Preprint, Stony Brook **9** (2000).

[25] S. van Strien, *Transitive maps which are not ergodic with respect to Lebesgue measure*, Ergodic Theory and Dynamical Systems **16** (1996), 833–848.

[26] D. Sullivan, *Conformal dynamical systems*, Lecture Notes in Mathematics **1007**, Springer, Berlin, 1983, pp. 725–752.