Ergod. Th. & Dynam. Sys. (2007), **27**, 743–768
doi:10.1017/S0143385706000976
© *2007 Cambridge University Press*
Printed in the United Kingdom

Markov extensions and lifting measures for complex polynomials

HENK BRUIN and MIKE TODD
Department of Mathematics, University of Surrey, Guildford, Surrey GU2 7XH, UK
(e-mail: h.bruin@surrey.ac.uk, m.todd@surrey.ac.uk)

(*Received* 26 *July* 2005 *and accepted in revised form* 26 *September* 2006)

Abstract. For polynomials f on the complex plane with a dendrite Julia set we study invariant probability measures, obtained from a reference measure. To do this we follow Keller [**K1**] in constructing canonical Markov extensions. We discuss 'liftability' of measures (both f-invariant and non-invariant) to the Markov extension, showing that invariant measures are liftable if and only if they have a positive Lyapunov exponent. We also show that δ-conformal measure is liftable if and only if the set of points with positive Lyapunov exponent has positive measure.

1. *Introduction*

Ergodic properties for polynomial or rational maps have been looked at for various measures and various types of Julia sets. One can consider the measures of maximal entropy, e.g. [**FLM, Zd**], or more generally, equilibrium states of certain Hölder potentials, see e.g. [**Ly, DPU, Ha**]. This approach is particularly natural when the map is hyperbolic and the potential is $-t \log |Df|$, where t is the Hausdorff dimension of the Julia set: then the equilibrium state is equivalent to conformal measure (as obtained by Sullivan, see [**Su**]). When the Julia set is parabolic, invariant measures equivalent to conformal measure are found in [**DU, U1, U2**]. In the case where there are recurrent critical points in the Julia set, the papers [**GS, Pr, Re**] focus on invariant probability measures that are absolutely continuous with respect to conformal measures, using assumptions on the derivatives on the critical orbits. See [**PU, U3**] for surveys. The theory has not yet developed to the same extent as, for example, interval maps, where the availability of induced maps and tower constructions (cf. [**Y1, Y2**]) allowed the investigation of several stochastic properties, including the rate of mixing and central limit theorem [**Y2, BLS**], return time statistics and related properties [**BSTV, BV, Co**] and invariance principles, see e.g. [**MN**]. However, during the preparation of this paper, we learned that some good results in this direction have been proved for rational maps satisfying the 'topological Collet–Eckmann' condition in [**PrRL**].

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

In the 1980s, Hofbauer and Keller constructed so-called canonical Markov extensions for piecewise monotone maps of the interval [**Ho, HK, K1**], which they used to study the topological and measure theoretical behaviour of these maps. These Markov extensions were considered in an abstract setting in [**K1, Bu**], where one of the aims was to extend the theory to higher dimensions. Indeed in [**Bu**] some higher dimensional examples are given. That paper focuses on the probability measures given by the symbolic dynamics obtained from the tower structure, an approach also used in [**Ne, BuS**]. In [**BuK**] results on transfer operators are proved in this higher dimensional setting and in [**BuPS**] conformal measures are found.

Our approach follows the papers of Keller [**K1, K2**]. In the first of these papers, results are proved about the liftability of probability measures on the original system to the associated Markov extension. In particular, the liftability of ergodic invariant measures with positive entropy is shown. While the abstract theory given there applies, in principle, in any dimension, the applications given are to interval maps. In the second paper it is shown that, given a smooth interval map, positive pointwise Lyapunov exponents implies the liftability of Lebesgue measure. The purpose of this paper is to extend those results to maps on the complex plane. We construct Markov extensions $(\hat{\mathcal{J}}, \hat{f})$ for complex polynomials f and study the liftability properties of probability measures supported on the Julia set $\mathcal{J}$. (This allows us to deal with some cases where critical points lie in $\mathcal{J}$.) Given a probability measure μ on $\mathcal{J}$, we construct a sequence of Cesaro means $\{\hat{\mu}_n\}_n$ on $\hat{\mathcal{J}}$, and we say that μ is *liftable* to the Markov extension if this sequence has a non-zero vague limit measure $\hat{\mu}$. The limit measure $\hat{\mu}$ is $\hat{f}$-invariant, even if the measure μ is not f-invariant. This technique is particularly useful for finding invariant probability measures that are absolutely continuous with respect to δ-conformal measure on the Julia set.

Among other things, we prove that an ergodic invariant probability measure μ is liftable if and only if its Lyapunov exponent is positive (cf. [**BK**]). Furthermore, for liftable measures, typical points are conical (i.e. go to large scale, see Lemma 9) with positive frequency. Similar results hold for (non-invariant) δ-conformal measure μ_δ. (The measure μ_δ is δ-*conformal* on $\mathcal{J}$ if $\mu_\delta(\mathcal{J}) = 1$ and $\mu(f(A)) = \int_A |Df|^\delta \, d\mu_\delta$ for all measurable sets A such that $f : A \to f(A)$ is one-to-one.) We prove that the pointwise lower Lyapunov exponent $\underline{\lambda}(z)$ is strictly greater than 0 for a set of positive μ_δ-measure if and only if μ_δ is liftable, and in this case there is an f-invariant probability measure equivalent to μ_δ. We note that this result applies to polynomials considered in [**GS, Pr, Re**], when the Julia sets of these polynomials are dendrites; see below.

When proving our results on the relation between liftability and positive Lyapunov exponents, we use the Koebe lemma: a one-dimensional tool. Work in progress aims at extending these results to higher dimensions. Our result on finding invariant probability measure absolutely continuous with respect to δ-conformal measure again uses the Koebe lemma and seems a more difficult type of result to generalize.

Markov extensions (popularly called Hofbauer towers) are less well known than the Young towers [**Y1, Y2**]. We wish to highlight the difference between these two constructions. In short, for the Young tower case, given an invariant measure μ and a subset Y of the phase space, a partition $Y = \bigcup_j Y_j \pmod{\mu}$ is constructed together with return times R_j such that $F : \bigcup_j Y_j \to Y$, $F|_{Y_j} = f^{R_j}|_{Y_j}$ and $f^{R_j} : Y_j \to Y$

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

is one-to-one and has good distortion and expansion properties. These are then used to study stochastic limit properties (e.g. mixing rates, the central limit theorem, invariance principles) of specific invariant measures. The construction is therefore linked to the choice of the measure, and may be quite involved in practical applications. The construction of the Hofbauer tower, on the other hand, is combinatorial and can be used to study all probability measures. In fact, it is exactly for the liftable invariant probability measures that Young towers can be constructed, in a canonical way, as first return maps to appropriate sets in the Markov extension; see [**Br**].

The structure of this paper is as follows. The construction of the canonical Markov extension occupies §2. We restrict our attention to polynomials f with locally connected full Julia sets (dendrites), as we need to find a finite partition $\mathcal{P}_1$ of the Julia set $\mathcal{J}$ such that f is univalent on each partition element. Such partitions may exist for the Julia set of many other rational maps as well, but is hard to give for rational maps in all generality. In §3 we describe the lifting procedure of measures. As remarked there, in contrast to subsequent sections, §3 is largely independent of the geometry of $\mathcal{J}$, and can be easily extended to Markov extensions in other settings. In §4 we introduce inducing constructions as a tool to prove that, for liftable measures, typical points will 'go to large scale' with positive frequency. Section 5 focuses on (ergodic) invariant probability measures μ and their Lyapunov exponents $\lambda(\mu)$. It is shown that μ is liftable if and only if $\lambda(\mu) > 0$. Section 6 gives a similar result for δ-conformal measure μ_δ. It shows that μ_δ is liftable if and only if the pointwise lower Lyapunov exponent $\underline{\lambda}(z) > 0$ for all z in a set of positive μ_δ-measure.

2. *The Markov extension*

Let $f : \mathbb{C} \to \mathbb{C}$ be a polynomial of degree d with a connected, locally connected and full Julia set $\mathcal{J}$ (i.e. $\mathbb{C} \setminus \mathcal{J}$ is connected). Consequently all critical points belong to $\mathcal{J}$. Let $\mathcal{C}r$ denote the critical set. It is easy to see that $\mathcal{J}$ is a dendrite, defined as follows (cf. [**Ku**]).

Definition. A metric space (X, d) is called a *dendrite* if it is connected, locally connected, and for any two points $x, y \in X$ there is a unique arc $\gamma : [0, 1] \to X$ connecting x to y.

The Fatou set $\mathcal{F}$ coincides with the basin of ∞. Let the Green function $G : \mathcal{F} \to \mathbb{R}$ be defined by $G(z) = \lim_{n\to\infty} \log|f^n(z)|/d^n$; see [**Mi**] for more details. The equipotentials (i.e. level sets) of the Green function form a foliation of $\mathcal{F}$ consisting of nested Jordan curves. The orthogonal foliation is the foliation of external rays. Each external ray is a copy of $\mathbb{R}$ embedded in $\mathcal{F}$, and if $\gamma : \mathbb{R} \to R$ is such an embedding such that $|\gamma(t)|$ is large for large t, then $\lim_{t\to\infty} \arg\gamma(t)$ is a well-defined number $\vartheta \in S^1$, called the external angle of R. Let R_ϑ denote the ray with external angle ϑ, and $\gamma_\vartheta : \mathbb{R} \to R_\vartheta$ its parameterization. It is convenient to parameterize external rays by the values of the Green function: $G(\gamma_\vartheta(t)) = t$ for each $\vartheta \in S^1$ and $t \in \mathbb{R}$. Note that $f(R_\vartheta) = R_{d\vartheta \bmod 1}$; more precisely: $f(\gamma_\vartheta(t)) = \gamma_{d\vartheta \bmod 1}(t + 1)$.

LEMMA 1. *There is a finite partition $\mathcal{P}_1$ of $\mathcal{J} \setminus \mathcal{C}r$ such that $f|_Z$ is univalent for each $Z \in \mathcal{P}_1$.*

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

Proof. Because $\mathcal{J}$ is locally connected, each external ray R_ϑ lands at a single point in $z \in \mathcal{J}$ and each $z \in \mathcal{J}$ is the landing point of at least one external ray. For each $c \in \mathcal{C}r$, select $\kappa_c \geqslant 1$ rays that land at $f(c)$ (note that Theorems 1.1 and 3.1 of [**Ki**] imply that there can be at most 2^d rays landing here). If c has degree d_c, there are $\kappa_c d_c \geqslant 2$ preimage rays landing at c. The union of these rays, together with c_j, is a locally compact set, separating the plane and also $\mathcal{J}$ into $\kappa_c d_c$ 'segments'. As $\mathcal{J}$ is closed, the closure of the segments of $\mathcal{J} \setminus \{c\}$ intersect only at c. Repeating the argument for all other critical points gives the assertion. □

Let $\mathcal{P}_0 = \{\mathcal{J}\}$ be the trivial partition of $\mathcal{J}$, and $\mathcal{P}_1$ be the partition of Lemma 1. Let $\mathcal{P}_n = \bigvee_{i=0}^{n-1} f^{-i}(\mathcal{P}_1)$, and for $z \notin \bigcup_{i=0}^{n-1} f^{-i}(\mathcal{C}r)$ let $Z_n[z]$ be the element of $\mathcal{P}_n$ containing z. Note that each $Z_n \in \mathcal{P}_n$ is connected (in fact, $\overline{Z_n}$ is a dendrite) and $f^n|_{Z_n}$ is univalent. We call Z_n an n-cylinder.

Definition. The canonical Markov extension $\hat{\mathcal{J}}$ is a disjoint union of copies D of subsets of $\mathcal{J}$, subject to an identification discussed below. We call sets D *domains* and denote their collection as by $\mathcal{D}$. Let $\pi : \hat{\mathcal{J}} \to \mathcal{J}$ be the inclusion map. Domains D are defined recursively as follows.

- The first domain $\hat{\mathcal{J}}_0$, called the *base* of the Markov extension, is a copy of $\mathcal{J}$.
- Given a domain $D \in \mathcal{D}$ and a non-empty set of the form $\overline{f(\pi(D) \cap Z)}$ for some $Z \in \mathcal{P}_1$, we let D' be a copy of $\overline{f(\pi(D) \cap Z)}$ and add it to $\mathcal{D}$. Write $D \to D'$ in this case.
- The collection $\mathcal{D}$ is such that $\hat{\mathcal{J}}_0 \in \mathcal{D}$, and $\mathcal{D}$ is closed under the previous operation.

Each $\hat{z} \in D$ can be represented by a pair (z, D) where $\hat{z} \in D$ and $\pi(\hat{z}) = z$. Moreover, any pair (z, D) defines a unique $\hat{z} \in \mathcal{J}$ whenever $z \in \pi(D)$. This allows us to define $\hat{f} : \hat{\mathcal{J}} \to \hat{\mathcal{J}}$.

- If $\hat{z} \in D$, $D \to D'$ and $\pi(\hat{z})$ belongs to the closure of $Z \in \mathcal{P}_1$ such that $\pi(D') = \overline{f(\pi(D) \cap Z)}$, then we let $\hat{f}(\hat{z}) = (f(z), D')$. Clearly
$$\pi \circ \hat{f} = f \circ \pi.$$
- If $\pi(\hat{z}) \in \mathcal{C}r$, then $\hat{f}$ can be multi-valued at $\hat{z}$, but a domain $D \in \mathcal{D}$ contains at most one of the images of $\hat{z}$. In all other cases, $\hat{f}(\hat{z})$ is a single point, belonging to a single domain D.

The next step is to define the *cutpoints*, their *ages* and *origins*, as well as the *level* of domains.

- The base $\hat{\mathcal{J}}_0$ contains no cutpoints.
- If $\hat{z} \in D$ is a cutpoint or $\pi(\hat{z}) \in \mathcal{C}r$, then each image $\hat{f}(\hat{z})$ is a cutpoint. Its *age* is
$$\begin{cases} 1 & \text{if } \pi(\hat{z}) \in \mathcal{C}r \text{ and } \hat{z} \text{ is not a cutpoint;} \\ a+1 & \text{if } \hat{z} \text{ is a cutpoint of age } a. \end{cases}$$
The set of cutpoints is denoted by $\mathcal{C}ut$.
- An *a-cutpoint* will be a cutpoint of age a. Each a-cutpoint $\hat{z}$ satisfies $\hat{z} = (f^a(c), D)$ for some $D \in \mathcal{D}$ and $c \in \mathcal{C}r$. This critical point c is called the *origin* of $\hat{z}$.
- Given a domain D, level(D) is 0 if there are no cutpoints in D, and is the maximal age of the cutpoints in D otherwise. Let $\hat{\mathcal{J}}_R$ be the union of all domains of level$(D) \leqslant R$.

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

The final step is the identification of domains wherever possible.

- Any two domains D and D' such that $\pi(D) = \pi(D')$, $\pi(D \cap \mathcal{C}ut) = \pi(D' \cap \mathcal{C}ut)$ and whose cutpoints have the same ages and origins are identified. The *canonical Markov extension* is the disjoint union of the domains, factorized over the identification described above.

The arrow relations $D \to D'$ give the Markov extension the structure of an (infinite) directed graph. This is the *Markov graph*, since by construction $(\hat{\mathcal{J}}, \hat{f})$ is Markov with respect to the partition of the domains of $\hat{\mathcal{J}}$.

For counting arguments later (see Lemma 2 and Appendix A), we must be aware of the possibility of 'moving sideways' in the Markov graph. That is, it is possible that for some domain D of $\hat{\mathcal{J}}$ there is an arrow $D \to D'$ where $\text{level}(D) = \text{level}(D')$. This occurs if D is a domain of level n containing one cutpoint $\hat{z}$ of age n and a cutpoint $\hat{z}'$ of age $n-1$. If the arc in D connecting these two cutpoints intersects $\pi^{-1}(\mathcal{C}r)$ (recall that since $\mathcal{J}$ is a dendrite, for any $z, z' \in \mathcal{J}$ there exists a unique arc in $\mathcal{J}$ connecting z to z'), then the domain D' containing $\hat{f}(\hat{z})$ will also have level n. So if $D \to D'$, then $\text{level}(D')$ can take *any* value $\leqslant \text{level}(D) + 1$.

Define $\hat{\mathcal{P}}_1$ to be the partition given by $\mathcal{D} \vee \pi^{-1}\mathcal{P}_1$. Let $\hat{\mathcal{P}}_n := \bigvee_{i=0}^{n-1} \hat{f}^{-i}(\mathcal{P}_1)$ and $\hat{\mathcal{P}}_n^R := \hat{\mathcal{P}}_n \cap \hat{\mathcal{J}}_R$.

Remark 1. The partition in Lemma 1, and hence the construction of the Markov extension, is not unique, because we have freedom in choosing the number of the rays κ_c for each critical value. However, any choice makes a valid partition. To illustrate this, assume that $f(z) = z^2 + c$ for $c \in (-2, -\frac{1}{4})$ such that $0 \in \mathcal{J}$. One is inclined to choose two (complex conjugate) rays landing at $c = f(0)$; see Figure 1 (left). This will lead to a canonical Markov extension which is very similar to the standard Markov extension constructed for interval maps. More precisely, select the domains $D \in \mathcal{D}$ such that $\pi(D) \cap \mathbb{R} \neq \emptyset$ and such that if $\pi(D) = f^n(\overline{Z_n})$, then for each $x \in \pi(D) \cap \mathbb{R}$, there is $x_0 \in \overline{Z_n} \cap \mathbb{R}$ such that $x = f^n(x_0)$. For each such D, retain $D \cap \pi^{-1}(\mathbb{R})$, and discard the rest of D as well as all other domains. Then this set with remaining graph structure is exactly the real Markov extension see Figure 1 (left, bold lines).

Choosing only one ray is possible as well; in this case, each domain in the Markov extension will be a copy of the whole Julia set, see Figure 1 (right), and they will be distinguished only by the fact that they have different (numbers of) cutpoints, and consequently different canonical neighbourhoods, see below.

We summarize some properties of $\hat{\mathcal{J}}$ in the following lemma.

LEMMA 2.

(a) *For any $a \geqslant 1$, each $D \in \mathcal{D}$ contains at most $\#\mathcal{C}r$ cutpoints of age a (and at most one for each different origin c).*

(b) *Let D and D' be domains in $\hat{\mathcal{J}}$ of the same level, sharing a cutpoint of maximal age, i.e. $\hat{p} \in D$ and $\hat{p}' \in D$ are cutpoints of age $a = \text{level}(D) = \text{level}(D')$ and $\pi(\hat{p}) = \pi(\hat{p}')$. Suppose also that $\hat{p}$ and $\hat{p}'$ have the same origin. Then $\pi(D) = \pi(D')$ or $\pi(D) \cap \pi(D') = \pi(\hat{p})$.*

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

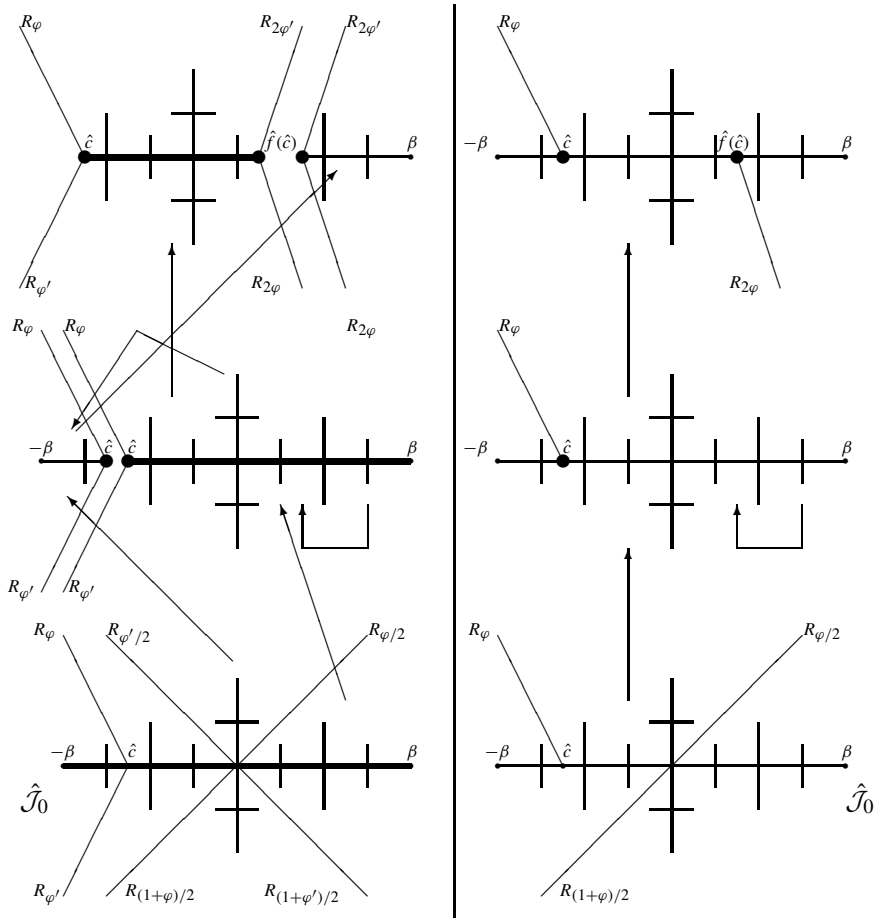


FIGURE 1. Schematic picture of some domains of the Markov extension of a quadratic map $z \mapsto z^2 + c$ using (left) two external rays R_φ and R'_φ for $\varphi' = 2\pi - \varphi$, landing at $\hat{c}$ and (right) one external ray R_φ. The pictures on the bottom line are $\hat{\mathcal{J}}_0$ and the rays landing at 0 defining the partition are shown. The point $\pi(\beta) > 0$ is fixed under f, and $\pi(-\beta)$ is its other preimage. The middle line gives domains of level 1 and the top line domains of level 2. Cutpoints are denoted by a filled circle. Arrows indicate the edges $D \to D'$ of the Markov graph. For clarity of the picture, if Z and $Z' \in \mathcal{P}_1$ are symmetric to each other, the arrow from only one of them is shown. The bold lines on the left pictures indicate the 'real Markov extension'.

(c) *The number of domains of level l is bounded by $\#\mathcal{C}r \prod_c \kappa_c$. Consequently, the number of domains in $\hat{\mathcal{J}}_R$ is at most $1 + R\#\mathcal{C}r \prod_c \kappa_c$.*

Notice that (b) implies that within a given level we can only 'move sideways' a uniformly bounded number of times.

Proof. For the proof of (a), note that if $\hat{p} \in \mathcal{C}ut \cap D$ has age a, then $\pi(\hat{p}) \in f^a(\mathcal{C}r)$. As $\hat{\mathcal{J}}$ contains no loops, only one such point exists for each c and a. So there are at most $\#\mathcal{C}r$ cutpoints of age a.

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

To prove (b), let D and D' be as in the statement, and assume that $\pi(\hat{p}) = \pi(\hat{p}') =: p = f^a(c)$, where a is the age of p and $\hat{p}$ and c their common origin. This means that there are dendrites E and E' intersecting at $f(c)$ such that $\pi(D) = f^{a-1}(E)$, $\pi(D) = f^{a-1}(E')$, and $f^{a-1}|_E$ and $f^{a-1}|_{E'}$ are homeomorphic.

Assume first that E and E' have at least an arc in common. If $E \neq E'$, say $x \in E \setminus E'$, then there is y such that $[x, f(c)] \cap E' = [y, f(c)]$. Here $[a, b]$ indicates the unique arc in E connecting a and b. By construction, each set $\overline{\mathcal{J} \setminus \pi(D')} \cap \pi(D')$ consists of postcritical points, and the same holds for $\overline{\pi(D) \setminus \pi(D')} \cap \pi(D')$. Since $f^{a-1}(E') = \pi(D')$, we have $f^n(y) \cap \mathcal{C}r \neq \emptyset$ for some $n \in \mathbb{Z}$. There are two possibilities:

- the first is that $y \in \bigcup_{n \geqslant 2} f^n(\mathcal{C}r)$, but then D' must have a cutpoint of age $> a$, contradicting maximality of a;
- the second possibility $y \in \bigcup_{n \leqslant 1} f^n(\mathcal{C}r)$. In this case there is $\tilde{c} \in \mathcal{C}r$ and $0 \leqslant s < a$ such that $f^{s-1}(y) \ni \tilde{c}$. Take y such that s is maximal with this property. Now $f^s(E)$ and $f^s(E')$ belong to the same sector defined by the $\kappa_{\tilde{c}}$ external rays landing at $f(\tilde{c})$. But then $f^s(E)$ and $f^s(E')$ both contain $f^s(x)$, a contradiction. Consequently, $\pi(D) = \pi(D')$.

Otherwise, E and E' intersect only at $f(c)$. First assume that orb(c) contains no further critical points. Then f^{a-1} is locally univalent at $f(c)$. Thus if $E \cap E' = f(c)$, then (as $\mathcal{J}$ is a dendrite, containing no loops) $f^{a-1}(E) \cap f^{a-1}(E) = p$. The final case is that there is s and $\tilde{c} \in \mathcal{C}r$ such that $f^{s-1}(p) = \tilde{c}$ and $f^s(E) \cap f^s(E')$ contains more than just $f^s(p) = f(\tilde{c})$. Then the previous argument shows that $f^s(E) = f^s(E')$, and we again obtain $\pi(D) = \pi(D')$.

Now to prove (c), note that for each $c \in \mathcal{C}r$ and domain D, $\pi^{-1}(c) \cap D$ has at most κ_c images under $\hat{f}$. Under further iteration of $\hat{f}$, this number does not increase, unless $f^s(c) = \tilde{c}$ for some $\tilde{c} \in \mathcal{C}r$, in which case the number of images can multiply by at most $\kappa_{\tilde{c}}$. The worst case is that there are $\prod_c \kappa_c$ images. In other words, for each l, there can be at most $\prod_c \kappa_c$ domains D of level l for which $\pi(D)$ pairwise intersect only at $f^l(c)$. Using (b), this gives at most $\#\mathcal{C}r \prod_c \kappa_c$ domains of level l altogether. □

LEMMA 3. *Given $\hat{z}, \hat{z}' \in \pi^{-1}(z)$, one of the following three cases occurs:*

(a) *there exists n such that $\hat{f}^n(\hat{z}) = \hat{f}^n(\hat{z}')$;*

(b) *$\bigcap_n Z_n[z]$ has positive diameter;*

(c) *$z \in \bigcup_{c \in \mathcal{C}r} \bigcup_{n \in \mathbb{Z}} f^n(c)$.*

In the latter two cases at least one of $\hat{z}, \hat{z}'$ visits any $\hat{\mathcal{J}}_R$ only finitely often.

Proof. Assume that z is not precritical, and that neither $\hat{z}$ nor $\hat{z}'$ is a cutpoint. Let D and D' be such that $\hat{z} \in D$ and $\hat{z}' \in D'$. If there is n such that the cylinder $Z_n[z]$ is contained in $\pi(D)$ as well as in $\pi(D')$, then $\hat{f}^n(D \cap \pi^{-1}(Z_n[z])) = \hat{f}^n(D' \cap \pi^{-1}(Z_n[z]))$, and $\hat{f}^n(\hat{z}) = \hat{f}^n(\hat{z}')$ as in case (a).

If on the other hand there is no such n, then $Z := \bigcap_m Z_m[z]$ has positive diameter as in case (b). Furthermore, Z contains no critical point in its interior (here we mean interior with respect to the relative topology on $\mathcal{J}$), and if $\pi(D) \not\supset Z$, then $\hat{Z} := D \cap \pi^{-1}Z$ contains a cutpoint of D. Let p be such a cutpoint of maximal age, say a. Then $\hat{f}^k(\hat{Z}) \ni \hat{f}^k(p)$ which has age $a + k$. It follows that all the sets $\hat{f}^k(\hat{Z})$, $k \geqslant 0$, are disjoint. As a result $\hat{Z}$ can remain in $\hat{\mathcal{J}}_R$ for at most R iterates.

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

If $\hat{z}$ (or $\hat{z}'$) is a cutpoint of age a then we are in case (c) and the age of $\hat{z}$ will increase under iteration of $\hat{f}$. So for any R, $\hat{f}^{R+k}(\hat{z})$ is outside $\mathcal{J}_R$ for any $k \geqslant 1$.

Finally, if z is precritical, then we are in case (c) again. It is possible that z belongs to the common boundary of several cylinder sets Z_n, and $\hat{f}^{n+1}$ is multi-valued at $\hat{z}$ and $\hat{z}'$. But each image $\hat{f}^{n+1}(\hat{z})$ and $\hat{f}^{n+1}(\hat{z}')$ is a cutpoint of its level, so it will eventually climb in the Markov extension. □

Running assumptions. We will repeatedly invoke the following assumptions on measures μ, to almost surely rule out cases (b) and (c) of Lemma 3, as explained in the next section. Typical cylinders should shrink,

$$\operatorname{diam} Z_n[z] \to 0 \text{ as } n \to \infty \text{ for } \mu\text{-a.e. } z \in \mathcal{J}, \tag{SC}$$

and the mass on the precritical points is 0,

$$\mu\left(\bigcup_{n\leqslant 0} f^n(\mathcal{C}r)\right) = 0. \tag{$\mathcal{C}r_0$}$$

By Theorem 3.2 of [**BL**], (SC) automatically follows from our assumption that f is a polynomial and $\mathcal{J}$ is locally connected and full. However, as we believe Markov extensions can be of use also when $\mathcal{J}$ is not locally connected (in which case one should think of a different partition $\mathcal{P}_1$ than the one based on external rays landing at $\mathcal{C}r$), we will refer to this property whenever we use it.

Canonical neighbourhoods. Let $U_{\hat{\mathcal{J}}_0}$ be a copy of the neighbourhood $U_{\mathcal{J}}$ of $\mathcal{J}$ bounded by the equipotential $\{G(z) = 0\}$. This is the *canonical neighbourhood* of $\hat{\mathcal{J}}_0$. We will define a canonical neighbourhood U_D for each $D \in \mathcal{D}$; they are copies of subsets of $\mathbb{C}$. The inclusion map π is extended to U_D in the natural way. In the proof of Lemma 1 we chose κ_c external rays landing at the critical value $f(c)$. The preimage rays landing at critical points (together with $\mathcal{C}r$) divide $\pi(U_{\hat{\mathcal{J}}_0})$ into $\#\mathcal{P}_1$ regions. The closure of each such region O contains exactly one element of $\mathcal{P}_1$: if $Z \in \mathcal{P}_1$, let O_Z be the corresponding region. For each $D = \overline{f(Z)}$, $Z \in \mathcal{P}_1$, let U_D be a copy of $f(O_Z) \cap U_{\mathcal{J}}$. This set is bounded by external rays landing at critical values and by the equipotential $\{G(z) = 0\}$. We call U_D the *canonical neighbourhood* of D, although it is not a neighbourhood in the strict sense: $D \setminus U_D$ consists of the cutpoints of D.

We continue recursively. If $D \to D'$ and U_D is the canonical neighbourhood of D, then $U_{D'}$ is a copy of $f(\pi(U_D) \cap O_Z) \cap U_{\mathcal{J}}$, where $Z \in \mathcal{P}_1$ is such that $\pi(D') = \overline{f(\pi(D) \cap Z)}$. It is bounded by external rays landing at cutpoints in D' and by $\{G(z) = 0\}$.

Let $\hat{U}$ be the disjoint union of all canonical neighbourhoods. Then $\hat{f}$ naturally extends univalently to $\hat{U}$ by

$$\hat{f}(z, U_D) = (f(z), U_{D'})$$

if $z \in \overline{\pi(U_D) \cap O_Z}$ where Z is such that $\pi(D') = \overline{f(Z \cap D)}$.

LEMMA 4. *The recursive definition of U_D is independent of the path $\hat{\mathcal{J}}_0 \to \cdots \to D$ by which D is reached.*

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

Proof. Since no path leads into $\hat{\mathcal{J}}_0$, its canonical neighbourhood is uniquely defined. Now take $D \in \mathcal{D}$, $D \neq \hat{\mathcal{J}}_0$ with at least two arrows leading to D. (To prove the lemma, we can restrict to domains D with two arrows rather than two paths leading to it, because when two paths eventually merge, it suffices to study those domains at which these paths merge.) For any cutpoint $\hat{z}$ of age a and origin $c \in \mathcal{C}r$, we can find O_Z with boundary point c such that $\pi(U_D) \cap B_\varepsilon(\pi(\hat{z}))$ intersects $f^a(O_Z \cap B_\varepsilon(c))$ for any $\varepsilon > 0$. Furthermore, there are rays R_φ and $R_{\varphi'}$ (or possibly only one ray) landing at c and intersecting ∂O_Z such that $f^a(R_\varphi)$ and $f^a(R'_\varphi)$ land at $\pi(\hat{z})$ and intersect $\partial\pi(U_D)$.

There are distinct arrows leading to D only if there is $D' \in \mathcal{D}$ which is identified with D. But D and D' are only identified if $\pi(D) = \pi(D')$, and the cutpoints, their origins and ages coincide. Therefore the boundaries of $\pi(U_D)$ and $\pi(U_{D'})$ comprise the same external rays together with $\{G(z) = 0\}$. It follows that $\pi(U_D) = \pi(U_{D'})$, proving the lemma. □

LEMMA 5. *The system $(\hat{U}, \hat{f})$ is Markov with respect to the partition of canonical neighbourhoods, in the sense that, if $\hat{f}(U_D) \cap U_{D'} \neq \emptyset$, then $\hat{f}(U_D) \supset U_{D'}$.*

Proof. This is a direct consequence of the previous proof. □

3. *Lifting measures*

In the previous section we introduced the Markov extensions and canonical neighbourhoods for complex polynomials. In this section we will discuss the 'liftability' properties of measures to the Markov extension in the sense of Keller. Our assumptions are (SC) and $(\mathcal{C}r_0)$. We explain how they replace conditions (2.2) and (2.3) of [**K1**]. This section gives the abstract theory which is applicable to more general settings with this type of Markov extension. In subsequent sections the precise geometry of $\mathcal{J}$, and thus the domains of $\hat{\mathcal{J}}$, play an important role again.

Given a Borel σ-algebra $\mathcal{B}$ on $\mathcal{J}$ and a Borel probability measure μ on $\mathcal{J}$, we will dynamically *lift* this measure to a Borel probability measure $\hat{\mu}$ on $\hat{\mathcal{J}}$. Our approach follows that of [**K1**]. We define a method of obtaining $\hat{\mu}$ and then show that it is $\hat{f}$-invariant and $\hat{\mu} \circ \pi^{-1} \ll \mu$.

We first introduce some notation. For some space X, we let $\mathcal{C}_0(X)$ denote the set of continuous functions $\varphi : X \to \mathbb{R}$ with compact support. For a set $A \subset X$, let $\chi_A : X \to \{0, 1\}$ be the characteristic function of A.

Let i be the trivial bijection mapping $\mathcal{J}$ to $\hat{\mathcal{J}}_0$ (note that $i^{-1} = \pi|_{\hat{\mathcal{J}}_0}$). Let

$$\hat{\mu}_0 \circ i = \mu \quad \text{and} \quad \hat{\mu}_n = \frac{1}{n} \sum_{k=0}^{n-1} \hat{\mu}_0 \circ \hat{f}^{-k}. \tag{1}$$

We will find some $\hat{\mu}$ to be a limit of a subsequence of these measures. Note that $\hat{\mathcal{J}}$ is in general not compact, so the sequence $\{\hat{\mu}_n\}_n$ may not have a subsequence with limit in the weak topology. Instead we use the vague topology; see for example [**Bi**]. Given a topological space, we say that a sequence of measures σ_n converges to a measure σ in the vague topology if, for any function $\varphi \in \mathcal{C}_0(X)$, we have $\lim_{n\to\infty} \sigma_n(\varphi) = \sigma(\varphi)$. The sequence $\{\hat{\mu}_n\}_n$ given in (1) has an accumulation point in the vague topology.

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

Definition. A probability measure μ on $\mathcal{J}$ is *liftable* if a vague limit $\hat{\mu}$ obtained in (1) is not identically 0.

Remark 2. Note that the measure $\mu \circ \pi$ on $\hat{\mathcal{J}}$ is in general σ-finite, and not $\hat{f}$-invariant. The lifted measure $\hat{\mu}$ distributes the mass of μ over the domains of $\hat{\mathcal{J}}$ so as to become invariant, as we shall see below. Indeed μ is already called liftable if part of the mass lifts to $\hat{\mathcal{J}}$; this is to accommodate non-ergodic measures μ. As in [**K1**], we generally wish to exclude the possibility of $\hat{\mu} \equiv 0$ (where all the mass escapes to infinity). Later we will find conditions to ensure that this will not happen.

The following theorem extends Theorem 2 of [**K1**] from ergodic invariant probability measures to general invariant probability measures.

THEOREM 1. *Suppose that μ is an invariant probability measure on $\mathcal{J}$ satisfying* (SC) *and* $(\mathcal{C}r_0)$. *If $\hat{\mu}$ is a vague limit point of any subsequence of measures given by (1) then it is $\hat{f}$-invariant and there is some measurable function $0 \leqslant \rho \leqslant 1$ such that $\hat{\mu} \circ \pi^{-1} = \rho \cdot \mu$.*

We will first state the theorem for ergodic invariant probability measures, and then use the ergodic decomposition to generalize to all invariant probability measures.

PROPOSITION 1. *Suppose that μ is an ergodic invariant probability measure satisfying* (SC) *and* $(\mathcal{C}r_0)$. *If $\hat{\mu}$ is a vague limit point of any subsequence of measures given by (1) and $\hat{\mu} \not\equiv 0$, then $\hat{\mu}$ is an ergodic invariant measure and $\hat{\mu} \circ \pi^{-1} = \mu$.*

Once we have shown that conditions (2.2) and (2.3) of [**K1**] can be replaced by (SC) and $(\mathcal{C}r_0)$ then the proposition follows from [**K1**, Theorem 2].

Theorem 1 of [**K1**] implies that any ergodic invariant probability measure μ can be lifted to a finite measure $\hat{\mu}$ by applying (1). The conclusions of that theorem also hold in our case. Conditions (2.2) and (2.3) of that paper need to be assumed there in order to show that the lifting process preserves ergodicity, and thus [**K1**, Theorem 2] holds. The following lemma, which takes the role of [**K1**, Lemma 1], shows that (SC) and $(\mathcal{C}r_0)$ are enough in our case to draw the same conclusion here (i.e. lifting preserves ergodicity and hence Proposition 1 holds).

Define $\mathcal{I} := \{A \in \mathcal{B} : f^{-1}(A) = A\}$ and $\hat{\mathcal{I}} := \{A \in \hat{\mathcal{B}} : \hat{f}^{-1}(A) = A\}$.

LEMMA 6. *Let μ satisfy* (SC) *and* $(\mathcal{C}r_0)$. *Suppose that $\hat{\mu}$ is a vague limit of a subsequence of $\{\mu_n\}_n$ such that $\hat{\mu} \circ \pi^{-1} = \mu$. Then $\pi^{-1}(\mathcal{I}) = \hat{\mathcal{I}} \mod \hat{\mu}$.*

Proof. Suppose that $A \in \mathcal{I}$. Then $\hat{f}^{-1} \circ \pi^{-1}(A) = \pi^{-1} \circ f^{-1}(A) = \pi^{-1}(A)$ and so $\pi^{-1}(\mathcal{I}) \subset \hat{\mathcal{I}}$.

Conversely, suppose that $\hat{A} \in \hat{\mathcal{I}}$ and let $A = \pi(\hat{A})$. Let $\hat{B} = \hat{A} \triangle \pi^{-1}(A)$. We will show that $\hat{\mu}(\hat{B}) = 0$. It follows from (SC) that $\operatorname{diam} \hat{Z}_n[\hat{z}] \to 0$ for $\hat{\mu}$-a.e. $\hat{z}$. Furthermore, $(\mathcal{C}r_0)$ implies that $\hat{\mu}_n(p) = 0$ for every n and $p \in \mathcal{C}ut$. Therefore $\hat{\mu}(p) = 0$ as well. Hence $\hat{\mu}$-a.e. z fulfilling the conditions of Lemma 3 must be in case (a) of that lemma. (This is the same as saying that $(\hat{\mathcal{J}}, \hat{f}, \hat{\mu})$ satisfies condition (2.3) of [**K1**].)

Therefore, for $\hat{\mu}$-a.e. $z_1 \in \hat{B}$, there exists $z_2 \in \hat{A}$ and $n \geqslant 1$ such that $\hat{f}^n(z_1) = \hat{f}^n(z_2)$. Hence $\hat{\mu}(\hat{B}) > 0$ implies that $\hat{A}$ is not invariant; a contradiction, whence $\hat{\mu}(\hat{A} \triangle \pi^{-1}(A)) = 0$. Thus, $\hat{\mathcal{I}} \subset \pi^{-1}(\mathcal{I}) \mod \hat{\mu}$ and the lemma is proved. □

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

Remark 3. If $\hat{\mu}$ is an ergodic invariant probability measure on $\hat{\mathcal{J}}$ such that $\hat{\mu} \circ \pi^{-1} = \mu$, then μ is liftable. This is because it can be shown that, for $\hat{\mu}_n$ defined as in (1), $\hat{\mu}_n(\hat{\mathcal{J}}_R) \geqslant \hat{\mu}(\hat{J}_R)$ for all $n, R \in \mathbb{N}$. Moreover the lift of μ is absolutely continuous (and therefore equal) to $\hat{\mu}$. Also, it follows from the proof of Theorem 2 in [**K1**] that given an ergodic invariant probability measure μ satisfying (SC) and $(\mathcal{C}r_0)$, there is at most one ergodic $\hat{f}$-invariant probability measure $\hat{\mu}$ such that $\hat{\mu} \circ \pi^{-1} = \mu$; so $\hat{\mu}$ is unique.

For liftable non-invariant measures, for example those considered in §6, the measures $\hat{\mu} \circ \pi^{-1}$ and μ are different.

Proof of Theorem 1. Let $\mathcal{B}$ the σ-algebra of μ-measurable sets, and let

$$\mu(\cdot) = \int_Y \mu_y(\cdot)\, d\nu(y) \tag{2}$$

be the ergodic decomposition of μ. More precisely, the measure space $(Y, \mathcal{C}, \nu)$ is used to index the collection of all ergodic invariant probability measures for $(\mathcal{J}, \mathcal{B})$ and the probability measure ν satisfies (2). The diagram

$$\begin{array}{ccc} (\mathcal{J}, \mathcal{B}, \mu) & \xrightarrow{f} & (\mathcal{J}, \mathcal{B}, \mu) \\ \Pi \downarrow & & \downarrow \Pi \\ (Y, \mathcal{C}, \nu) & \xrightarrow{\mathrm{id}} & (Y, \mathcal{C}, \nu) \end{array}$$

commutes, the map Π is such that $\Pi(z) = \Pi(z')$ if $f^n(z) = f^m(z')$ for some $n, m \geqslant 0$, and $\mathcal{C}$ is the finest σ-algebra such that Π is $\mathcal{B}$-measurable. For each $y \in Y$, $\Pi^{-1}(y)$ is called the *carrier* of μ_y; it is unique up to sets of μ_y-measure 0. For each $y \in Y$, Proposition 1 states that there exists a lifted measure $\hat{\mu}_y$ as the vague limit of $\{\hat{\mu}_{y,n}\}_n$ constructed as in (1) (note that the vague limit was independent of the subsequence chosen), and either $\hat{\mu}_y \equiv 0$ or $\hat{\mu}_y(\hat{\mathcal{J}}) = 1$. Let $L = \{y \in Y : \mu_y \text{ is liftable}\}$.

CLAIM. *We claim that $L \in \mathcal{C}$; more precisely, there exists $L' \in \mathcal{B}$ such that $\Pi : L' \to \mathcal{C}$ is well-defined pointwise, and $\Pi(L') = L$.*

Proof. To prove this claim, fix a countable C^0-dense subset $\hat{\Phi} := \{\hat{\varphi}_k\}_k$ of $\mathcal{C}_0(\hat{\mathcal{J}})$ and a countable collection of open intervals $\{\mathcal{U}_l\}_l$ generating the standard topology of $\mathbb{R}$. For each $y \in L$, we can use the set T_y of μ_y-typical points as carrier. (Recall that z is called μ_y-typical if the ergodic average $(1/n)\sum_{i=0}^{n-1} \varphi \circ f^i(z) \to \int \varphi\, d\mu_y$ for each continuous function $\varphi : \mathcal{J} \to \mathbb{R}$.) As μ_y is liftable, the lifted measure $\hat{\mu}_y$ has its set $\hat{T}_y$ of $\hat{\mu}_y$-typical points. Obviously $\hat{T}_y \subset \pi^{-1}(T_y)$, and if $\hat{z}_0, \hat{z}_1 \in \pi^{-1}(z) \cap \hat{T}_y$, then by Lemma 3, there is n such that $\hat{f}^n(\hat{z}_0) = \hat{f}^n(\hat{z}_1)$. Therefore $\hat{\mu}_y(\pi^{-1}(T_y) \triangle \hat{T}_y) = 0$, and, *a fortiori*, $i \circ \pi(\hat{z}) \in \hat{T}_y$ for each $\hat{z} \in \hat{T}_y$. Let L' be the set of points $z \in \mathcal{J}$ such that $i(z)$ is typical for $\hat{\mu}_y$ for some $y \in Y$. This is exactly the set of points $z \in \mathcal{J}$ such that the ergodic averages $(1/n)\sum_{j=0}^{n-1} \hat{\varphi}_k \circ \hat{f}^j(i(z))$ converge for each $k \in \mathbb{N}$, and at least one of the limits is not equal to 0 (otherwise z could only be typical for a non-liftable measure $\mu_y \in \mathcal{C}$). Let

$$X_{n,k,l} := \left\{ z \in \mathcal{J} : \frac{1}{n} \sum_{j=0}^{n-1} \hat{\varphi}_k \circ \hat{f}^j(i(z)) \in \mathcal{U}_l \right\},$$

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

then

$$L' = \Big(\bigcap_{l\in\mathbb{N}}\bigcap_{k\in\mathbb{N}}\bigcup_{N\in\mathbb{N}}\bigcap_{n\geqslant N} X_{n,k,l}\Big)\cap\Big(\bigcup_{\{l\in\mathbb{N}:0\notin\mathcal{U}_l\}}\bigcup_{k\in\mathbb{N}}\bigcup_{N\in\mathbb{N}}\bigcap_{n\geqslant N} X_{n,k,l}\Big).$$

This set is obtained using countable operations on $\mathcal{B}$-measurable sets $X_{n,k,j}$, so it belongs to the σ-algebra $\mathcal{B}$. This proves the claim. □

Let ρ be the indicator function of L'. Define

$$\hat\mu(\hat A) := \int_Y \hat\mu_y\, d\nu(y) = \int_Y\int_A \rho\circ\pi(\hat z)\, d\mu_y(\hat z)\, d\nu(y),$$

whence $\hat\mu\circ\pi^{-1} = \rho\cdot\mu$. It remains to show that $\hat\mu$ is the vague limit of the measures $\{\hat\mu_n\}_n$ constructed in (1).

Given $\varepsilon>0$, $\hat\varphi\in\mathcal{C}_0(\hat{\mathcal{J}})$, for each $y\in Y$ we can find $N=N(\varepsilon,\hat\varphi,y)$ such that

$$|\hat\mu_{y,n}(\hat\varphi)-\hat\mu_y(\hat\varphi)|<\varepsilon \quad \text{for all } n\geqslant N.$$

If μ_y is non-liftable, then $\hat\mu_y\equiv 0$; in this case $|\hat\mu_{y,n}(\hat\varphi)|<\varepsilon$ for $n\geqslant N$.

Take N_0 so large that if $Y_0=\{y\in Y : N(\varepsilon,\hat\varphi,y)>N_0\}$ then $\nu(Y_0)<\varepsilon$. Then for $n\geqslant N_0$,

$$\begin{aligned}|\hat\mu_n(\hat\varphi)-\hat\mu(\hat\varphi)| &\leqslant \int_{Y\setminus Y_0}|\hat\mu_{y,n}(\hat\varphi)-\hat\mu(\hat\varphi)|\,d\nu(y)+\int_{Y_0}|\hat\mu_{y,n}(\hat\varphi)-\hat\mu(\hat\varphi)|\,d\nu(y)\\ &\leqslant \int_{Y\setminus Y_0}\varepsilon\,d\nu(y)+2\sup\hat\varphi\;\nu(Y_0)\\ &\leqslant (1+2\sup\hat\varphi)\varepsilon.\end{aligned}$$

Since ε is arbitrary, $\hat\mu_n(\hat\varphi)\to\hat\mu(\hat\varphi)$ as required. □

We will use the following lemma often in the forthcoming sections.

LEMMA 7. *Suppose that $\hat\mu$ is some measure on $\hat{\mathcal{J}}$ obtained from applying (1) to the probability measure μ on $\mathcal{J}$. If $\hat\mu\not\equiv 0$ and $\hat\mu\circ\pi^{-1}\ll\mu$ then $\hat\nu:=\hat\mu/\hat\mu(\hat{\mathcal{J}})$ is an invariant probability measure on $\hat{\mathcal{J}}$.*

Note that the property $\hat\mu\circ\pi^{-1}\ll\mu$ is immediate if μ is invariant. Indeed, in this case $\hat\mu_n\circ\pi^{-1}=\mu$ for all n. So the lemma is useful when μ is not invariant.

Proof. Let $\hat\varphi\in\mathcal{C}_0(\hat{\mathcal{J}})$. Define $\hat\nu_0:=\nu\circ i^{-1}$. Then for any $R,n\geqslant 1$,

$$\begin{aligned}\hat\nu_n(\hat\varphi\circ\hat f\cdot\chi_{\hat{\mathcal{J}}_R}) &= \frac1n\sum_{j=0}^{n-1}\int_{\hat{\mathcal{J}}_R}\hat\varphi\circ\hat f\,d(\hat\nu_0\circ\hat f^j)=\frac1n\sum_{j=0}^{n-1}\int_{\hat{\mathcal{J}}_R}\hat\varphi\circ\hat f^{j+1}\,d\hat\nu_0\\ &= \frac1n\Big(\sum_{j=0}^{n-1}\int_{\hat{\mathcal{J}}_R}\hat\varphi\circ\hat f^j\,d\hat\nu_0+\int_{\hat{\mathcal{J}}_R}\hat\varphi\circ\hat f^n\,d\hat\nu_0-\int_{\hat{\mathcal{J}}_R}\hat\varphi\,d\hat\nu_0\Big).\end{aligned}$$

Therefore,

$$|\hat\nu_n(\hat\varphi\circ\hat f\cdot\chi_{\hat{\mathcal{J}}_R})-\hat\nu_n(\hat\varphi\cdot\chi_{\hat{\mathcal{J}}_R})|<\frac{2\sup|\hat\varphi|}{n}.$$

Letting $n,R\to\infty$ we have proved the lemma. □

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

A dynamical system (X, T, μ) is said to be *dissipative* if there is a *wandering set* of positive measure, i.e. a set $A \subset X$ with $\mu(A) > 0$ such that $\mu(T^{-n}(A) \cap A) = 0$ all $n > 0$. Otherwise the system is *conservative*. The system is *totally dissipative* if there is no set Y with $\mu(Y) > 0$ and $\mu(T^{-1}(Y) \triangle Y) = 0$ such that $(Y, T, \mu|_Y)$ is conservative.

In [**AL**], the dissipativity/conservativity of various quadratic polynomials with Feigenbaum combinatorics is investigated. For a lifted measure, we only see a conservative part of the dynamics. This can be seen in the following lemma.

LEMMA 8. *Suppose that $(\mathcal{J}, f, \mu)$ is totally dissipative and $\hat{\mu}$ is a measure obtained by applying (1) to the probability measure μ. If $\hat{\mu} \circ \pi^{-1} \ll \mu$ then $\hat{\mu} \equiv 0$.*

Proof. We start with the following claim: no measure $\hat{\mu}$ obtained by applying (1) to a probability measure μ can have wandering sets of positive $\hat{\mu}$-measure. Indeed, suppose that $\hat{A} \subset \hat{\mathcal{J}}$ has $\hat{f}^{-n}(\hat{A}) \cap \hat{A} = \emptyset$ for all $n \geqslant 1$. Suppose that $\hat{\mu}$ is a vague limit of $\{\hat{\mu}_{n_k}\}_k$. We will show that $\hat{\mu}(\hat{A}) = 0$. Note that $\sum_{i=0}^{n-1} \hat{\mu}_0(\hat{f}^{-i}(\hat{A})) \leqslant 1$ since the domains $\hat{f}^{-i}(\hat{A})$ are disjoint. Thus

$$\hat{\mu}(\hat{A}) = \lim_{k\to\infty} \frac{1}{n_k} \sum_{i=0}^{n_k-1} \hat{\mu}_0(\hat{f}^{-i}(\hat{A})) = 0.$$

This proves the claim, and hence $\hat{\mu}$ is conservative.

Now suppose that $\hat{\mu} \not\equiv 0$, then $\hat{\mu} \circ \pi^{-1}$ is an f-invariant measure which can be normalized, say

$$\mu_0 := \frac{1}{\hat{\mu}(\hat{\mathcal{J}})}\, \hat{\mu} \circ \pi^{-1}.$$

Let Y be the carrier of μ_0 (or more precisely, the union of the carriers of all liftable ergodic measures present in the ergodic decomposition of μ_0), then since $\hat{\mu} \circ \pi^{-1} \ll \mu$, $\mu(Y) > 0$ and $\mu(f^{-1}(Y) \triangle Y) = 0$. Since μ is totally dissipative, there must be a wandering set $A \subset Y$, and hence $\pi^{-1}(A)$ is wandering for $\hat{\mu}$. This contradiction proves the lemma. □

4. *Inducing*

A particularly useful property of Markov extensions is that they easily enable one to construct uniformly expanding induced systems with bounded distortion, provided the measure μ is liftable. In fact, any first return map on the Markov extension corresponds to an induced (jump) transformation of the original system, and, under mild conditions, the reverse is true as well; cf. [**Br**]. If $\hat{W} \subset \hat{\mathcal{J}}$, let us write $\hat{F}_{\hat{W}}$ for the first return map to $\hat{W}$, i.e. $\hat{F}(z) = \hat{f}^{\tau(z)}(z)$ where $\tau = \tau_{\hat{W}} : \hat{W} \to \mathbb{N}$ is the first return time to $\hat{W}$. Let $\tau^n(z)$ denote the nth return time, i.e. $\tau^1(z) = \tau(z)$ and $\tau^n(z) = \tau^{n-1}(z) + \tau(\hat{F}^{n-1}(z))$. For our purposes, we are most interested in subsets $\hat{W}$ of some domain $D \in \mathcal{D}$ that are bounded away from the cutpoints of D. As a result, any such set $\hat{W}$ is compactly contained in the canonical neighbourhood U_D of D, and by the Markov property of $(\sqcup_D U_D, \hat{f})$, any branch of $\hat{F}^n_{\hat{W}} = \hat{f}^{\tau^n(\hat{W}_0)} : \hat{W}_0 \to \hat{W}$ for any $n \in \mathbb{N}$ is extendible to a univalent onto map $\hat{f}^{\tau^n} : V_0 \to U_D$.

Given $\delta > 0$ and $M > 0$ we say that z *reaches large scale at time* j if there are neighbourhoods $\mathbb{C} \supset V_0 \supset V_1 \ni z$ such that $f^j : V_0 \to f^j(V_0)$ is univalent, $f^j(V_1)$

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

contains a round ball of radius δ (measured in Euclidean distance) and $\mathrm{mod}(V_0, V_1) > M$; see [**Mi**] for definitions. It follows from the Koebe distortion theorem, see [**Po**, Theorem 1.3], that there exists $K = K(M)$ such that the distortion

$$\mathrm{dist}(f^j|_{V_1}) := \sup_{z,z' \in V_1} \frac{|Df^j(z)|}{|Df^j(z')|} \leqslant K.$$

LEMMA 9. *Let μ be an ergodic f-invariant probability measure satisfying (SC) and ($\mathcal{C}r_0$). Then μ is liftable if and only if there exist $\delta > 0$, $v > 0$ and $M > 0$ such that for μ-a.e. $z \in \mathcal{J}$*

$$\liminf_{n\to\infty} \frac{1}{n} \#\{0 \leqslant j < n : z \text{ reaches large scale for } \delta, M \text{ at time } j\} \geqslant v. \tag{3}$$

In this case $\hat{\mu} \circ \pi^{-1} = \mu$.

Proof. First assume that μ is liftable and let $\hat{\mu}$ be the lifted measure. Let $D \in \mathcal{D}$ and let $\hat{W} \subset U_D$ be an open set bounded away from the cutpoints of D such that (using ($\mathcal{C}r_0$)) $v := \hat{\mu}(\hat{W}) > 0$. Take δ such that U_D and $\hat{W}$ contain round balls of radius δ. Also $\hat{W}$ is compactly contained in U_D, so $M := \mathrm{mod}(U_D, \hat{W}) > 0$. Let z be a typical point for μ and let $\hat{z} = i(z)$. By Birkhoff's ergodic theorem,

$$\lim_{n\to\infty} \frac{1}{n} \#\{0 \leqslant j < n : \hat{f}^j(\hat{z}) \in \hat{W}\} = \hat{\mu}(\hat{W}).$$

By the Markov property, z reaches large scale for δ, M at time j if $\hat{f}^j(\hat{z}) \in \hat{W}$. It follows that $\pi(\hat{z})$ has reached large scale at time i as well and so the first implication follows.

Conversely, suppose that μ-a.e. z satisfies (3). We say that $z \in H_R$ if, given $\hat{z}$ such that $\pi(\hat{z}) = z$ and $B_\delta(\hat{z})$ has $\mathrm{mod}(U_D, B_\delta) > M$, then $\pi^{-1}(Z_R[z])$ contains no cutpoint of D. By assumptions (SC) and ($\mathcal{C}r_0$), $\mu(H_R) \to 1$ as $R \to \infty$, say $\mu(H_R) > 1 - \eta(R)$ where $\lim_{R\to\infty} \eta(R) = 0$.

If z reaches large scale for $\delta > 0$ and $M > 0$, at iterate j, then for $\hat{z} = i(z)$ and the domain $D \ni \hat{f}^j(\hat{z})$, $\hat{Z}_R[\hat{f}^j(\hat{z})]$ contains no cutpoint of D. It follows that $\hat{f}^{j+R}(\hat{z}) \in \hat{\mathcal{J}}_R$. Therefore, given $\varepsilon > 0$ and a μ-typical point z, there exists $n_0(z)$ such that for $n \geqslant n_0(z)$

(i) $\dfrac{1}{n} \#\{0 \leqslant j < n : f^j(z) \in H_R\} > 1 - 2\eta(R)$,

(ii) $\dfrac{1}{n} \#\{0 \leqslant j < n : \hat{f}^j(\hat{z}) \in \hat{\mathcal{J}}_R\} > \dfrac{v}{1+\varepsilon}$.

Take M so large that $n_0(z) \leqslant M$ for all z in a set of μ-measure $\geqslant 1 - \varepsilon$. Then

$$\hat{\mu}_n(\hat{\mathcal{J}}_R) = \frac{1}{n} \sum_{j=1}^{n-1} \hat{\mu}_0 \circ \hat{f}^{-j}(\hat{\mathcal{J}}_R) \geqslant v\left(\frac{r-1}{r}\right)\left(\frac{1-\varepsilon}{1+\varepsilon}\right)(1 - 2\eta(R)),$$

for all $n \geqslant rM$. As $r \in \mathbb{N}$ and $\varepsilon > 0$ are arbitrary, any vague limit point of $\{\hat{\mu}_n\}_n$ satisfies $\hat{\mu}(\hat{\mathcal{J}}_R) \geqslant v(1 - 2\eta(R))$, which is positive for R sufficiently large. By Proposition 1 this means that the ergodic measure μ is liftable and $\hat{\mu} \circ \pi^{-1} = \mu$. □

Remark 4. Our notion of 'reaching large scale' with positive frequency is stronger than the notion of *induced hyperbolicity* in [**GS**]. Note also that in fact the proof above shows that

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

given $\delta, M > 0$, if

$$\liminf_{n\to\infty} \frac{1}{n}\#\{0 \leqslant j < n : z \text{ reaches large scale for } \delta, M \text{ at time } j\} > 0$$

on a positive measure set for *any* probability measure μ then the measure $\hat{\mu}$ obtained from (1) is non-zero.

If a point $\hat{z}$ visits a compact part $\hat{\mathcal{J}}_R$ of the Markov extension with positive frequency, then the majority of these visits are at a certain distance away from cutpoints in $\hat{\mathcal{J}}_R$. This is made precise in the following lemma. As a consequence, $z = \pi(\hat{z})$ will go to large scale (with bounded distortion) with positive frequency.

LEMMA 10. *Suppose that* $(\mathcal{C}r_0)$ *is satisfied. For each domain* $D \in \mathcal{D}$ *and* $\varepsilon > 0$*, there exists* $\delta > 0$ *such that if* $\hat{X} = \bigcup_{p\in\mathcal{C}ut\cap D} B_\delta(p)$*, then for every invariant probability measure* $\hat{\mu}$ *on* $\hat{\mathcal{J}}$*,* $\hat{\mu}(\hat{X}) < \varepsilon$.

Proof. Suppose that $D \in \mathcal{D}$ has level$(D) = n$ and $p \in \mathcal{C}ut_D := \mathcal{C}ut \cap D$. Then p has age $m \leqslant n$. Since the domain containing $\hat{f}^{n-m+j}(p)$ must have level at least $j + n$, p can return to D under iteration by $\hat{f}$ a maximum of $n - m$ times. Therefore, there exists $n_0 \geqslant 1$ such that $f^{n_0+k}(\mathcal{C}ut_D) \notin D$ for all $k \geqslant 1$. So for any j_0, there exists $\delta > 0$ such that $\hat{f}^{j+n_0}(B_\delta(p))\cap D = \emptyset$ for $1 \leqslant j \leqslant j_0$. Take $\hat{X} := \bigcup_{p\in\mathcal{C}ut_D} B_\delta(p)$. If $k \in \{0, \ldots, j_0\}$ and $l > k$ is such that $\hat{f}^{-k}(\hat{X}) \cap f^{-l}(\hat{X}) \neq \emptyset$, then $\hat{f}^{l-k}(\hat{X}) \cap D \neq \emptyset$. Therefore there are at most $2n_0$ numbers $l \in \{0, \ldots, n_0 + j_0\}$ such that $\hat{f}^{-k}(\hat{X}) \cap f^{-l}(\hat{X}) \neq \emptyset$. Furthermore $\hat{\mu}(\hat{f}^{-k}(\hat{X})) = \hat{\mu}(\hat{X})$. It follows by Lemma 7 that

$$1 \geqslant \hat{\mu}\left(\bigcup_{k=0}^{j_0} \hat{f}^{-k}(\hat{X})\right) \geqslant \frac{j_0}{2n_0+1}\hat{\mu}(\hat{X}).$$

To complete the proof, take $j_0 > 2n_0/\varepsilon$ and get $\delta > 0$ accordingly. □

5. *Liftability and positive Lyapunov exponents*

Given a dynamical system (X, g, ν), let $\varphi_g = \log|Dg|$ wherever this is defined and $\lambda_g(\nu) = \int \varphi_g \, d\nu$ be the Lyapunov exponent of ν. The pointwise (upper and lower) Lyapunov exponents at a point $x \in X$ are denoted as $\lambda_g(x)$ (and $\overline{\lambda}_g(x)$ and $\underline{\lambda}_g(x)$ respectively) wherever these are well defined.

PROPOSITION 2. *Suppose that* $(\mathcal{C}r_0)$ *holds. If* μ *is an ergodic invariant liftable probability measure, with lifted measure* $\hat{\mu}$*, then*

(a) φ_f *is integrable with respect to* μ*;*

(b) $\lambda_f(\mu) = \lambda_{\hat{f}}(\hat{\mu}) > 0$.

Proof. Note that $\hat{\mu}$ is an invariant measure: for example see Lemma 7. By Lemma 10, we may take domain $D \in \mathcal{D}$ and $\hat{W} \in D \cap \hat{\mathcal{P}}_n$ such that $\hat{W}$ is compactly contained in U_D and $\hat{\mu}(\hat{W}) > 0$. By the Poincaré recurrence theorem, $\hat{F}_{\hat{W}} : \bigcup_j \hat{W}_j \to \hat{W}$, the first return map to $\hat{W}$, is defined $\hat{\mu}$-a.e. Given $z \in \hat{W}_j \setminus \partial\hat{W}_j$ there is an open neighbourhood U of x such that $\hat{F}_{\hat{W}}$ extends to this neighbourhood. In particular $D\hat{F}_{\hat{W}}$ is defined for all $z \in \bigcup_j \hat{W}_j \setminus \partial\hat{W}_j$. In particular, since $\hat{\mu}(\partial\hat{W}_j) = 0$ (otherwise $(\mathcal{C}r_0)$ is contradicted),

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

the derivative is defined for $\hat{\mu}$-a.e. $z \in \hat{W}$. Each branch of $\hat{F}^n_{\hat{W}}$ is extendible to U_D, so by the Koebe distortion theorem, $\kappa := \inf\{|D\hat{F}(z)| : \hat{F}(z) \text{ is well defined}\} > 0$. In fact, there is N such that $\inf\{|D\hat{F}^N(z)| : \hat{F}^N(z) \text{ is well defined}\} \geqslant 2$ (one consequence of this is given in Remark 5 below).

The measure $\hat{\mu}_{\hat{W}} := [1/\hat{\mu}(\hat{W})]\hat{\mu}|_{\hat{W}}$ is an $\hat{F}_{\hat{W}}$-invariant probability measure, and Kac's lemma implies that

$$\int \tau \, d\hat{\mu}_{\hat{W}} = \frac{1}{\hat{\mu}(\hat{W})} < \infty,$$

where $\tau = \tau_{\hat{W}}$ is the first return time by $\hat{f}$ to $\hat{W}$. Moreover $D\hat{F}^n(\hat{z}) = D\hat{f}^{\tau^n(z)}(z)$ and if $\hat{z}$ is typical for τ, then denoting $L_f = \sup_{z\in\mathcal{J}} |Df(z)|$,

$$\begin{aligned} 0 &< \hat{\mu}(\hat{W}) \liminf_{n\to\infty} \frac{1}{n} \log |D\hat{F}^n(\hat{z})| \\ &= \hat{\mu}(\hat{W}) \lim_{n\to\infty} \frac{\tau^n(\hat{z})}{n} \liminf_{n\to\infty} \frac{1}{\tau^n(\hat{z})} \log |D\hat{f}^{\tau^n(\hat{z})}(\hat{z})| \\ &\leqslant \overline{\lambda}_{\hat{f}}(\hat{z}) \leqslant \log L_f < \infty. \end{aligned}$$

For $L < \infty$, take $\hat{\Phi}_L = \min\{L, \log |D\hat{F}(z)|\}$. Then $\hat{\Phi}_L$ is bounded and hence $\hat{\mu}_{\hat{W}}$-integrable, and for $\hat{\mu}_{\hat{W}}$-a.e. $\hat{z}$

$$\begin{aligned} 0 &< \int_{\hat{W}} \hat{\Phi}_L \, d\hat{\mu}_{\hat{W}} = \lim_{n\to\infty} \sum_{k=0}^{n-1} \hat{\Phi}_L(\hat{F}^k(\hat{z})) \\ &\leqslant \lim_{n\to\infty} \sum_{k=0}^{n-1} \log |D\hat{F}(\hat{F}^k(\hat{z}))| = \lim_{n\to\infty} \frac{1}{n} \log |D\hat{F}^n(\hat{z})| \\ &\leqslant \frac{1}{\hat{\mu}(\hat{W})} \log L_f < \infty. \end{aligned}$$

The monotone convergence theorem gives that $\log |D\hat{F}| = \lim_{L\to\infty} \hat{\Phi}_L$ is $\hat{\mu}_{\hat{W}}$-integrable and

$$\int_{\hat{W}} \log |D\hat{F}| \, d\hat{\mu}_{\hat{W}} = \lim_{n\to\infty} \frac{1}{n} \log |D\hat{F}^n(z)| \quad \hat{\mu}_{\hat{W}}\text{-a.e.}$$

We can apply the same argument to $\varphi_U := \max\{-U, \log |Df|\}$, which is μ-integrable: for μ-a.e. $z \in W := \pi(\hat{W})$ and $\hat{z} \in \hat{W}$ such that $\pi(\hat{z}) = z$,

$$\begin{aligned} \log L_f &\geqslant \int \varphi_U \, d\mu = \lim_{n\to\infty} \frac{1}{\tau^n(\hat{z})} \sum_{k=0}^{\tau^n(\hat{z})-1} \varphi_U(f^k(z)) \\ &= \lim_{n\to\infty} \frac{n}{\tau^n(\hat{z})} \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \sum_{j=\tau^k(\hat{z})}^{\tau^{k+1}(\hat{z})-1} \varphi_U(f^j(z)) \\ &\geqslant \hat{\mu}(\hat{W}) \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \log |D\hat{F}(\hat{F}^k(\hat{z}))| \\ &= \hat{\mu}(\hat{W}) \int_{\hat{W}} \log |D\hat{F}| \, d\hat{\mu}_{\hat{W}} > 0. \end{aligned}$$

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

The monotone convergence theorem implies that $\log|Df| = \lim_{U\to\infty}\varphi_U$ is μ-integrable. Hence $\underline{\lambda}_f(z) = \overline{\lambda}_f(z) = \lambda_f(z)$ for $\hat{\mu}$-a.e. z and

$$\lambda_f(z) = \lambda_{\hat{f}}(z) = \hat{\mu}(\hat{W})\lambda_{\hat{F}}(z) \geqslant \hat{\mu}(\hat{W})\frac{\log 2}{N}. \qquad \square$$

Remark 5. Our set-up of dendrite Julia sets necessarily excludes the existence of neutral periodic cycles, but also when the construction is extended to more general Julia sets, for example with Siegel discs or Leau–Fatou petals (cf. [**Mi**]), the proof of this proposition shows that Dirac measures on parabolic periodic points are not liftable to the Markov extension.

In the next result, let $\hat{W}$ and $\hat{F}$ be as in the proof of Proposition 2.

PROPOSITION 3. *If (SC) and ($\mathcal{C}r_0$) hold and μ is invariant, ergodic and liftable, then*

$$h_\mu(f) = h_{\hat{\mu}}(\hat{f}) = \hat{\mu}(\hat{W})h_{\hat{\mu}_{\hat{W}}}(\hat{F}).$$

Proof. The first equality can be shown in the same way as Theorem 3 from [**K1**]. Note that (SC) by itself does not imply that the partition $\mathcal{P}_1$ generates the Borel σ-algebra; the condition used by Keller. But Keller's proof relies on the Shannon–McMillan–Breiman theorem, which only uses that $Z_n[z] \to 0$ μ-a.e., which is indeed condition (SC). Otherwise this equality follows from the fact that a countable-to-one factor map preserves entropy, provided the Borel sets are preserved by lifting; see [**DS**].

The second equality is Abramov's formula; see [**Ab**]. $\square$

Given an invariant probability measure μ on $(\mathcal{J}, f)$, let $(\tilde{\mathcal{J}}, \tilde{f}, \tilde{\mu})$ be the natural extension. Each $\tilde{z} \in \tilde{\mathcal{J}}$ is represented as a sequence $(\tilde{z}_0, \tilde{z}_1, \dots)$ such that $\tilde{z}_j = f^j(\tilde{z}) \in \mathcal{J}$, and $\tilde{f}(\tilde{z}_0, \tilde{z}_1, \dots) = (f(\tilde{z}_0), \tilde{z}_0, \tilde{z}_1 \dots)$. Define $\tilde{\pi}_k(\tilde{z}) := \tilde{z}_k \in \mathcal{J}$.

The motivation for the following lemma and the idea for the next theorem are based on a result of [**L1**]. For the remainder of this section, we will always consider subsets of $\mathcal{J}$, and we will use the relative topology on $\mathcal{J}$.

LEMMA 11. *Suppose that μ is an invariant probability measure satisfying* (SC). *Given $\tilde{z} \in \tilde{\mathcal{J}}$, let $W(\tilde{z}) = \bigcap_{k\geqslant 1} f^k Z_k[\tilde{z}_k]$ and $r(\tilde{z}) = \sup\{\rho \geqslant 0$: in $\sup\{\rho \geqslant 0$; $B_\rho(\tilde{z}_0) \subset W(\tilde{z})\}$. If $r(\tilde{z}) > 0$ for $\tilde{\mu}$-a.e. $\tilde{z}$, then applying (1) to μ gives rise to an invariant probability measure $\hat{\mu}$ on $\hat{\mathcal{J}}$.*

Proof. Let $\varepsilon > 0$ be arbitrary and $r_0 > 0$ be such that $\tilde{\mu}(A) > 1 - \varepsilon$ for $A = \{\tilde{z} \in \tilde{\mathcal{J}} : r(\tilde{z}) > r_0\}$. Then for each $k \geqslant 0$, $A_k := \tilde{\pi}_k(A)$ satisfies $1 - \varepsilon \leqslant \mu(A_k) \leqslant \mu \circ f^{-k}(A_0)$, and for each $z \in A_k$, $B_{r_0}(f^k(z)) \subset f^k(Z_k[z])$.

Given $\hat{z} \in \hat{\mathcal{J}}$, define $D_{\hat{z}} \in \mathcal{D}$ to be the domain containing $\hat{z}$. Take R so large that if $\hat{z} \in \hat{\mathcal{J}}$ and $B_{r_0}(\hat{z})$ contains no cutpoints of $D_{\hat{z}}$ then $\hat{Z}_R[\hat{z}] \subset B_{r_0}(\hat{z})$. Define $K_R := \{\hat{z} \in \hat{\mathcal{J}} : \hat{Z}_R[\hat{z}]$ contains no cutpoint of $D_{\hat{z}}\}$. Note that if $\hat{z} \in K_R$ then $\hat{f}^R(\hat{z}) \in \hat{\mathcal{J}}_R$. If $z \in A_k$, then $f^k(Z_k[z]) \supset B_{r_0}(f^k(z))$ and by the Markov property of $(\hat{\mathcal{J}}, \hat{f})$, letting $\hat{z} := i(z)$ we have $\hat{f}^k(\hat{Z}_k[\hat{z}]) = D_{\hat{f}^k(\hat{z})} \supset B_{r_0}(\hat{f}^k(\hat{z})) \supset \hat{Z}_R[\hat{f}^k(\hat{z})]$. It follows that $\hat{f}^k(\hat{z}) \in K_R$, and therefore $\hat{f}^{k+R}(\hat{z}) \in \hat{\mathcal{J}}_R$. This shows that

$$\hat{\mu}_0 \circ \hat{f}^{-(k+R)}(\hat{\mathcal{J}}_R) \geqslant \hat{\mu}_0 \circ \hat{f}^{-k}(K_R) \geqslant \hat{\mu}_0 \circ i(A_k) = \mu(A_k) \geqslant 1 - \varepsilon$$

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

for all $k \geqslant 0$. Therefore any vague limit point $\hat{\mu}$ of $\{\hat{\mu}_n\}_n$ satisfies $\hat{\mu}(\hat{\mathcal{J}}_R) \geqslant 1-\varepsilon$, and because $\varepsilon > 0$ was arbitrary, $\hat{\mu} \not\equiv 0$ and in fact $\hat{\mu}(\hat{\mathcal{J}}) = 1$. □

THEOREM 2. *Let μ be an f-invariant probability measure satisfying* (SC) *and* ($\mathcal{C}r_0$)*, such that $\lambda_f(z) > 0$ μ-a.e. Then μ is liftable and $\hat{\mu} \circ \pi^{-1} \ll \mu$.*

Proof. We can apply [**EL**, Theorem 3.17], which says that in this setting there exists a partition η of $\tilde{\mathcal{J}}$ into open sets (recall that we are using the relative topology on $\tilde{\mathcal{J}}$ here) such that for $\tilde{\mu}$-a.e. $\tilde{z}$, $\tilde{f}$ has bounded distortion on the element $\eta(\tilde{z})$ of η containing $\tilde{z}$. More precisely, there is constant $K(\tilde{z}) \geqslant 1$ such that

$$K(\tilde{z})^{-1} \leqslant \frac{|Df^n(\tilde{x}_n)|}{|Df^n(\tilde{y}_n)|} \leqslant K(\tilde{z})$$

for all $n \geqslant 0$ and components $\tilde{x}_n$, $\tilde{y}_n$ of $\tilde{x}$, $\tilde{y} \in \eta(\tilde{x})$. Define $r(\tilde{z}) := \sup\{\rho \geqslant 0 : B_\rho(\tilde{z}_0) \subset \pi(\eta(\tilde{z}))\}$. This is a $\tilde{\mu}$-measurable function which is strictly positive $\tilde{\mu}$-a.e.

The general idea behind this result is from Ruelle, see [**Ru**], and is usually presented as a 'local unstable manifold theorem'. It is given in the complex setting in [**L2**] and is discussed in [**EL**]. An alternative proof is presented in §9 of [**PU**].

Now notice that for any $\tilde{z} \in \tilde{\mathcal{J}}$, $\eta(\tilde{z}) \subset W(\tilde{z})$ (as defined in Lemma 11), otherwise $\eta(\tilde{z}) \cap \partial f^k Z_k[\tilde{z}_k] \neq \emptyset$ for some k, which implies that distortion is unbounded; a contradiction. Let $r'(\tilde{z}) := \sup\{\rho \geqslant 0 : B_\rho(\tilde{z}_0) \subset W(\tilde{z})\}$, then $r'(\tilde{z}) \geqslant r(\tilde{z}) > 0$, $\tilde{\mu}$-a.e. By Lemma 11, applying (1) to μ gives a measure $\hat{\mu} \not\equiv 0$. It follows from Theorem 1 that $\hat{\mu} \circ \pi^{-1} \ll \mu$ and $\hat{\mu}$ is invariant. □

COROLLARY 1. *Let μ be an invariant probability measure satisfying* (SC) *and* ($\mathcal{C}r_0$)*. If the measure theoretical entropy $h_\mu(f) > 0$, then μ is liftable.*

Proof. It follows from the Ruelle inequality [**PU**] that for ergodic invariant measures, ν, $h_\nu(f) \leqslant 2\lambda(\nu)$. Therefore, if we consider the ergodic decomposition, our assumption implies that there is a positive μ-measure set of z with $\lambda_f(z) > 0$. Thus Theorem 2 implies that μ is liftable. (In the interval case, Keller [**K1**] gave a proof based on a counting argument of paths high up in the tower. This type of proof can be used here too; see Appendix A for our counting argument, which is to be used in the next section.) □

6. *Conformal measure*

In this section we discuss the liftability properties of conformal measure. Sullivan [**Su**] showed that all rational maps on the Riemann sphere have a conformal measure for at least one minimal $\delta \in (0, 2]$. We would like to emphasize that μ_δ is not invariant, but when μ_δ is liftable, then $\hat{\mu}_\delta$ (normalized) projects to an invariant probability measure, say $\nu = \alpha \cdot \hat{\mu}_\delta \circ \pi^{-1}$, where $\alpha \geqslant 1$ is the normalizing constant.

Our first lemma is that ν is absolutely continuous, generalizing Proposition 1 to δ-conformal measure. It can be expected that this lemma generalizes to other non-invariant probability measures too, provided there is distortion control.

LEMMA 12. *Suppose that a conformal measure μ_δ on $\mathcal{J}$ satisfies ($\mathcal{C}r_0$). Let $\hat{\mu}$ be a measure on $\hat{\mathcal{J}}$ obtained as a vague limit of (1). Then $\hat{\mu} \circ \pi^{-1} \ll \mu_\delta$.*

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

Proof. We suppose that $\hat{\mu}_\delta \not\equiv 0$, otherwise there is nothing to prove. Suppose that $\hat{\mu}_{n_k} \to \hat{\mu}_\delta$ as $k \to \infty$.

If the lemma is not satisfied then there exists $\varepsilon > 0$ and a set $\hat{A} \subset \hat{\mathcal{J}}$ which has $\hat{\mu}_\delta(\hat{A}) > \varepsilon$, but $\mu_\delta(A) = 0$ for $A = \pi(\hat{A})$. We may assume that $\hat{A}$ is contained in some domain $D \in \mathcal{D}$.

Due to $(\mathcal{C}r_0)$, we can assume that $\hat{A}$ is compactly contained inside U_D. Therefore $\hat{\mu}_\delta(U_D) > 0$. Choose some $\hat{B}$ compactly contained in U_D with $\mu_\delta(B) > 0$ where $B = \pi(\hat{B})$. We take some neighbourhood U containing both A and B which is compactly contained in U_D. There is some $C > 0$ such that for any $x, y \in U$, for each branch of the inverse map we have

$$\left|\frac{D\hat{f}^{-n}(x)}{D\hat{f}^{-n}(y)}\right| < C \quad \text{for all } n \geqslant 1.$$

Supposing that $\delta > 0$ is the exponent of the conformal measure, we have

$$\frac{\hat{\mu}_\delta(\hat{A})}{\hat{\mu}_\delta(\hat{B})} = \frac{\lim_{k\to\infty}(1/n_k)\sum_{j=0}^{n_k-1}\hat{\mu}_0(\hat{f}^{-j}(\hat{A}))}{\lim_{k\to\infty}(1/n_k)\sum_{j=0}^{n_k-1}\hat{\mu}_0(\hat{f}^{-j}(\hat{B}))} \leqslant C^{2\delta}\,\frac{\mu_\delta(A)}{\mu_\delta(B)}.$$

But while the left-hand side is positive, the right-hand side is 0, so we have a contradiction. Thus we obtain absolute continuity as required. □

Combining Remark 4 and Lemmas 7 and 12, we get the following corollary.

COROLLARY 2. *Suppose that μ_δ is a conformal measure satisfying (SC) and $(\mathcal{C}r_0)$. If for a given δ', $M > 0$ the set of z satisfying*

$$\liminf_{n\to\infty}\frac{1}{n}\#\{0 \leqslant j < n : z \text{ reaches large scale for } \delta', M \text{ at time } j\} > 0$$

has positive μ_δ measure, then μ_δ is liftable to some non-zero $\hat{\mu}_\delta$. Moreover $\hat{\mu}_\delta \circ \pi^{-1} \ll \mu_\delta$.

The following lemma and theorem are similar to part of the statement of Theorem B in [**L2**]. We supply a proof for completeness.

LEMMA 13. *Assume that conformal measure μ_δ satisfies (SC) and $(\mathcal{C}r_0)$. If μ_δ is liftable, and $\nu = \hat{\mu}_\delta \circ \pi^{-1}$, then μ_δ and ν are equivalent. Moreover, μ_δ and ν are ergodic.*

Proof. It was shown in Lemma 12 that $\nu \ll \mu_\delta$. Let us prove that $\psi := d\nu/d\mu_\delta$ is a positive density.

Let $A \subset J$ be an open set such that $\nu(A) > 0$. Let $\hat{\nu}$ be the measure obtained from applying (1) to ν. As $\nu = \hat{\nu} \circ \pi^{-1}$, there must be some $D \in \mathcal{D}$ such that $\hat{\nu}(D \cap \pi^{-1}(A)) > 0$. By replacing A by an appropriate cylinder set $Z \in \mathcal{P}_n$, we can assume (using (SC) and $(\mathcal{C}r_0)$) that $\hat{Z} := \pi^{-1}(Z) \cap D$ is bounded away from the cutpoints of D and such that $\hat{\nu}(\hat{Z}) > 0$. Moreover, as $\hat{Z} \in \hat{\mathcal{P}}_n$, no boundary point of $\hat{Z}$ (relative to D) returns to $\hat{Z}$.

Let $\hat{F}$ be the first return map to $\hat{Z}$. We will show that we may apply the folklore theorem to this map. First note that if $\hat{Z}$ is chosen sufficiently small, then all the branches of $\hat{F}$ are expanding and the Koebe lemma implies that they have bounded distortion. Let $\hat{B} \subset \hat{Z}$ be the set of points in $\hat{Z}$ which never return to $\hat{Z}$. We now wish to check that $\mu_\delta \circ \pi(\hat{B}) = 0$.

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

We use the same technique as in the proof of Lemma 12. By Poincaré recurrence we know that $\hat{\mu}_\delta(\hat{B}) = 0$. We let $B := \pi(\hat{B})$ and suppose that $\mu_\delta(B) > 0$, and will show that this leads to a contradiction. As in the proof of Lemma 12, we can use a distortion argument to show that

$$0 = \frac{\hat{\mu}_\delta(\hat{B})}{\hat{\mu}_\delta(\hat{Z})} = \frac{\lim_{k\to\infty}(1/n_k)\sum_{j=0}^{n_k-1}\hat{\mu}_0(\hat{f}^{-j}(\hat{B}))}{\lim_{k\to\infty}(1/n_k)\sum_{j=0}^{n_k-1}\hat{\mu}_0(\hat{f}^{-j}(\hat{Z}))} \geqslant \frac{1}{C^{2\delta}}\,\frac{\mu_\delta(B)}{\mu_\delta(Z)}.$$

But since the right-hand side is bounded away from zero, we have a contradiction.

We can now apply the folklore theorem to $[1/\mu_\delta(\pi(\hat{Z}))]\mu_\delta \circ \pi|_{\hat{Z}}$, which yields an ergodic $\hat{F}$-invariant probability measure $\hat{\nu}_{\hat{Z}}$, with density $\hat{\psi} = d\hat{\nu}_{\hat{Z}}/[d\mu_\delta \circ \pi|_{\hat{Z}}]$ bounded above and bounded away from zero. Since μ_δ is liftable and, by Lemma 12, the lifted measure $\hat{\nu}$ satisfies $\hat{\nu}\circ\pi^{-1} \ll \mu_\delta$ we have $[1/\hat{\nu}(\hat{Z})]\hat{\nu}|_{\hat{Z}} \ll \hat{\nu}_{\hat{Z}}$. Since $\hat{\nu}_{\hat{Z}}$ is ergodic and both $\hat{\nu}_{\hat{Z}}$ and $[1/\hat{\nu}(\hat{Z})]\hat{\nu}|_{\hat{Z}}$ are $\hat{F}$-invariant probability measures, $\hat{\nu}_{\hat{Z}} = [1/\hat{\nu}(\hat{Z})]\hat{\nu}|_{\hat{Z}}$. Recall that $\psi := d\nu/d\mu_\delta$. By projecting $\hat{\nu}_{\hat{Z}}$ down to the Julia set, we find that $\psi \geqslant \hat{\psi} \circ \pi^{-1} > 0$ on Z. Let $\psi_0 = \inf\{\psi(z) : z \in Z\}$. Since $Z = U \cap \mathcal{J}$ for some open set U in $\mathbb{C}$, we can find M such that $f^M(U) \supset \mathcal{J}$. Let us now prove that $\psi > 0$ for other points as well. Let $z \in \mathcal{J} \setminus \bigcup_{i=1}^{M} f^i(\mathcal{C}r)$ be arbitrary, and let $B \ni z$ be a neighbourhood of z such that $\operatorname{diam}(B) \leqslant d\big(B, \bigcup_{i=1}^{M} f^i(\mathcal{C}r)\big)$. Then there is a subset $B_0 \subset U$ such that $f^M : B_0 \to B$ is univalent. It follows that

$$\nu(B) \geqslant \nu(B_0) \geqslant \psi_0\mu_\delta(B_0) \geqslant \psi_0 \inf\{|Df^M(z)|^{-\delta} : z \in Z\}\mu_\delta(B).$$

This implies that $\psi(z) \geqslant \psi_0 \inf\{|Df^M(z)|^{-\delta} : z \in Z\} > 0$. □

The following result clarifies some properties of δ and μ_δ. The uniqueness part is due to [**DMNU**] and parts (a) and (b) are due to [**BMO**, Lemma 4.2]. We let

$$L(f) := \bigcup_{M>0}\bigcup_{\delta>0}\{z \in \mathcal{J} \text{ goes to large scale for } \delta, M > 0 \text{ infinitely often}\}.$$

Points in this set are often referred to as *conical* points. For a system (X, T, μ) we say that A is *lim sup full* if $\limsup_n \mu(T^n A) = 1$. We say that T is lim sup full if this property holds for all sets of positive measure.

THEOREM 3. *Suppose that μ_δ is a δ-conformal measure with $\mu_\delta(L(f)) > 0$. Then μ_δ is the unique measure with this property and*

(a) *f is lim sup full, exact, ergodic, conservative, μ is non-atomic,* $\operatorname{supp}(\mu) = \mathcal{J}$ *and $\omega(z) = \mathcal{J}$ for μ-a.e. $z \in \mathbb{C}$;*

(b) *δ is the minimal exponent for which a conformal measure with support on $\mathcal{J}$ exists.*

Note that (b) implies the well-known fact that for any f and $\delta > 0$ satisfying the conditions of the theorem, $\mu_{\delta'}(L(f)) = 0$ for each $\delta' > \delta$. Mayer [**Ma**] gives an example of a polynomial f such that $\mu_\delta(L(f)) = 0$ for all δ such that δ-conformal measure μ_δ exists.

We next make an alternative assumption on the behaviour of points under iteration by f which guarantees that there is some lifted measure.

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

THEOREM 4. *Suppose that* (Cr_0) *is satisfied. Let* μ_δ *be a* δ*-conformal measure on* $\mathcal{J}$*, then the following are equivalent.*

(a) *There exists* $\lambda > 0$ *such that* $\underline{\lambda}(z) \geqslant \lambda$ *for all* z *in a set of positive* μ_δ*-measure.*

(b) *The measure* μ_δ *is liftable.*

All of the situations considered in [**GS, Pr, Re**] give invariant probability measures $\mu \ll \mu_\delta$ for some δ-conformal measure with $\lambda_f(\mu) > 0$. Therefore, all of those cases fit into our setting. The closest result to ours that we know of is [**GS**] where the measure μ was obtained whenever the rational function f satisfied a summability condition on the derivatives of critical orbits.

COROLLARY 3. *If there is a liftable probability measure* $\mu \ll \mu_\delta$*, then* $\delta = \dim_{\mathrm{H}}(\mu)$ *(where* $\dim_{\mathrm{H}}$ *stands for Hausdorff dimension).*

Proof. Since μ is liftable, so is μ_δ. By Lemma 13, μ is ergodic and by Theorem 4, μ must have positive Lyapunov exponent. Pesin's formula in [**PU**, Ch. 10] implies that $\delta = \dim_{\mathrm{H}}(\mu)$. □

To prove Theorem 4, we will need the following results. Define $s_R(n, D)$ to be the maximal number of n-paths originating from an element $D \in \mathcal{D}$, $\mathrm{level}(D) = R$ and not re-entering $\hat{\mathcal{J}}_R$. Let

$$s_R(n) := \max\{s_R(n, D) : \mathrm{level}(D) = R\}.$$

LEMMA 14. *Let* $N := \#\mathcal{P}_1 \leqslant \sum_c \kappa_c d_c$*. For each* R*, there exists a* $C > 0$ *such that , for all* $0 \leqslant j < R$*,*

$$s_R(nR + j) \leqslant C(2RN)^{n+1}.$$

The proof of this lemma is in Appendix A.

Proof of Theorem 4 assuming Lemma 14. First assume that (b) holds, and let $\hat{\mu}_\delta$ be the lifted measure. By Lemma 7, $\hat{\mu}_\delta$ is invariant and by Lemma 12, $\hat{\mu}_\delta \circ \pi^{-1} \ll \mu$. Therefore Proposition 2 implies that (a) holds.

Now assume that (a) holds. We will use a counting argument to prove that a positive measure set of points must return to some $\hat{\mathcal{J}}_R$ with positive frequency, from which liftability follows.

For $1 < \lambda_0 < \lambda$, $R, n \geqslant 1$ and $\varepsilon > 0$ we consider the set

$$B_{\lambda_0,R,n}(\varepsilon) := \left\{ z : |Df^n(z)| > \lambda_0^n \text{ and } \frac{1}{n}\#\{0 \leqslant j < n : \hat{f}^j(i(z)) \in \hat{\mathcal{J}}_R\} \leqslant \varepsilon \right\}.$$

We let $\mathcal{P}_{B,n}$ denote the collection of cylinder sets of $\mathcal{P}_n$ which intersect $B_{\lambda_0,R,n}$. Since μ_δ is δ-conformal, we can compute that $\mu_\delta(B_{\lambda_0,R,n}(\varepsilon)) \leqslant \lambda_0^{-\delta n}\#\mathcal{P}_{B,n}$. We will prove that by taking $R_0 \geqslant 1$ and $\varepsilon > 0$ appropriately, this is arbitrarily small in n, which leads us to conclude that a positive measure set must visit $\hat{\mathcal{J}}_{R_0}$ with positive frequency.

Notice that any $Z \in \mathcal{P}_n$ uniquely determines a path $D_0(\hat{Z}) \to \cdots \to D_{n-1}(\hat{Z})$ in $\hat{\mathcal{J}}$ given by $\hat{Z} = i(Z)$, $\hat{f}^j(\hat{Z}) \subset D_j(\hat{Z}) \in \mathcal{D}$, and *vice versa*. In our case, given $P \in \mathcal{P}_{B,n}$, we let $\hat{P} = i(P)$, we have a path defined in $\hat{\mathcal{J}}$. Moreover, $D_j(\hat{P}) \cap \hat{\mathcal{J}}_R = \emptyset$ for at most εn of the times $j = 0, \ldots, n-1$. We will estimate $\#\mathcal{P}_{B,n}$ in terms of these paths. Define

$$S(\varepsilon, n) := \{M \subset \{0, \ldots, n-1\} : \#M \leqslant \varepsilon n\}.$$

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

The following well-known result estimates the cardinality of this set. For $x \in (0,1)$, define $l(x) := -x\log x - (1-x)\log(1-x)$.

LEMMA 15. *Let* $S(\varepsilon,n) := \{M \subset \{0,\dots,n-1\} : \#M \leqslant \varepsilon n\}$. *Then for* n *large,* $\#S(\varepsilon,n) \leqslant e^{n(\varepsilon+l(\varepsilon))}$.

Observe that for $M \in S(\varepsilon,n)$, the set $\{0,\dots,n-1\}\setminus M$ consists of at most $1+\#M$ integer-intervals. The number of 1-paths in $\hat{\mathcal{J}}_R$ is bounded by the number of domains in $\hat{\mathcal{J}}_{R+1}$. By Lemma 2(c), this is bounded above by $1+(R+1)\#\mathcal{C}r\prod_c \kappa_c$. Then choosing some large $n \geqslant 1$,

$$\begin{aligned}\#\mathcal{P}_{B,n} &\leqslant \sum_{M\in S(\varepsilon,n)} \#\{Z\in\mathcal{P}_n : j\notin M \Rightarrow D_j(i(Z))\cap\hat{\mathcal{J}}_R=\emptyset\}\\ &\leqslant \sum_{M\in S(\varepsilon,n)}\left(1+(R+1)\#\mathcal{C}r\prod_c\kappa_c\right)^{\#M} s_R(n-\#M)\\ &\leqslant \#S(\varepsilon,n)\left(1+(R+1)\#\mathcal{C}r\prod_c\kappa_c\right)^{\varepsilon n} s_R(n).\end{aligned}$$

Therefore, using Lemma 14, there exist $R_0 \geqslant 1$ and $\varepsilon_0 > 0$ such that for some $\gamma < \lambda_0^\delta$ and $C > 0$, $\#\mathcal{P}_{B,n} < C\gamma^n$. Therefore we have $\mu_\delta(B_{\lambda_0,R_0,n}(\varepsilon_0)) \leqslant C(\gamma/\lambda_0^\delta)^n$. Whence, $\mu_\delta(B_{\lambda_0,R_0,n}(\varepsilon_0)) \to 0$ as $n\to\infty$. Since, by assumption, we have $\lim_{n\to\infty}\mu_\delta\{z : |Df^n(z)| > \lambda_0^n\} > 0$, there must exist some $\varepsilon_1,\alpha>0$ such that for large enough $n\geqslant 1$,

$$\mu_\delta\left\{z : \frac{1}{n}\#\{0\leqslant j<n : \hat{f}^j(i(z))\in\hat{\mathcal{J}}_{R_0}\} > \varepsilon_1\right\} > \alpha.$$

It is now easy to see that for any vague limit $\hat{\mu}_\delta$ of measures obtained as in (1), we have $\hat{\mu}_\delta(\hat{\mathcal{J}}) \geqslant \hat{\mu}_\delta(\hat{\mathcal{J}}_{R_0}) > \alpha\varepsilon_1$. □

Acknowledgements. We are grateful to J. Hawkins, M. Urbański and J. Rivera-Letelier for fruitful discussions. Also the referee's suggestions for improving the clarity of the paper are gratefully acknowledged.

A. *Appendix*

This appendix is devoted to proving Lemma 14. We fix some $R \geqslant 1$ and a domain $D\in\mathcal{D}$, $\mathrm{level}(D) = R$. We say that:

- a t-path *survives* if the path starts in D and never falls into $\hat{\mathcal{J}}_R$;
- a domain *is surviving at time* t if it is the terminal domain of a surviving t-path;
- a cutpoint z is a *surviving cutpoint at time* t if it lies in the terminal domain of a surviving t-path.

Define

$$L_t(m) := \#\{\text{surviving } m\text{-cutpoints in } t\text{-paths starting from } D\}.$$

Since for a path $D\to\cdots\to D'$, each cutpoint in D has only one image in D', we have

$$L_t(m) \leqslant L_{t-l}(m-l) \quad \text{for } 1\leqslant l\leqslant m\leqslant t, \tag{A.1}$$

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

which is a rule we will apply repeatedly. Moreover, since the terminal domain of each surviving t-path contains at least one l-cutpoint for $R < l \leqslant R + t$, we find $L_t(j) = 0$ for $j > t + R$ and

$$L_t(1) \leqslant N \cdot \#\{\text{surviving } (t-1)\text{-paths}\} \leqslant N\Bigg(\sum_{l=R+1}^{R+t-1} L_{t-1}(l)\Bigg), \tag{A.2}$$

where $N = \#\mathcal{P}_1$. Using these rules, we prove the following lemma.

LEMMA 16. *Suppose that* $(n-1)R < t \leqslant nR$; *then*

$$L_t(j) \leqslant \begin{cases} 2^n R^n N^{n+1} & \text{if } 0 < j \leqslant t; \\ N & \text{if } t < j \leqslant t + R; \\ 0 & \text{if } t + R < j. \end{cases} \tag{A.3}$$

Proof. Since every terminal domain of a t-path has level $\leqslant t+R$, $L_t(j) = 0$ for $j > t+R$. This proves the third inequality. Before we prove the remaining part by induction, let us compute what happens for $t \leqslant R$.

$t = 0$. The maximal number of 1-cutpoints possible in a single domain is $\#\mathcal{C}r \leqslant N$. So in particular $L_0(1) \leqslant N$. By Lemma 2, $L_0(l) \leqslant N$ for $1 \leqslant l \leqslant R$.

$t = 1$. By rule (A.1), $L_1(j+1) = L_0(j) \leqslant N$ for $1 \leqslant j \leqslant R$. By rule (A.2), $L_1(1) \leqslant N$.

$t = 2$. By rule (A.1), $L_2(2 + j) = L_0(j) \leqslant N$ for $1 \leqslant j \leqslant R$. Similarly, $L_2(2) = L_1(1) \leqslant N$. Also $L_2(1) \leqslant N^2$ by rule (A.2).

$t \leqslant R$. As before

$$L_t(j) = L_0(j-t) \leqslant N \quad \text{for } t < j \leqslant t + R,$$

and

$$\begin{aligned} L_t(j) &\leqslant L_{t-j+1}(1) \leqslant N\#\{(t-j)\text{-surviving paths}\} \\ &\leqslant N \sum_{l=R+1}^{t-j+R} L_{t-j}(l) \leqslant (t-j)N^2 \leqslant RN^2 \quad \text{by rule (A.2),} \end{aligned}$$

for $0 < j \leqslant t$.

It follows that if $t < j \leqslant t + R$, then $L_t(j) = L_0(j-t) \leqslant N$. So now we have proved the second inequality of the lemma for all t, and the first inequality for $t \leqslant R$.

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

We continue by induction on n. So assume that (A.3) holds for n and that $nR < t \leqslant (n+1)R$ and $j \leqslant t$. Then

$$\begin{aligned}
L_t(j) = L_{t-j+1}(1) &\leqslant N \sum_{l=R+1}^{t-j+R} L_{t-j}(l) \quad \text{by rule (A.2)} \\
&\leqslant RN^2 + N \sum_{l=R+1}^{t-j} L_{t-j}(l) \quad \text{by the induction hypothesis for } l > t-j \\
&\leqslant RN^2 + N \sum_{l=R+1}^{t-j} L_{t-j-l+1}(1) \quad \text{by rule (A.1)} \\
&\leqslant RN^2 + N \sum_{s=1}^{t-j-R} L_s(1) \\
&\leqslant RN^2 + N \sum_{d=1}^{n} \sum_{s=(d-1)R+1}^{dR} L_s(1) \\
&\leqslant RN^2 + N \sum_{d=1}^{n} \sum_{s=(d-1)R+1}^{dR} 2^d R^d N^{d+1} \quad \text{by the induction hypothesis} \\
&\leqslant RN^2 + RN^2 \sum_{d=1}^{n} (2RN)^d \leqslant RN^2 \left(1 + \frac{(2RN)^{n+1} - 2RN}{2RN - 1}\right) \\
&\leqslant RN^2 \left(\frac{(2RN)^n}{1 - 1/(2RN)}\right) \leqslant 2^{n+1} R^{n+1} N^{n+2}.
\end{aligned}$$

This proves the induction step, and hence the lemma. □

Proof of Lemma 14. Since there cannot be more surviving nR-paths than surviving cutpoints at time nR we can estimate s_R using $L_t(j)$. We use rules (A.1) and (A.2) as in the previous proof:

$$\begin{aligned}
s_R(nR + j) &= \#\{(nR+j)\text{-surviving paths}\} \\
&\leqslant \sum_{l=R+1}^{nR+j+R} L_{nR+j}(l) \quad \text{by rule (A.2)} \\
&\leqslant RN^2 + \sum_{l=R+1}^{nR+j} L_{nR+j-l+1}(1) \quad \text{by rule (A.1) and Lemma 16} \\
&\leqslant RN^2 + \sum_{s=1}^{(n-1)R+j} L_s(1) \\
&\leqslant RN^2 + \sum_{d=1}^{n} \sum_{s=(d-1)R+1}^{dR} L_s(1) \\
&\leqslant RN^2 + \sum_{d=1}^{n} \sum_{s=(d-1)R+1}^{dR} (2RN)^{d+1} \quad \text{by Lemma 16} \\
&\leqslant C(2RN)^{n+1},
\end{aligned}$$

for some constant $C > 0$. □

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

REFERENCES

[Ab] L. M. Abramov. The entropy of a derived automorphism. *Dokl. Akad. Nauk SSSR* **128** (1959), 647–650.

[AL] A. Avila and M. Lyubich. Hausdorff dimension and conformal measures of Feigenbaum Julia sets. *Preprint*, IMS at Stony Brook #2004/05.

[Bi] P. Billingsley. *Probability and Measure*, 2nd edn. Wiley, 1986.

[BL] A. Blokh and G. Levin. An inequality for laminations, Julia sets and 'growing trees'. *Ergod. Th. & Dynam. Sys.* **22** (2002), 63–97.

[BMO] A. Blokh, J. C. Mayer and L. G. Oversteegen. Recurrent critical points and typical limit sets of rational maps. *Proc. Amer. Math. Soc.* **127** (1999), 1215–1220.

[Br] H. Bruin. Induced maps, Markov extensions and invariant measures in one-dimensional dynamics. *Comm. Math. Phys.* **168** (1995), 571–580.

[BK] H. Bruin and G. Keller. Equilibrium states for unimodal maps. *Ergod. Th. & Dynam. Sys.* **18** (1998), 765–789.

[BLS] H. Bruin, S. Luzzatto and S. van Strien. Decay of correlations in one-dimensional dynamics. *Ann. Sci. École Norm. Sup.* **36** (2003), 621–646.

[BSTV] H. Bruin, B. Saussol, S. Troubetzkoy and S. Vaienti. Return time statistics via inducing. *Ergod. Th. & Dynam. Sys.* **23** (2003), 991–1013.

[BV] H. Bruin and S. Vaienti. Return time statistics for unimodal maps. *Fund. Math.* **176** (2003), 77–94.

[Bu] J. Buzzi. Markov extensions for multi-dimensional dynamical systems. *Israel. J. Math.* **112** (1999), 357–380.

[BuK] J. Buzzi and G. Keller. Zeta functions and transfer operators for multidimensional piecewise affine and expanding maps. *Ergod. Th. & Dynam. Sys.* **21** (2001), 689–716.

[BuPS] J. Buzzi, F. Paccaut and B. Schmitt. Conformal measures for multidimensional piecewise invertible maps. *Ergod. Th. & Dynam. Sys.* **21** (2001), 1035–1049.

[BuS] J. Buzzi and O. Sarig. Uniqueness of equilibrium measures for countable Markov shifts and multidimensional piecewise expanding maps. *Ergod. Th. & Dynam. Sys.* **23** (2003), 1383–1400.

[Co] P. Collet. Statistics of closest return times for some non uniformly hyperbolic systems. *Ergod. Th. & Dynam. Sys.* **21** (2001), 401–420.

[DMNU] M. Denker, R. D. Mauldin, Z. Nitecki and M. Urbański. Conformal measures for rational functions revisited. *Fund. Math.* **157** (1998), 161–173.

[DPU] M. Denker, F. Przytycki and M. Urbański. On the transfer operator for rational functions on the Riemann sphere. *Ergod. Th. & Dynam. Sys.* **16** (1996), 255–266.

[DU] M. Denker and M. Urbański. Absolutely continuous invariant measures for expansive rational maps with rationally indifferent periodic points. *Forum Math.* **3** (1991), 561–579.

[DS] T. Downarowicz and J. Serafin. Fiber entropy and conditional variational principles in compact non-metrizable spaces. *Fund. Math.* **172** (2002), 217–247.

[EL] A. Erëmenko and M. Lyubich. The dynamics of analytic transformations. *Algebra i Analiz 1* **3** (1989), 1–70; translation in *Leningrad Math. J. 1* **3** (1990), 563–634.

[FLM] A. Freire, A. Lopes and R. Mañé. An invariant measure for rational maps. *Bol. Soc. Brasil. Mat.* **14** (1983), 45–62.

[GS] J. Graczyk and S. Smirnov. Non-uniform hyperbolicity in complex dynamics I, II. *Preprint*, http://www.math.kth.se/∼stas/papers/.

[Ha] N. Haydn. Convergence of the transfer operator for rational maps. *Ergod. Th. & Dynam. Sys.* **19** (1999), 657–669.

[Ho] F. Hofbauer. Piecewise invertible dynamical systems. *Probab. Theory Related Fields* **72**(3) (1986), 359–386.

[HK] F. Hofbauer and G. Keller. Quadratic maps without asymptotic measure. *Comm. Math. Phys.* **127** (1990), 319–337.

[K1] G. Keller. Lifting measures to Markov extensions. *Monatsh. Math.* **108** (1989), 183–200.

[K2] G. Keller. Exponents, attractors and Hopf decompositions for interval maps. *Ergod. Th. & Dynam. Sys.* **10** (1990), 717–744.

[Ki] J. Kiwi. Wandering orbit portraits. *Trans. Amer. Math. Soc.* **354** (2002), 1473–1485.

[Ku] K. Kuratowski. *Introduction to Set Theory*. Polish Scientific Publishers, Warsaw, 1977.

[L1] F. Ledrappier. Some properties of absolutely continuous invariant measures on an interval. *Ergod. Th. & Dynam. Sys.* **1** (1981), 77–93.

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press

[L2] F. Ledrappier. Quelques propriétés ergodiques des applications rationnelles. *C. R. Acad. Sci. Paris Sér. I Math.* **299** (1984), 37–40.

[Ly] M. Lyubich. Entropy properties of rational endomorphisms of the Riemann sphere. *Ergod. Th. & Dynam. Sys.* **3** (1983), 351–385.

[Ma] V. Mayer. Rational functions without conformal measures on the conical set. *Preprint*, http://math.univ-lille1.fr/∼mayer/preprints.html.

[MN] I. Melbourne and M. Nicol. Almost sure invariance principle for nonuniformly hyperbolic systems. *Comm. Math. Phys.* **260** (2005), 131–146.

[Mi] J. Milnor. *Dynamics in One Complex Variable*, 2nd edn. Vieweg, 2000.

[Ne] S. Newhouse. On some results of Hofbauer on maps of the interval. *Dynamical Systems and Related Topics (Nagoya, 1990) (Adv. Ser. Dynam. Systems, 9)*. World Scientific, River Edge, NJ, 1991, pp. 407–421.

[Po] C. Pommerenke. *Boundary Behaviour of Conformal Maps (Grundlehren der Mathematischen Wissenschaften, 299)*. Springer, Berlin, 1992.

[Pr] F. Przytycki. Iterations of holomorphic Collet–Eckmann maps: conformal and invariant measures. *Trans. Amer. Math. Soc.* **350** (1998), 717–742.

[PrRL] F. Przytycki and J. Rivera-Letelier. Statistical properties of Collet–Eckmann maps. *Preprint*, www.arXiv.org/math.DS/063459.

[PU] F. Przytycki and M. Urbański. Fractals in the plane—the ergodic theory methods. *Preprint*, http://www.math.unt.edu/˜urbanski/book1.html.

[Re] M. Rees. Positive measure sets of ergodic rational maps. *Ann. Sci École Norm. Sup.* **19** (1986), 383–407.

[Ru] D. Ruelle. Characteristic exponents and invariant manifolds in Hilbert space. *Ann. of Math.* (2) **115** (1982), 243–290.

[Su] D. Sullivan. Conformal dynamical systems. *Geometric Dynamics (Rio de Janeiro, 1981) (Lecture Notes in Mathematics, 1007)*. Springer, Berlin, 1983, pp. 725–752.

[U1] M. Urbański. On the Hausdorff dimension of a Julia set with a rationally indifferent periodic point. *Studia Math.* **97** (1991), 167–188.

[U2] M. Urbański. Geometry and ergodic theory of conformal non-recurrent dynamics. *Ergod. Th. & Dynam. Sys.* **17** (1997), 1449–1476.

[U3] M. Urbański. Measures and dimensions in conformal dynamics. *Bull. Amer. Math. Soc.* **40** (2003), 281–321.

[Y1] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.* **147** (1998), 585–650.

[Y2] L.-S. Young. Recurrence times and rates of mixing. *Israel. J. Math.* **110** (1999), 153–188.

[Zd] A. Zdunik. Harmonic measure on the Julia set for polynomial-like maps. *Invent. Math.* **128** (1997), 303–327.

https://doi.org/10.1017/S0143385706000976 Published online by Cambridge University Press