

## An invariant measure for rational maps

Alexandre Freire, Artur Lopes and Ricardo Mañé

### Introduction

Let  $\bar{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$  be the Riemann sphere and  $f : \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}$  be an analytic endomorphism of degree  $d \geq 2$ . Then  $f$  can be written as a rational function  $f(z) = P(z)/Q(z)$  where  $P$  and  $Q$  are relatively prime polynomials and either  $P$  or  $Q$  has degree  $d$ . Set  $f^n = f \circ \dots \circ f$ . The purpose of this paper is to construct an  $f$ -invariant probability that describes the asymptotic random distribution of the roots of the equation  $f^n(z) = a$ , when  $n \rightarrow +\infty$ . More precisely, denote  $z_i^{(n)}(a)$ ,  $i = 1, \dots, d^n$ , the roots of the equation  $f^n(z) = a$  (counted with algebraic multiplicity), and define a probability  $\mu_n(a)$  by:

$$\mu_n(a) = \frac{1}{d^n} \sum_{i=1}^{d^n} \delta_{z_i^{(n)}(a)}$$

Let  $\mathcal{M}$  be the space of probabilities on the Borel  $\sigma$ -algebra of  $\bar{\mathbb{C}}$  endowed with the weak topology, i.e., the unique metrizable topology on  $\mathcal{M}$  such that a sequence  $\{\mu_n | n \geq 1\} \subset \mathcal{M}$  converges to  $\mu \in \mathcal{M}$  if and only if:

$$\lim_{n \rightarrow +\infty} \int_{\bar{\mathbb{C}}} \phi d\mu_n = \int_{\bar{\mathbb{C}}} \phi d\mu$$

for every continuous  $\phi : \bar{\mathbb{C}} \rightarrow \mathbb{R}$ . We shall prove that for every  $a \in \bar{\mathbb{C}}$  (with the possible exception of two values that can be explicitly characterized) the sequence  $\mu_n(a)$  converges to an  $f$ -invariant probability  $\mu_f \in \mathcal{M}$ , independent of  $a$ , that exhibits certain interesting ergodic properties. To give the full statement of our theorem, we have to recall first the definition of the Julia set  $J(f)$  of  $f$ .  $J(f)$  is the set of points  $z \in \bar{\mathbb{C}}$  such that for every neighborhood  $U$  of  $z$ , the family  $\{f^n/U | n \geq 0\}$  is not normal. It is easy to check that  $J(f)$  is compact and satisfies  $f^{-1}(J(f)) = J(f)$ . Moreover,  $J(f)$  is the closure of the set of sources of  $f$ , i.e., points  $z$  such that  $f^n(z) = z$  and  $|(f^n)'(z)| > 1$  for some  $n \geq 1$  (Julia [4], Fatou [3]). The definition of  $J(f)$  easily implies that every  $z \notin J(f)$  has a neighborhood  $U$  where the family of iterates  $f^n/U \rightarrow \bar{\mathbb{C}}$  is equicontinuous. In other words, in the complement of  $J(f)$  the dynamics of  $f$  is extremely stable (in the sense of Lyapounov). On the other hand, every neighborhood of a point in



the Julia set is expanded under forward iteration. In fact, if  $U$  is a neighborhood of  $z \in J(f)$ , the non-normality of  $\{f^n/U \mid n \geq 1\}$  implies (by Montel's characterization of normal families) that for large values of  $n$ ,  $\bigcup_{m \geq n} f^m(U)$  covers the whole sphere except for at most two points. With some more work (see Brolin [1]), it can be proved that there exists a set  $\text{Exc}(f) \subset \bar{\mathbb{C}}$ , whose elements are called exceptional points, containing at most two points, and such that for every neighborhood  $U$  of a point in  $J(f)$ , there exists  $N > 0$  such that  $f^n(U) = \text{Exc}(f)^c$  for every  $n \geq N$ . Moreover, the points of  $\text{Exc}(f)$  can be described as follows. If  $\text{Exc}(f)$  contains only one point  $p$ , then it must satisfy  $f^{-1}(p) = \{p\}$ . Then if  $L: \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}$  is a Möbius transformation such that  $L(p) = \infty$ , it is easy to check that  $(LfL^{-1})^{-1}(\infty) = \{\infty\}$ , and this implies that  $LfL^{-1}$  is a polynomial. If  $\text{Exc}(f)$  contains two points  $p$  and  $q$ , they must satisfy  $f^{-1}(\{p, q\}) = \{p, q\}$ . Taking a Möbius transformation  $L: \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}$  such that  $L(p) = \infty$ ,  $L(q) = 0$ , it follows that  $(L^{-1}fL)^{-1}(\{0, \infty\}) = \{0, \infty\}$ . This property implies that  $(L^{-1}fL)^{-1}(z) = \alpha z^{\pm d}$ , for some  $\alpha \in \bar{\mathbb{C}}$ .

**Theorem.** *There exists an  $f$ -invariant probability  $\mu_f$  satisfying the following properties:*

- $\lim_{n \rightarrow +\infty} \mu_n(a) = \mu_f$  for every  $a \notin \text{Exc}(f)$ . Moreover, this convergence is uniform when  $a$  varies in a compact subset of  $\text{Exc}(f)^c$ .
- The support of  $\mu_f$  is  $J(f)$ .
- $f$  is a  $K$ -system with respect to  $\mu_f$ .
- $u_f(f(A)) = d\mu_f(A)$  for every Borel set  $A \subset \bar{\mathbb{C}}$  such that  $f|_A$  is injective. Conversely,  $\mu_f$  is the unique  $f$ -invariant probability satisfying this property.
- $h_{u_f}(f) = \log d$ .

Since the definition of  $K$ -system used sometimes in Ergodic Theory applies only to invertible transformations, that is not our case, it is perhaps useful for the reader to explain property (c). We shall that if  $\mathcal{A}$  is the Borel  $\sigma$ -algebra of  $\bar{\mathbb{C}}$ , then  $\bigcap_{n \geq 0} f^{-n}(\mathcal{A})$  contains only sets of measure 0 or 1. It is natural to include these transformations (as is done by several authors) in the class of  $K$ -systems. As in the invertible case (and for the same reasons), it implies the mixing property but is much stronger.

In [1], Brolin proved the existence of  $\mu_f$  satisfying (a), (b) and a weaker form of (c) (with mixing instead of  $K$ -system) for the case of polynomial mappings. His methods, based in Potential Theory, do not extend to general rational maps. On the other hand, when  $f$  is a polynomial, these methods give a remarkable identification of  $\mu_f$ , namely, that  $\mu_f$

is the equilibrium distribution (in the sense of Potential Theory) associated to  $J(f)$ . This means, roughly speaking, that  $\mu_f$  describes the way a unit positive electric charge would be distributed in  $J(f)$  under equilibrium conditions (for the formal definition, see Brolin [1]). Unfortunately, this beautiful characterization of  $\mu_f$  doesn't extend to general rational maps. For instance, take

$$f(z) = \prod_{i=1}^m \left( \frac{z - a_i}{1 - \bar{a}_i z} \right)$$

where  $|a_i| < 1$  for all  $i$  and not all are zero. It is easy to check that  $J(f)$  is the unit circle and that  $f|_{J(f)}$  is an expanding endomorphism, i.e.,  $\lim_{n \rightarrow +\infty} |(f^n)'(z)v| = +\infty$  for every  $z \in J(f)$  and every  $0 \neq v$  tangent to  $J(f)$  at  $z$ . The equilibrium distribution in the unit circle is the Lebesgue measure  $\lambda$ . We shall show that  $\mu_f$  is singular with respect to  $\lambda$ . By the theory of expanding endomorphisms of manifolds, there exists an  $f$ -invariant ergodic probability  $\nu$  on  $J(f)$  that is equivalent to  $\lambda$  and such that  $d\nu/d\lambda$  is a strictly positive continuous function. Since both  $\nu$  and  $\mu_f$  are ergodic, they are either singular or equal. Suppose that  $\nu = \mu_f$  and set  $H = d\mu_f/d\lambda$ . From property (d) it follows that:

$$m = |f'(z)| \frac{H(f(z))}{H(z)}$$

for a.e.  $z$ . Hence  $H$  is continuous, this property holds for every  $z$ . Therefore

$$|f'(z)| = m$$

for every fixed point  $z$  of  $f$ . It is not difficult to show that there exist values of  $a_1, \dots, a_m$  such that this condition is not satisfied. Then, for these values the probabilities  $\mu_f$  and  $\nu$  cannot coincide. Hence, they are singular and  $\mu_f$  is singular with respect to the Lebesgue measure of the unit circle.

## I. Proof of the Theorem

The proof of the Theorem will be based in the following definitions and lemma. We say that a set  $\gamma \subset \bar{\mathbb{C}}$  is an arc if it is homeomorphic to the interval  $[0, 1]$ . A set  $U \subset \bar{\mathbb{C}}$  is a topological disk if it is homeomorphic to the disk  $D = \{z \mid |z| < 1\}$ .

**Definition I.** We say that a set  $U \subset \bar{\mathbb{C}}$  is  $(N, \varepsilon)$ -adapted if for all  $n \geq N$  there exist topological disks  $S_i^{(n)}$ ,  $i = 1, \dots, \ell_n$  and integers  $1 \leq k_i^{(n)}$  such that:



- a)  $f^n/S_i^{(n)}$  is a  $k_i^{(n)}$ -to-1 map onto  $\bar{U}$ .  
 b)  $\text{diam}(S_i^{(n)}) \leq \varepsilon$   
 c)  $\sum_{i=1}^{\ell_n} k_i^{(n)} \geq (1 - \varepsilon)d^n$ .  
 d)  $\lim_{n \rightarrow +\infty} (\sup_i \text{diam}(S_i^{(n)})) = 0$ .

**Definition II.** We say that two points  $z_i \in \bar{\mathbb{C}}$ ,  $i = 1, 2$ , are  $(N, \varepsilon)$ -related if for all  $n \geq N$  the roots  $z_i^{(n)}(z_1)$ ,  $i = 1, \dots, d^n$  of the equation  $f^n(z) = z_1$  and the roots  $z_i^{(n)}(z_2)$ ,  $i = 1, \dots, d^n$  of the equation  $f^n(z) = z_2$  can be indexed in such a way that:

$$d(z_i^{(n)}(z_1), z_i^{(n)}(z_2)) \leq \varepsilon$$

for all  $1 \leq i \leq t_n$ , where  $t_n$  satisfies

$$t_n \geq (1 - \varepsilon)d^n.$$

An important property that links these two definitions is that if  $z_i$  belongs to an  $(N, \varepsilon)$ -adapted set  $U_i$ ,  $i = 1, 2$ , and  $U_1 \cap U_2 \neq \emptyset$ , then  $z_1$  and  $z_2$  are  $(N, 2\varepsilon)$ -related. The proof is immediate and we leave it to the reader.

**Fundamental Lemma.** Given  $\varepsilon > 0$ ,  $z \notin \text{Exc}(f)$  and an arc  $\gamma$  containing  $z$  and such that  $\gamma - \{z\}$  doesn't contain critical values of  $f^n$  for all  $n \geq 1$ , there exists an  $(N, \varepsilon)$ -adapted set  $U \supset \{\gamma\}$  for some  $N \geq 1$ .

The proof of this Lemma will be given in the next section. Now let us show some of its corollaries.

**Corollary I.** Given a compact set  $K \subset \text{Exc}(f)^c$  and  $\varepsilon > 0$ , there exists  $N = N(K, \varepsilon) > 0$  such that any couple of points in  $K$  is  $(N, \varepsilon)$ -related.

*Proof.* First, we shall prove that if  $z_1$  and  $z_2$  are in  $K$ , there exist  $N > 0$  and an open set  $V \supset \{z_1, z_2\}$  such that any couple of points in  $V$  is  $(N, \varepsilon/2)$ -related. For this, take arcs  $\gamma_i \supset \{z_i\}$ ,  $i = 1, 2$  satisfying the hypothesis of the Fundamental Lemma and  $\gamma_1 \cap \gamma_2 \neq \emptyset$ . Then there exist  $(N_i, \varepsilon/2)$ -adapted sets  $U_i \supset \gamma_i$ ,  $i = 1, 2$ . Taking  $N = \max\{N_1, N_2\}$ , the remark after Definition II concludes the proof of the property. Now define  $\tilde{N} : K \times K \rightarrow \mathbb{Z}^+$  by the following property:  $N(z_1, z_2)$  is the minimum  $N > 0$  such that there exist neighborhoods  $V_i$  of  $z_i$ ,  $i = 1, 2$ , such that every point in  $V_1$  is  $(N(z_1, z_2), \varepsilon)$ -related to every point in  $V_2$ . The previous property shows that  $N$  is well defined. Moreover, it is obviously upper semicontinuous. Then it is bounded. Let  $N > 0$  be an upper bound. Clearly,  $N$  satisfies the required property.

**Corollary II.** If  $K \subset \text{Exc}(f)^c$  is compact, for all  $\varepsilon_1 > 0$  and every continuous function  $\phi : \bar{\mathbb{C}} \rightarrow \mathbb{R}$ , there exists  $N = N(K, \varepsilon_1, \phi) > 0$  such that:

$$\left| \int_{\bar{\mathbb{C}}} \phi d\mu_n(z_1) - \int_{\bar{\mathbb{C}}} \phi d\mu_n(z_2) \right| \leq \varepsilon_1$$

for every  $z_1$  and  $z_2$  in  $K$  and  $n \geq N$ .

*Proof.* Take  $\varepsilon > 0$  such that:

$$\varepsilon \sup_z |\phi(z)| \leq \varepsilon_1/2$$

and  $|\phi(z') - \phi(z'')| \leq \varepsilon_1/2$  if  $d(z', z'') \leq \varepsilon$ . Take  $N = N(K, \varepsilon)$  given by Corollary I. If  $n \geq N$ , by the definition of  $(N, \varepsilon)$ -related points, given  $z_1$  and  $z_2$  in  $K$ , we can arrange the roots  $z_i^{(n)}(z_1)$ ,  $i = 1, \dots, d^n$  of  $f^n(z) = z_1$ , and the roots  $z_i^{(n)}(z_2)$ ,  $i = 1, \dots, d^n$  of  $f^n(z) = z_2$ , in such a way that  $d(z_i^{(n)}(z_1), z_i^{(n)}(z_2)) \leq \varepsilon$  for  $i = 1, \dots, s_n$ , where  $s_n$  satisfies  $s_n \geq (1 - \varepsilon)d^n$ . Then:

$$\begin{aligned} \left| \int_{\bar{\mathbb{C}}} \phi d\mu_n(z_1) - \int_{\bar{\mathbb{C}}} \phi d\mu_n(z_2) \right| &\leq \frac{1}{d^n} \sum_{i=1}^{d^n} |\phi(z_i^{(n)}(z_1)) - \phi(z_i^{(n)}(z_2))| \leq \\ &\leq \frac{1}{d^n} \sum_{i=1}^{s_n} |\phi(z_i^{(n)}(z_1)) - \phi(z_i^{(n)}(z_2))| + \frac{d^n - s_n}{d^n} \sup_z |\phi(z)| \leq \\ &\leq \frac{1}{d^n} \frac{\varepsilon_1}{2} s_n + \varepsilon \sup_z |\phi(z)| \leq \frac{\varepsilon_1}{2} + \varepsilon \sup_z |\phi(z)| \leq \frac{\varepsilon_1}{2} + \frac{\varepsilon_1}{2} = \varepsilon_1. \end{aligned}$$

**Corollary III.** Given a compact set  $K \subset \text{Exc}(f)^c$ ,  $\varepsilon_1 > 0$  and a continuous function  $\phi : \bar{\mathbb{C}} \rightarrow \mathbb{R}$ , there exists  $N > 0$  such that:

$$\left| \int_{\bar{\mathbb{C}}} \phi d\mu_n(z) - \int_{\bar{\mathbb{C}}} \phi d\mu_m(z) \right| \leq \varepsilon_1$$

for every  $z \in K$  and  $m \geq n \geq N$ .

*Proof.* Set  $\tilde{K} = \bigcup_{n \geq 0} f^{-n}(K)$  and take  $N = N(\tilde{K}, \phi, \varepsilon_1)$  given by Corollary II.

Using the notation of the introduction, we can write for all  $m > n$  and  $z \in \bar{\mathbb{C}}$ :

$$\int_{\bar{\mathbb{C}}} \phi(z) d\mu_m(z) = \frac{1}{d^k} \sum_{i=1}^{d^k} \int_{\bar{\mathbb{C}}} \phi_n d\mu_n(z_i^{(k)}(z))$$

where  $k = m - n$ . Then, if  $z \in K$ ,

$$\left| \int_{\bar{\mathbb{C}}} \phi d\mu_n(z) - \int_{\bar{\mathbb{C}}} \phi d\mu_m(z) \right| \leq \frac{1}{d^k} \sum_{i=1}^{d^k} \left| \int_{\bar{\mathbb{C}}} \phi d\mu_n(z) - \int_{\bar{\mathbb{C}}} \phi d\mu_n(z_i^{(k)}(z)) \right|.$$



But  $z_i^{(k)} \in \tilde{K}$  for all  $n \geq 0$ . Hence, if  $n \geq N$ , the last term above is bounded by  $d^{-k}(d^k \varepsilon_1) = \varepsilon_1$  by Corollary II.

Now we are ready to prove the theorem. By Corollary III, if  $z \in \text{Exc}(f)^c$ , the sequence  $\{\mu_n(z) \mid n \geq 0\}$  converges, in the topology of  $\mathcal{M}$ , to a probability  $\mu_f(z)$ . Moreover, also by Corollary III, this convergence is uniform on compact sets of  $\text{Exc}(f)^c$ . By Corollary II, the probability  $\mu_f(z)$  is independent of  $z$ . Denote it  $\mu_f$ . Moreover,  $\mu_f$  is invariant because for every continuous  $\phi: \bar{C} \rightarrow \mathbb{R}$ , taking any  $a \in \text{Exc}(f)^c$

$$\begin{aligned} \int_{\bar{C}} (\phi \circ f) d\mu_f &= \lim_{n \rightarrow +\infty} \frac{1}{d^n} \sum_{i=1}^{d^n} \phi(f(z_i^{(n)}(a))) = \\ &= \lim_{n \rightarrow +\infty} \frac{1}{d^{n-1}} \sum_{i=1}^{d^{n-1}} \phi(z_i^{(n-1)}(a)) = \int_{\bar{C}} \phi d\mu_f. \end{aligned}$$

This completes the proof of the existence of an  $f$ -invariant probability  $\mu_f$  satisfying (a). Now let us prove property (b). Take  $a \in J(f)$ . Then the support of  $\mu_n(a)$  is contained in  $J(f)$  for all  $n \geq 1$ . Then the support of  $\mu_f$  is contained in  $J(f)$ . Conversely, if  $p \in J(f)$ , we shall prove that it belongs to the support of  $\mu_f$ . Take a point  $q$  in the support of  $\mu_f$ . Since  $p$  and  $q$  are not exceptional points (because they belong to  $J(f)$ ), there exist a sequence  $p_n \rightarrow p$  and a sequence of integers  $m_n \rightarrow +\infty$  such that  $f^{m_n}(p_n) = q$ . Since the support is closed, it is sufficient to show that  $p_n$  belongs to the support of  $\mu_f$  for all  $n$ . Given any neighborhood  $V$  of  $p_n$ , we take a neighborhood  $U$  of  $p_n$  such that  $U \subset V$ ,  $\mu_f(\partial U) = 0$  and also  $\mu_f(\partial(f^{m_n}(U))) = 0$ . Moreover,  $f^{m_n}(U)$  is an open set containing  $q$ . Since the boundaries of  $U$  and  $f^{m_n}(U)$  have measure zero, we can write:

$$(1) \quad \mu_f(U) = \lim_{n \rightarrow +\infty} \mu_n(a)(U) = \lim_{j \rightarrow +\infty} \frac{1}{d^j} \# \{z \in U \mid f^j(z) = a\}$$

$$(2) \quad \mu_f(f^{m_n}(U)) = \lim_{j \rightarrow +\infty} \mu_j(a)(f^{m_n}(U)) = \lim_{j \rightarrow +\infty} \frac{1}{d^j} \# \{z \in f^{m_n}(U) \mid f^j(z) = a\}$$

where  $a$  is any point in  $\text{Exc}(f)^c$  that is not a critical value of any  $f^j$  (in order to grant that every root of  $f^n(z) = a$  is simple). Moreover:

$$(3) \quad \# \{z \in U \mid f^j(z) = a\} \geq \# \{z \in f^{m_n}(U) \mid f^{j-m_n}(z) = a\}.$$

From (1), (2) and (3) it follows that  $\mu_f(U) \geq d^{m_n} \mu_f(f^{m_n}(U))$ . But  $\mu_f(f^{m_n}(U)) > 0$  because  $f^{m_n}(U)$  is a neighborhood of  $q$  that is in the support of  $\mu_f$ . Hence,  $\mu_f(U) > 0$  and then  $\mu_f(V) \geq \mu_f(U) > 0$ .

To prove (c), we shall show that if  $\mathcal{A}$  denotes the Borel  $\sigma$ -algebra of  $J(f)$ , then  $\bigcap_{n \geq 0} f^{-n}(\mathcal{A})$  contains only sets of measure 0 or 1. Suppose

that  $S \in \bigcap_{n \geq 0} f^{-n}(\mathcal{A})$  and  $\mu_f(S) > 0$ . We have to prove that  $\mu_f(S) = 1$ . Take  $0 < \varepsilon < \mu_f(S)/2$ . Take  $N > 0$ ,  $m > 0$  such that there exists a family  $\mathcal{P} = \{P_1, \dots, P_r\}$  of disjoint  $(N, \varepsilon/2)$ -adapted sets such that:

$$\mu_f\left(\bigcup_i P_i\right)^c = 0.$$

The existence of this partition is based on the fact that a topological disk contained in an  $(N, \varepsilon/2)$ -adapted set is also an  $(N, \varepsilon/2)$ -adapted set. Then to every point  $z$  in  $J(f)$ , by the Fundamental Lemma and the previous remark, we can associate an integer  $N(z)$  and a disk  $B(z)$  centered at  $z$ , with  $\mu_f(\partial B(z)) = 0$ , that is  $(N(z), \varepsilon/2)$ -adapted. Take  $z_i \in J(f)$ ,  $i = 1, \dots, t$  such that  $J(f) \subset \bigcup_i B(z_i)$  and set  $N = \max_i N(z_i)$ . Finally, take as sets

$P_i$ , all the intersections of sets  $B(z_i)$  or  $B^c(z_i)$  that have measure  $\neq 0$ . Denote  $P_{i,j}^{(n)}$ ,  $k_{i,j}^{(n)}$ ,  $j = 1, \dots, \ell(n, i)$  the sets and integers associated to  $P_i$  by Definition I. We shall need two lemmas:

**Lemma I.** Let  $V \subset \bar{C}$  be an open set with  $\mu_f(\partial V) = 0$  such that there exist open sets  $V_1, \dots, V_d$  such that for all  $1 \leq i \leq d$ ,  $f|_{V_i}$  is a homeomorphism of  $V_i$  onto  $V$ . Then

$$d\mu_f(f^{-1}(A) \cap V_i) = \mu_f(A)$$

for every Borel set  $A \subset V$  and all  $1 \leq i \leq d$ .

*Proof.* First we shall prove the lemma when  $A$  is open and  $\mu_f(\partial A) = 0$ . By the hypothesis  $\mu_f(\partial A) = 0$ , we can calculate  $\mu_f(A)$  by:

$$(1) \quad \mu_f(A) = \lim_{n \rightarrow +\infty} \mu_n(a)(A) = \lim_{n \rightarrow +\infty} \frac{1}{d^n} \# \{z \in A \mid f^n(z) = a\}$$

where  $a$  is any point that is not a critical value of  $f^n$  for all  $n \geq 1$ . If  $\mu_f(\partial A) = 0$ , then  $\mu_f(\partial(f^{-1}(A) \cap V_i)) \leq \mu_f(f^{-1}(\partial A)) = \mu_f(\partial A) = 0$ . Hence,

$$(2) \quad \mu_f(f^{-1}(A) \cap V_i) = \lim_{n \rightarrow +\infty} \frac{1}{d^n} \# \{z \in f^{-1}(A) \cap V_i \mid f^n(z) = a\}.$$

By the relation  $f(f^{-1}(A) \cap V_i) = A$  and the injectivity of  $f|_{V_i}$ ,

$$(3) \quad \# \{z \in f^{-1}(A) \cap V_i \mid f^n(z) = a\} = \frac{1}{d} \# \{z \in A \mid f^{n-1}(z) = a\}.$$

Then, (1), (2) and (3) prove the lemma in this case. If  $A \subset V$  is any Borel set, we can write it as  $A = \left(\bigcap_{n \geq 0} A_n\right) \cup A$  where  $A_1 \supset A_2 \supset \dots$  is a sequence of open sets with  $\mu(\partial A_i) = 0$  for all  $i$  and  $\mu_f(A_\infty) = 0$ . Then:



$$\begin{aligned}\mu_f(f^{-1}(A) \cap V_i) &= \lim_{n \rightarrow +\infty} \mu_f(f^{-1}(A_n) \cap V_i) - \mu_f(f^{-1}(A_\infty) \cap V_i) = \\ &= \lim_{n \rightarrow +\infty} \mu_f(f^{-1}(A_n) \cap V_i) = \frac{1}{d} \lim_{n \rightarrow +\infty} \mu_f(A_n) = \frac{1}{d} \mu_f(A).\end{aligned}$$

**Corollary.**  $\mu_f(f(A)) \leq d\mu_f(A)$  for every Borel set  $A$ .

*Proof.* Take a family  $V_1, V_2, \dots$  of disjoint topological disks satisfying:

- $f|_{V_i}$  is a homeomorphism onto  $f(V_i)$ .
- $\mu_f(\bigcup_i V_i) = 0$  and  $\mu_f(\partial f(V_i)) = 0$  for all  $i$ .
- $f(V_i)$  doesn't contain critical values of  $f$ , for all  $i$ . Then,

$$\begin{aligned}\mu_f(f(A)) &= \mu_f(f((\bigcup_i (A \cap V_i)) \cup (\bigcup_i A \cap \partial V_i))) \leq \\ &\leq \sum_i \mu_f(f(A \cap V_i)) + \sum_i \mu_f(f(A \cap \partial V_i)) \leq \sum_i \mu_f(f(A \cap V_i)) + \\ &\quad + \sum_i \mu_f(f(A) \cap \partial f(V_i)) = \sum_i \mu_f(f(A \cap V_i)).\end{aligned}$$

But by (a) and (c), we can apply Lemma I to obtain:

$$\mu_f(f(A \cap V_i)) = d\mu_f(A \cap V_i).$$

Hence, by (b):

$$\mu_f(f(A)) \leq d \sum_i \mu_f(A \cap V_i) = d\mu_f(A \cap (\bigcup_i V_i)) = d\mu_f(A).$$

**Lemma II.** If  $A$  is a Borel set contained in a set  $P_{i,j}^{(n)}$ ,  $n \geq N$ :

$$\mu_f(A) = k_{i,j}^{(n)} d^{-n} \mu_f(f^n(A)).$$

*Proof.* Take a topological disk  $V \subset P_i$  with  $\mu_f(\partial V) = 0$ ,  $\mu_f(P_i \setminus V) = 0$  and not containing critical values of  $f^n$ . To obtain  $V$ , it is enough to join the critical values of  $f^n$  in  $P_i$  to its boundary with disjoint arcs having zero measure and the defining  $V$  as the complement in  $P_i$  of these curves. Define  $W = f^{-n}(V) \cap P_{i,j}^{(n)}$ . Since there are no critical values of  $f$  in  $V$ , there exist  $k_{i,j}^{(n)}$  branches of  $f^{-n}$  on  $V$ , i.e., analytic functions  $\phi_i: V \rightarrow W$ ,  $i = 1, \dots, k_{i,j}^{(n)}$ , such that  $\phi_i(f(z)) = z$  for all  $z \in W$  and  $\phi_i(w) \neq \phi_j(w)$  for all  $w \in V$ ,  $1 \leq i < j \leq k_{i,j}^{(n)}$ . It is clear that:

- $\phi_i(V) \cap \phi_j(V) = \emptyset$ ,  $1 \leq i < j \leq k_{i,j}^{(n)}$ .
- $f^n(\phi_i(V)) = V$ ,  $i = 1, \dots, k_{i,j}^{(n)}$ .
- $f^n/\phi_i(V)$  is injective,  $i = 1, \dots, k_{i,j}^{(n)}$ .
- $\bigcup_i \phi_i(V) = W$ .

By properties (1)-(4), we can apply Lemma I to obtain:

$$(5) \quad \mu_f(A \cap \phi_i(V)) = d^{-n} \mu_f(f^n(A) \cap V).$$

By (1), (4) and (5):

$$\mu_f(A \cap W) = \sum_i \mu_f(A \cap \phi_i(V)) = k_{i,j}^{(n)} d^{-n} \mu_f(f^n(A) \cap V) = k_{i,j}^{(n)} d^{-n} \mu_f(f(A)).$$

But  $\mu_f(A \cap W) = \mu_f(A)$  because:

$$\mu_f(A) - \mu_f(A \cap W) \leq \mu_f(P_{i,j}^{(n)} \setminus W) \leq \mu_f(f^{-n}(P_i \setminus V)) = \mu_f(P_i \setminus V) = 0.$$

Now set  $U_n = \bigcup_{i,j} P_{i,j}^{(n)}$ . By Lemma II:

$$\begin{aligned}\mu_f(U_n) &= \sum_{i,j} \mu_f(P_{i,j}^{(n)}) = d^{-n} \sum_{i,j} k_{i,j}^{(n)} \mu_f(P_i) = \\ &= d^{-n} \sum_i \left( \sum_j d_{i,j}^{(n)} \right) \mu_f(P_i) \geq d^{-n} \left( 1 - \frac{\varepsilon}{2} \right) d^n \sum_i \mu_f(P_i) = 1 - \frac{\varepsilon}{2}.\end{aligned}$$

For each  $n \geq N$ , the sets  $P_{i,j}^{(n)} \cap J(f)$ ,  $i = 1, \dots, 4$ ,  $j = 1, \dots, \ell(n, i)$  are a partition of  $U \cap J(f)$ . This partition can be extended to a partition  $P_n$  of  $J(f)$  in such a way that:

$$\lim_{n \rightarrow +\infty} (\sup_{P \in \mathcal{P}_n} \text{diam}(P)) = 0.$$

This property and standard derivation theorems imply that if  $\mathcal{P}_n(x)$  denotes the atom of  $\mathcal{P}_n$  containing  $x$ , then the sequence of functions  $F_n: J(f) \rightarrow \mathbb{R}$  defined by:

$$F_n(x) = \lim_{n \rightarrow +\infty} \frac{\mu_f(S \cap \mathcal{P}_n(x))}{\mu_f(\mathcal{P}_n(x))}$$

converges in measure to the characteristic function  $f_S$  of  $S$ . From this property, we shall prove the following claim: if we take Borel sets  $A_n$ ,  $n \geq 1$ , such that  $f^{-n}(A_n) = S$ , there exists an atom  $P_i$  and a sequence  $n_j \rightarrow +\infty$  such that  $\mu_f(P_i \setminus A_{n_j}) \rightarrow 0$  when  $j \rightarrow +\infty$ . This implies that  $\mu_f(S) = 1$  because there exists  $n > 0$  such that  $f^n(P_i) \supset J(f)$ , and by the corollary of Lemma I:

$$\begin{aligned}\lim_{j \rightarrow \infty} \mu_f(f^{n_j}(P_i \cap A_{n_j})) &= \lim_{j \rightarrow +\infty} \mu_f((f^{n_j}(P_i) - f^{n_j}(P_i \setminus A_{n_j}))) = \\ &= 1 - \lim_{j \rightarrow +\infty} \mu_f(f^{n_j}(P_i \setminus A_{n_j})) \geq 1 - d^n \lim_{j \rightarrow +\infty} \mu_f(P_i \setminus A_{n_j}) = 1.\end{aligned}$$

But:

$$f^n(P_i \cap A_{n_j}) \subset f^n(P_i) \cap f^n(A_{n_j}) = f^n(A_{n_j}) = A_{n_j+n}.$$



Then  $\lim_{j \rightarrow +\infty} \mu_f(A_{n_j+n}) = 1$  and:

$$\mu_f(S) = \mu_f(f^{-(n_j+n)}(A_{n_j+n})) = \mu_f(A_{n_j+n}).$$

Hence,  $\mu_f(S) = 1$ . To prove the claim observe that the convergence in measure of the sequence  $F_n$  to  $f_S$  implies that for every  $k > 0$  there exists  $n_k$  such that the set of points  $x$  satisfying:

$$(6) \quad \left| \frac{\mu_f(S \cap \mathcal{P}_{n_k}(x))}{\mu_f(\mathcal{P}_{n_k}(x))} - f_S(x) \right| \leq \frac{1}{k}$$

has measure  $\geq 1 - (\varepsilon/2)$ . Then this set intersects  $S \cap U_{n_k}$  because

$$\mu_f(S \cap U_{n_k}) = \mu_f((S^c \cup U_{n_k}^c)^c) \geq \left( (1 - 2\varepsilon) - \frac{\varepsilon}{2} \right) = \frac{3}{2}.$$

Let  $x_k$  be a point in the intersection. Since it belongs to  $U_{n_k}$ , we have  $\mathcal{P}_{n_k}(x_k) = P_{i,j}^{(n_k)}$  for some  $i$  and  $j$ . By (6):

$$(7) \quad \frac{\mu_f(S \cap P_{i,j}^{(n_k)})}{\mu_f(P_{i,j}^{(n_k)})} \geq 1 - \frac{1}{k}.$$

By Lemma II:

$$(8) \quad d^{-n_k} k_{i,j}^{(n_k)} \mu_f(f^{n_k}(S \cap P_{i,j}^{(n_k)})) = \mu_f(S \cap P_{i,j}^{(n_k)})$$

$$(9) \quad d^{-n_k} k_{i,j}^{(n_k)} \mu_f(f^{n_k}(P_{i,j}^{(n_k)})) = \mu_f(P_{i,j}^{(n_k)}).$$

But  $f^{n_k}(P_{i,j}^{(n_k)}) = P_i$ . Hence, (7), (8) and (9) imply:

$$\begin{aligned} 1 - \frac{1}{k} &\leq \frac{\mu_f(S \cap P_{i,j}^{(n_k)})}{\mu_f(P_{i,j}^{(n_k)})} = \frac{\mu_f(f^{n_k}(S \cap P_{i,j}^{(n_k)}))}{\mu_f(P_i)} \leq \\ &\leq \frac{\mu_f(f^{n_k}(S) \cap f^{n_k}(P_{i,j}^{(n_k)}))}{\mu_f(P_i)} = \\ &= \frac{\mu_f(A_{n_k} \cap P_i)}{\mu_f(P_i)}. \end{aligned}$$

Hence:

$$\begin{aligned} \mu_f(P_i \setminus A_{n_k}) &= \mu_f(P_i) - \mu_f(P_i \cap A_{n_k}) \leq \\ &\leq \frac{1}{k} \mu_f(P_i) \leq \frac{1}{k} \sup_i \mu_f(P_i) \end{aligned}$$

thus completing the proof of the claim.

To prove (d) take a family  $\{U_1, \dots, U_m\}$  of topological disks not containing critical values of  $f$  and such that:

$$(1) \quad \mu_f\left(\left(\bigcup_{i=1}^m U_i\right)^c\right) = 0.$$

Then for every  $1 \leq i \leq m$ , there exist  $d$  branches  $g_j^{(i)} = U_i \rightarrow \bar{\mathbb{C}}$ ,  $j = 1, \dots, d$  of  $f^{-1} \setminus U_i$ . Set  $U_j^{(i)} = g_j^{(i)}(U_i)$ . From (1) it follows that:

$$(2) \quad \mu_f\left(\left(\bigcup_{i=1}^m \bigcup_{j=1}^d U_j^{(i)}\right)^c\right) = 0.$$

Suppose  $A \subset \bar{\mathbb{C}}$  is a Borel set such that  $f/A$  is injective. It follows from the injectivity of  $f/A$  that the sets  $f(A \cap U_j^{(i)})$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq d$ , are disjoint. This property together with (1), (2) and Lemma I yield

$$\begin{aligned} \mu_f(f(A)) &= \mu_f\left(\bigcup_i (f(A) \cap U_i)\right) = \mu_f\left(\bigcup_{i,j} f(A \cap U_j^{(i)})\right) = \\ &= \sum_{i,j} \mu_f(f(A \cap U_j^{(i)})) = d \sum_{i,j} \mu_f(A \cap U_j^{(i)}) = \\ &= d \mu_f(A \cap \left(\bigcup_{i,j} U_j^{(i)}\right)) = d \mu_f(A). \end{aligned}$$

To prove that  $\mu_f$  is the unique  $f$ -invariant probability satisfying (d), consider another  $f$ -invariant probability  $\mu$  satisfying (d). We shall prove that  $\mu \ll \mu_f$ . Then the ergodicity of  $\mu_f$  implies  $\mu = \mu_f$ . To show  $\mu \ll \mu_f$ , we have to find for every  $\varepsilon > 0$ , a  $\delta > 0$  such that if  $K \subset J(f)$  is a compact set with  $\mu_f(K) \leq \delta$ , then  $\mu(K) \leq \varepsilon$ . Given  $\varepsilon > 0$  take  $N > 0$  such that there exists a family  $\mathcal{P} = \{P_1, \dots, P_r\}$  of  $(N, \varepsilon/2)$ -adapted sets such that

$$\begin{aligned} \mu_f\left(\left(\bigcup_i P_i\right)^c\right) &= 0 \\ \mu\left(\left(\bigcup_i P_i\right)^c\right) &= 0. \end{aligned}$$

This family is constructed as the family used in the proof of (c). Take  $\delta > 0$  satisfying:

$$(1) \quad \delta \frac{\mu(P_i)}{\mu_f(P_i)} \leq \frac{\varepsilon}{4}$$

for every  $1 \leq i \leq r$ . As in the proof of (c), denote  $P_{i,j}^{(n)}$  and  $k_{i,j}^{(n)}$   $j = 1, \dots, \ell(n, i)$  the sets and integers associated to  $P_i$  by Definition I. By Lemma II, we have:

$$(2) \quad \mu_f(P_{i,j}^{(n)}) = k_{i,j}^{(n)} d^{-n} \mu_f(P_i).$$



Since in the proof of Lemma II the only property of  $\mu_f$  used is precisely (d), we can apply Lemma II to  $\mu$  instead of  $\mu_f$ . Hence:

$$(3) \quad \mu(P_{i,j}^{(n)}) = k_{i,j}^{(n)} d^{-n} \mu(P_i).$$

From (3) and part (c) of Definition I, we obtain:

$$\begin{aligned} \mu\left(\bigcup_j P_{i,j}^{(n)}\right) &= d^{-n} \mu(P_i) \sum_j k_{i,j}^{(n)} \geq d^{-n} \mu(P_i) \left(1 - \frac{\varepsilon}{2}\right) d^n = \\ &= \left(1 - \frac{\varepsilon}{2}\right) \mu(P_i). \end{aligned}$$

The same argument, replacing (3) by (2), shows that:

$$\mu_f\left(\bigcup_j P_{i,j}^{(n)}\right) \geq \left(1 - \frac{\varepsilon}{2}\right) \mu_f(P_i).$$

Then

$$(4) \quad \mu\left(\bigcup_{i,j} P_{i,j}^{(n)}\right) \geq 1 - \frac{\varepsilon}{2},$$

$$(5) \quad \mu_f\left(\bigcup_{i,j} P_{i,j}^{(n)}\right) \geq 1 - \frac{\varepsilon}{2}.$$

Suppose that  $K \subset J(f)$  is a compact set with  $\mu_f(K) \leq \delta$ . We want to show that  $\mu(K) \leq \varepsilon$ . By (5):

$$\mu\left(K \cap \left(\bigcup_{i,j} P_{i,j}^{(n)}\right)^c\right) \leq \frac{\varepsilon}{2}.$$

It remains to prove:

$$(6) \quad \mu\left(K \cap \left(\bigcup_{i,j} P_{i,j}^{(n)}\right)\right) \leq \frac{\varepsilon}{2}.$$

Set

$$\mathcal{S}_n = \{(i,j) \mid P_{i,j}^{(n)} \cap K \neq \emptyset\}$$

$$K_n = \bigcup \{P_{i,j}^{(n)} \mid (i,j) \in \mathcal{S}_n\}$$

Since the diameters of the atoms of  $P_{i,j}^{(n)}$  converge to zero uniformly in  $(i,j)$  (by part (d) of Definition I) and by the compactness of  $K$ , it follows that

$$\lim_{n \rightarrow +\infty} \mu_f(K_n) = \mu_f(K),$$

$$(7) \quad \lim_{n \rightarrow +\infty} \mu(K_n) = \mu(K).$$

Then  $\mu_f(K_n) \leq 2\delta$  if  $n$  is large. It follows that:

$$2\delta \geq \mu_f(K_n) \geq \sum_{(i,j) \in \mathcal{S}_n} \mu_f(P_{i,j}^{(n)}).$$

By (2) and (3):

$$2\delta \geq \sum_{(i,j) \in \mathcal{S}_n} \mu_f(P_{i,j}^{(n)}) = \sum_{(i,j) \in \mathcal{S}_n} \frac{\mu_f(P_i)}{\mu(P_i)} \mu(P_{i,j}^{(n)}).$$

Hence:

$$2\delta \sup_i \frac{\mu(P_i)}{\mu_f(P_i)} \geq \sum_{(i,j) \in \mathcal{S}_n} \mu(P_{i,j}^{(n)}) = \mu(K_n).$$

By (1):

$$\frac{\varepsilon}{2} \geq 2\delta \sup_i \frac{\mu(P_i)}{\mu_f(P_i)} \geq \mu(K_n).$$

Then, by (7),  $\mu(K) \leq \varepsilon/2$ .

To prove (c), take a family  $\mathcal{P} = \{P_1, \dots, P_r\}$  of disjoint topological disks such that  $f/P_i$  is injective for all  $1 \leq i \leq r$  and  $\mu_f\left(\left(\bigcup_i P_i\right)^c\right) = 0$ .

Denote  $\mathcal{P}_n = \bigvee_{j=0}^n f^j(\mathcal{P})$  and let  $\mathcal{P}_n(x)$  be the atom of  $\mathcal{P}_n$  containing  $x$ .

Observe that  $f^n/\mathcal{P}_n(x)$  is injective for all  $n \geq 1$ . In fact, this property holds for  $n=0$  and if it is true for  $n=m$  then

$$f^{m+1}/\mathcal{P}_{m+1}(x) = (f/\mathcal{P}_{m+1}(x)) \circ (f^m/\mathcal{P}_m(f(x))).$$

But  $f/\mathcal{P}_{m+1}(x)$  is injective because  $\mathcal{P}_{m+1}(x) \subset \mathcal{P}(x)$  and  $f^m/\mathcal{P}_m(f(x))$  is injective by the induction hypothesis. Now we shall prove that:

$$(10) \quad \limsup_{n \rightarrow +\infty} \left( -\frac{1}{n} \log \mu_f(\mathcal{P}_n(x)) \right) \geq \log d$$

for  $\mu_f$ -a.e.  $x$ . We have:  $f^n(\mathcal{P}_n(x)) \subset \mathcal{P}(f^n(x))$ . Then

$$\mu_f(f^n(\mathcal{P}_n(x))) \leq \mu_f(\mathcal{P}(f^n(x))).$$

By (d) and by the injective of  $f^n/\mathcal{P}_n(x)$ :

$$\mu_f(\mathcal{P}_n(x)) = d^{-n} \mu_f(f^n(\mathcal{P}_n(x))) \leq d^{-n} \mu_f(\mathcal{P}(f^n(x))).$$

Then:

$$-\frac{1}{n} \log \mu_f(\mathcal{P}_n(x)) \geq \log d - \frac{1}{n} \log \mu_f(\mathcal{P}(f^n(x))).$$



From this (10) follows. But (10) implies that

$$h_{\mu_f}(f) \geq \log d.$$

On the other hand, Gromov proved in [3] that  $h_{top}(f) = \log d$ . Then  $h_{\mu_f}(f) \leq \log d$ .

## II. Proof of the Fundamental Lemma

If  $w \in \bar{\mathbb{C}}$  and  $n \in \mathbb{Z}^+$ , denote  $\widehat{m}_n(w)$  the multiplicity of  $w$  as root of the equation  $f^n(z) = f^n(w)$ . Set  $m_n(z) = \max \{ \widehat{m}_n(w) \mid w \in f^{-n}(z) \}$ . We shall need the following lemma:

**Lemma.** For every  $z \notin \text{Exc}(f)$  there exist  $N_1 > 0$  and  $1 < d_0 < d$  such that  $m_n(z) \leq d_0^n$  for all  $n \geq N_1$ .

*Proof.* Define  $\mathcal{B}(z)$  as the set of functions  $\theta : \mathbb{Z}^+ \rightarrow \bar{\mathbb{C}}$  such that  $\theta(0) = z$  and  $f(\theta(j+1)) = \theta(j)$  for all  $j \geq 0$ . Define  $\mathcal{B}_0(z)$  as the set of  $\theta \in \mathcal{B}(z)$  such that  $\theta(j)$  is a critical value of  $f$  only for a finite set of values of  $j$ . Then it is easy to see that  $\mathcal{B}(z) \neq \mathcal{B}_0(z)$  if and only if  $z$  belongs to the orbit of a periodic critical point, and that in this case, there is only one element  $\gamma \in \mathcal{B}(z) \setminus \mathcal{B}_0(z)$  that is periodic i.e., for some  $t$ , satisfies  $\gamma(t+j) = \gamma(j)$  for all  $j \geq 0$ . Now define

$$\alpha_n(z) = \{ \theta(n) \mid \theta \in \mathcal{B}(z) \},$$

$$\alpha_n^0(z) = \{ \theta(n) \mid \theta \in \mathcal{B}_0(z) \}.$$

From the fact that  $f$  has only finitely many critical points it follows that there exists  $N_2 > 0$  such that  $\theta(n)$  is not a critical point of  $f$  for all  $n \geq N_2$ ,  $\theta \in \mathcal{B}_0(z)$ . Hence

$$\widehat{m}_1(w) = 1$$

for all  $w \in \alpha_n^0(z)$  and  $n \geq N_2$ . Then, if  $w \in \alpha_n^0(z)$  with  $n \geq N_2$ , we obtain

$$\widehat{m}_n(w) = \prod_{j=0}^{n-1} \widehat{m}_1(f^j(w)) = \prod_{j=0}^{N_2-1} \widehat{m}_1(f^j(w)) \leq d^{N_2}.$$

If  $\mathcal{B}_0(z) = \mathcal{B}(z)$ , this concludes the proof because

$$m_n(z) = \max \{ \widehat{m}_n(w) \mid w \in \alpha_n(z) \} = \max \{ \widehat{m}_n(w) \mid w \in \alpha_n^0(z) \} \leq d^{N_2}$$

for all  $n \geq N_2$ . If  $\mathcal{B}_0(z) \neq \mathcal{B}(z)$  and if  $\gamma$  is the unique element in  $\mathcal{B}(z) \setminus \mathcal{B}_0(z)$ :

$$m_n(z) = \max \{ \widehat{m}_n(w) \mid w \in \alpha_n(z) \} \leq \max \{ d^{N_2}, \widehat{m}_n(\gamma(n)) \}.$$

Therefore the problem is reduced to show that there exist  $G > 0$  and  $1 < d_1 < d$  such that  $\widehat{m}_n(\gamma(n)) \leq Gd_1^n$ , for large values of  $n$ . But if  $n = kt + r$ ,  $k \in \mathbb{Z}^+$ ,  $r \in \mathbb{Z}^+$ ,  $0 \leq r < t$ , we have, using the periodicity of  $\gamma$ , that

$$\widehat{m}_n(\gamma(n)) = \prod_{j=0}^{n-1} \widehat{m}_1(\gamma(n-j)) = \prod_{j=0}^r \widehat{m}_1(\gamma(n-j)) \left( \prod_{j=0}^{t-1} \widehat{m}_1(\gamma(j)) \right)^k.$$

Set:

$$d_1 = \left( \prod_{j=0}^{t-1} \widehat{m}_1(\gamma(j)) \right)^{\frac{1}{t}}.$$

Then:

$$\widehat{m}_n(\gamma(n)) \leq d^r d_1^{kt} \leq d^t d_1^n.$$

for every  $n \geq t$ . This reduces our problem to prove that  $d_1 < d$ . But since every factor in the definition of  $d_1$  is  $\leq d$ , it is sufficient to show that we cannot have  $\widehat{m}_1(\gamma(j)) = d$  for all  $0 \leq j \leq t$ . But  $\widehat{m}_1(\gamma(j)) = d$  for all  $0 \leq j \leq t$  implies that  $f^{-1}(\pi) = \pi$ , where  $\pi$  is the periodic orbic  $\pi = \{ \gamma(0), \dots, \gamma(t-1) \}$ . Then  $f^{-1}(\pi^c) \subset \pi^c$ . Therefore  $(\bigcup_{n \geq 0} f^n(U)) \cap \pi = \emptyset$  for every subset  $U$  of  $\pi^c$ . But it is clear that  $\pi \cap J(f) = \emptyset$ . Hence, we can take  $U$  as an open set intersecting  $J(f)$ . Then, as we explained in the introduction,  $\bigcup_{n \geq 0} f^n(U) \supset \text{Exc}(f)^c$ . This, together with  $(\bigcup_{n \geq 0} f^n(U)) \cap \pi = \emptyset$  yields  $\pi \subset \text{Exc}(f)$ , contradicting the assumption  $z \notin \text{Exc}(f)$ .

Now let us prove the Fundamental Lemma. Take  $N_0$  so large that the sequence  $nd^{-n}$ ,  $n \geq 1$ , is decreasing for  $n \geq N_0$  and  $m_{N_0}(z)N_0d^{-N_0} \leq \varepsilon/2$ . Such  $N_0$  exists by the previous Lemma. Then

$$(*) \quad m_{N_0}(z)nd^{-n} \leq \frac{\varepsilon}{2}.$$

for every  $n \geq N_0$ . Assume that  $N_0$  is large enough to satisfy

$$(**) \quad 4m_{N_0}(z)d^{-N_0} \frac{d^2}{1-d} \leq \frac{\varepsilon}{2}.$$

Set  $m = m_{N_0}(z)$ . Since the only critical value of  $f^{N_0}$  contained in  $\gamma$  can be  $z$ , it follows that the connected components  $\gamma_1, \dots, \gamma_r$  of  $f^{-N_0}(\gamma)$  are either arcs or a union of arcs with a unique point of intersection. Therefore, each  $\gamma_i$  is simply connected. We can then take a topological disk  $U_0 \supset \gamma$ , so thin that there exist disjoint topological disks  $V_i \supset \gamma_i$ ,  $i = 1, \dots, r$  such that

$$f^{N_0}(V_i) = U_0$$



and  $f^{N_0}/V_i : V_i \rightarrow U_0$  is a  $k_i$ -to-1 map for all  $1 \leq i \leq r$ . Now set:

$$\begin{aligned} \varepsilon_{N_0} &= 0 \\ \varepsilon_{n+1} &= \varepsilon_n + 4md^{-na^2} \end{aligned}$$

for  $n \geq N_0$ . Observe that, by (\*\*):

$$\varepsilon_n \leq \varepsilon_{N_0} + 4md^2 \sum_{j=N_0}^{\infty} d^{-j} = 4md^{-N_0} \frac{d^3}{d-1} \leq \frac{\varepsilon}{2}.$$

We claim that for every  $n \geq N_0$ ,  $f^{-n}(U_0)$  contains a union of disjoint topological disks  $W_i^{(n)}$ ,  $i = 1, \dots, \tilde{\ell}_n$ , such that, for all  $i$ ,  $f^n(W_i^{(n)}) = U_0$  and  $f^n/W_i^{(n)} : W_i^{(n)} \rightarrow U_0$  is a  $k_i^{(n)}$ -to-1 map, where  $1 \leq k_i^{(n)} \leq m$ ,  $1 \leq i \leq \tilde{\ell}_n$  are integers satisfying

$$\sum_{i=1}^{\tilde{\ell}_n} k_i^{(n)} \geq \left(1 - \frac{1}{2} \varepsilon_n\right) d^n.$$

Clearly, the property is true for  $n = N_0$  just taking  $W_i^{(N_0)} = V_i$ ,  $i = 1, \dots, r$ ,  $k_i^{(N_0)} = k_i$  because

$$\sum_{i=1}^r k_i^{(N_0)} = \sum_{i=1}^r k_i = d^n.$$

The proof of the claim will now be completed by induction. Suppose constructed  $W_i^{(n)}$ ,  $k_i^{(n)}$ ,  $1 \leq i \leq \tilde{\ell}_n$ . Let  $H$  be the set of integers  $t$  between 1 and  $\tilde{\ell}_n$  such that  $W_i^{(n)}$  doesn't contain critical values of  $f$ . For every  $t \in H$  there exist disjoint topological disks  $D_j^{(t)}$ ,  $1 \leq j \leq d$ , such that  $f$  maps  $D_j^{(t)}$  homeomorphically onto  $W_t^{(n)}$ . Define as  $W_i^{(n+1)}$ ,  $i = 1, \dots, \tilde{\ell}_{n+1}$ , the sets  $D_j^{(t)}$ ,  $t \in H$ ,  $1 \leq j \leq d$ , and, if  $W_i^{(n+1)} = D_j^{(t)}$ , set  $k_i^{(n+1)} = k_t^{(n)}$ . Then:

$$\begin{aligned} \sum_{i=1}^{\tilde{\ell}_{n+1}} k_i^{(n+1)} &= d \sum_{i \in H} k_i^{(n)} = d \sum_{i=1}^{\tilde{\ell}_n} k_i^{(n)} - d \sum_{i \notin H} k_i^{(n)} \geq \\ &\geq d \sum_{i=1}^{\tilde{\ell}_n} k_i^{(n)} - d(\tilde{\ell}_n - \# H)m \geq \left(1 - \frac{1}{2} \varepsilon_n\right) d^{n+1} - d(\tilde{\ell}_n - \# H)m. \end{aligned}$$

But  $\tilde{\ell}_n - \# H$  is bounded by the number of critical values of  $f$ , that is,  $2d - 2$ . Hence:

$$\begin{aligned} \sum_{i=1}^{\tilde{\ell}_{n+1}} k_i^{(n+1)} &\geq \left(1 - \frac{\varepsilon_n}{2}\right) d^{n+1} - 2d^2m = \\ &= \left(1 - \frac{1}{2}(\varepsilon_n + 4d^2md^{-(n+1)})\right) d^{n+1} = \\ &= \left(1 - \frac{1}{2} \varepsilon_{n+1}\right) d^{n+1}. \end{aligned}$$

This completes the proof of the claim. The next step is to restrict, for each  $n$ , the family  $W_i^{(n)}$ ,  $1 \leq i \leq \tilde{\ell}_n$ , to those values of  $i$  that satisfy:

$$\lambda(W_i^{(n)}) \leq \frac{1}{n},$$

where  $\lambda$  denotes the Lebesgue measure.

Suppose that those values of  $i$  are  $1, \dots, \ell_n$ . Then:

$$\sum_{i=1}^{\ell_n} k_i^{(n)} \geq \sum_{i=1}^{\tilde{\ell}_n} k_i^{(n)} - (\ell_n - \tilde{\ell}_n)m.$$

To bound  $(\ell_n - \tilde{\ell}_n)$ , observe that

$$1 \geq \lambda\left(\bigcup_{i>\ell_n} W_i^{(n)}\right) \geq (\tilde{\ell}_n - \ell_n) \frac{1}{n}.$$

Then:

$$(1) \quad \sum_{i=1}^{\ell_n} k_i^{(n)} \geq \left(1 - \frac{1}{2} \varepsilon\right) d^n - nm = \left(1 - \frac{1}{2}(\varepsilon + 2mnd^{-n})\right) d^n.$$

By (\*), the factor of  $d^n$  is bounded by  $(1 - \varepsilon)$ . Finally, to complete the proof of the Lemma, we shall prove that for any topological disk  $U$  whose closure is contained in  $U_0$ , we have:

$$(2) \quad \lim_{n \rightarrow +\infty} (\sup_i \text{diam}(f^{-n}(U) \cap W_i^{(n)})) = 0.$$

If this property is true, the lemma is proved taking  $U$  containing  $\gamma$  and with closure contained in  $U_0$ . Then we define:

$$S_i^{(n)} = f^{-n}(U) \cap W_i^{(n)}.$$

By (2), there exists  $N \geq N_0$  such that property (b) of Definition I is satisfied. Property (d) also follows from (2). Finally,  $f^n/S_i^{(n)}$  is a  $k_i^{(n)}$ -to-1 map onto  $U$ , and from (1) and (\*) it follows that the integers  $k_i^{(n)}$  satisfy property (c). It remains to prove (2). By the way the sets  $W_i^{(n)}$  were constructed, we know that  $f^{n-N_0}/W_i^{(n)}$  is a conformal representation onto some  $V_j$ . Let  $\phi_i^{(n)} : V_j \rightarrow W_i^{(n)}$  be its inverse. Set  $D_r = \{z \mid |z| < r\}$ . Let  $\alpha_j : D_1 \rightarrow V_j$  be a conformal representation. Define  $\psi_i^{(n)} : D_1 \rightarrow W_i^{(n)}$  as  $\psi_i^{(n)} = \phi_i^{(n)} \alpha_j$ . Instead of (2), we shall prove

$$(3) \quad \lim_{n \rightarrow +\infty} (\sup_i \text{diam}(\psi_i^{(n)}(D_r))) = 0$$

for all  $0 < r < 1$ . This implies (2) because  $\psi_i^{(n)}(D_r) \supset f^{-(n-N_0)}(U) \cap W_i^{(n)}$  for all  $n \geq N$ , if  $r$  is near enough to 1. To prove (3), recall that the Distortion Theorem for univalent functions states that for all  $0 < r < 1$



there exists  $K(r)$  such that every univalent function  $\phi : D_1 \rightarrow \mathbb{C}$  satisfies  $|\phi'(a)/\phi'(b)| \leq K(r)$  for all  $a$  and  $b$  in  $D_r$ . In particular, if  $\lambda(\cdot)$  denotes Lebesgue measure,  $\lambda(\phi(D_r)) \geq K(r)^{-1} |\phi'(a)|^2 \lambda(D_r)$ , for all  $a \in D_r$ . In our case:

$$\frac{1}{n} \geq \lambda(W_i^{(n)}) \geq \lambda(\psi_i^{(n)}(D_r)) \geq K(r)^{-1} \lambda(D_r) |(\psi_i^{(n)})'(z)|^2$$

for all  $0 < r < 1$ ,  $z \in D_r$ . Then:

$$\lim_{n \rightarrow +\infty} \left( \sup_{i, z \in D_r} |(\psi_i^{(n)})'(z)| \right) = 0$$

for all  $0 < r < 1$ , and this obviously implies (3).

### References

- [1] H. Brolin, *Invariant sets under iteration of rational functions*, Arkiv für Mathematik, Band G, nr. 6 (1966).
- [2] M. Fatou, *Sur les fonctions satisfaisant certes equations fonctionelles*, Bull. de la Soc. Math. de France 47-48 (1919-1920).
- [3] M. Gromov, *On the entropy of holomorphic maps*, Preprint I.H.E.S.
- [4] G. Julia, *Sur l'iteration des fonctions rationnelles*, Journal de Mathematique Pure et Appliquée, 8 (1918).

Artur Lopes  
 Instituto de Matemática  
 Universidade Federal de Rio Grande do Sul  
 Porto Alegre, Brazil

Alexandre Freire  
 Instituto de Matemática Pura e Aplicada  
 Rio de Janeiro, Brazil

Ricardo Mañé  
 Instituto de Matemática Pura e Aplicada  
 Rio de Janeiro, Brazil