KARI ASTALA
ZOLTAN BALOGH
HANS MARTIN REIMANN

Lempert mappings and holomorphic motions in C^n

Astérisque, tome 261 (2000), p. 1-12
<http://www.numdam.org/item?id=AST_2000__261__1_0>

Astérisque
261, 2000, p. 1–12

LEMPERT MAPPINGS AND HOLOMORPHIC MOTIONS IN $\mathbf{C}^n$

by

Kari Astala, Zoltan Balogh & Hans Martin Reimann

Abstract. — The purpose of this note is twofold: to discuss the concept of holomorphic motions and phenomena of Mañé-Sad-Sullivan type in several complex variables and secondly, to compare the different notions of Beltrami differentials in CR-geometry which have appeared in [4] and [7].

1. Introduction

Holomorphic motions in the complex plane $\mathbf{C}$ are isotopies of subsets $A \subset \mathbf{C}$ for which the dependence on the "time" parameter is holomorphic. This simple notion has been important in explaining a number of different questions in complex analysis, in particular the rigidity phenomena in complex dynamics and the role of quasiconformal mappings in holomorphic deformations.

It was Mañé, Sad and Sullivan **[9]** who first realized that for time-holomorphic isotopies one can forget all smoothness requirements in space variables and thus produce almost automatic rigidity results in various contexts. Given a subset $A \subset \overline{\mathbf{C}}$, it is simply enough to define a holomorphic motion of A as a mapping $f : \Delta \times A \to \overline{\mathbf{C}}$, where $\Delta = \{\lambda \in \mathbf{C} : |\lambda| < 1\}$, such that

(i) for any fixed $a \in A$, the map $\lambda \to f(\lambda, a)$ is holomorphic in Δ
(ii) for any fixed $\lambda \in \Delta$, the map $a \to f(\lambda, a) = f_\lambda(a)$ is an injection and
(iii) the mapping f_0 is the identity on A.

Then f is automatically continuous in $\overline{A} \times \overline{\mathbf{C}}$ and the restrictions $f_\lambda(.)$ are quasisymmetric mappings **[9]**; in case $A = \overline{\mathbf{C}}$ they are quasiconformal with the precise

1991 ***Mathematics Subject Classification.*** — 30C65, 32F25, 32A30.
Key words and phrases. — Holomorphic motions, Lempert mappings, quasiconformal mappings on contact structures.

bound on the dilatation

$$K(f_\lambda) \leq \frac{1+|\lambda|}{1-|\lambda|}. \tag{1}$$

The picture was then completed by Slodkowski **[12]** who proved the so called Generalized Λ-lemma, that a holomorphic motion of any set $A \subset \overline{\mathbf{C}}$ extends to a motion of the whole $\overline{\mathbf{C}}$.

In this setting it is natural to look for similar phenomena in several complex variables, when the sets moving are of higher dimension. However, one quickly sees that the simple minded generalization does not work: Let, for example, $S(x) = x/|x|$ for $x \in \mathbf{R} \setminus \{0\}$, $S(0) = 0$ and define

$$f_\lambda(z, w) = (z + \lambda S(\operatorname{Re}\{w\}), w).$$

Then f_λ is holomorphic in λ and injective but not even continuous in $\mathbf{C}^2$.

Our first goal in this note is to introduce the proper notion or point of view to holomorphic motions in several complex variables and then show the existence of the first nontrivial examples, results of Mañé-Sad-Sullivan type. We expect that similar phenomena occur, in fact, in much larger setups.

Remark. — The generalizations to the case where the *parameter* space is higher dimensional were studied by Adrien Douady in his work **[3]**.

If there are to be holomorphic motions in $\mathbf{C}^n$, the one-dimensional theory suggests that they are connected to a notion of quasiconformality. Therefore recall that in several complex variables the appropriate concepts are the quasiconformal mappings on CR-structures **[4]**, or mappings on boundaries of pseudoconvex domains which firstly are contact transforms, *i.e.* preserve the horizontal (complex) lines of the tangent spaces

$$H_p\partial D = T_p\partial D \cap JT_p\partial D$$

where J is the complex structure as a mapping of $T_p\mathbf{C}^2$, and secondly, are there quasiconformal with respect to the corresponding Levi Form, *i.e.*

$$K(p) = \frac{\sup\{L(F_*X, F_*X) : X \in H_p\partial D,\ L(X,X) = 1\}}{\inf\{L(F_*X, F_*X) : X \in H_p\partial D,\ L(X,X) = 1\}} \leq K \tag{2}$$

for all $p \in \partial D$.

The same direction is, actually, suggested also by the approach of Slodkowski **[12]**. He viewed holomorphic motions (or their graphs) as disjoint analytic disks in $\mathbf{C}^2$. Namely given such a motion $f : \Delta \times A \to \mathbf{C}$ each point $a \in A$ defines a holomorphic disk $D_a \subset \mathbf{C}^2$, a holomorphic image of Δ, by

$$D_a = \{(\lambda, f(\lambda, a)) : \lambda \in \Delta\} \tag{3}$$

and these disks are clearly pointwise disjoint. Conversely, given a family of analytic disks of the form (3) with $D_a \cap D_b = \varnothing$ when $a \neq b$, they define a holomorpic motion $\Psi(\lambda, f(0,a)) = f(\lambda, a)$. (The extension of a given motion was then obtained by studying certain totally real tori whose polynomial hulls were shown to consists of disjoint families of suitable analytic disks.)

Interpreting the Mañé-Sad-Sullivan result in the language (3) of disjoint analytic disks, for a motion of the whole complex plane the disks D_a, $a \in \mathbf{C}$, fill in the domain $\Delta \times \mathbf{C}$. And when we move along the disks with λ, in the transverse direction *i.e.* on the *complex* lines of the corresponding tangent spaces of $\partial\Delta(|\lambda|) \times \mathbf{C}$ the mappings $(0,a) \mapsto (\lambda, f(\lambda, a))$ are now quasiconformal by the original λ-lemma.

This picture makes it very suggestive that similar phenomena should occur in other situations in $\mathbf{C}^2$ or $\mathbf{C}^n$ as well. That is, for suitable families of analytic disks one should expect that moving holomorphically along these disks yields automatic quasiconformality in transverse horizontal directions, quasiconformality in the Koranyi-Reimann sense, with the bound (1) on the dilatation. The philosophy of holomorphic motions in $\mathbf{C}^n$ would then be not that there is one strict definition of these motions but rather that there are several natural situations that share the common features described here.

To show that there do exist nontrivial holomorphic motions in the above sense in $\mathbf{C}^2$ (the choice $n = 2$ is made for simplicity) we make use of the theory developed by Lempert **[5]**-**[8]** and consider bounded strictly $\mathbf{R}$-convex smooth subdomains $D \subset \mathbf{C}^2$ and their generalizations the strictly linearly convex domains. The latter class consists of smooth bounded domains with the property that for each boundary point $p \in \partial D$ the horizontal space $H_p\partial D$ does not intersect $\bar{D}\backslash\{p\}$ and that $H_p\partial D$ has precisely first order contact with ∂D at p. That is, there exists $c > 0$ such that

$$\operatorname{dist}(q, H_p\partial D) \geq c \cdot \operatorname{dist}(p,q)^2, \qquad q \in D.$$

In particular, strictly convex domains are strictly linearly convex which in turn are strictly pseudoconvex.

As shown by Lempert in strictly linearly convex domains extremal Kobayashi disks are especially well behaved. For this recall that in any bounded domain containing the origin the Kobayashi indicatrix $\bar{I}$ of D is defined by

$$\bar{I} = \{f'(0) : \ f : \Delta \to D \text{ is holomorphic and } f(0) = 0\}.$$

If $v \in \partial\bar{I}$, a holomorphic mapping $f = f_v : \Delta \to D$ such that $f(0) = 0$ and $f'(0) = v$ is then called an extremal map corresponding to the vector v. In strictly linearly convex domains extremal disks are uniquely determined by v, a fact no longer true for general pseudoconvex domains. This enables us to simply define

$$\Psi : \Delta \times \partial\bar{I} \to \mathbf{C}^2, \qquad \Psi(\lambda, v) = \frac{f_v(\lambda)}{\lambda}$$

and we can describe a full counterpart of the Mañé-Sad-Sullivan result, a holomorphic motion of $\partial\bar{I}$.

Theorem 1. — *Let D be a strictly linearly convex domain containing the origin and $\Psi : \Delta \times \partial \bar{I} \to \mathbf{C}^2$ be defined by $\Psi(\lambda, v) = \lambda^{-1} f_v(\lambda)$. Then Ψ satisfies the following properties:*

(1) $\Psi(0, \cdot) = \mathrm{Id}\,|_{\partial \bar{I}}$;
(2) $\Psi(\cdot, v) : \Delta \to \mathbf{C}^2$ *is holomorphic;*
(3) $\Psi(\lambda, \cdot) : \partial \bar{I} \to A_\lambda$ *is a contact mapping where* $A_\lambda = \Psi(\lambda, \partial \bar{I})$ *is the boundary of a strictly pseudoconvex domain. In particular,* Ψ *is continuous in* $\Delta \times \partial \bar{I}$;
(4) $\Psi(\lambda, \cdot) : \partial \bar{I} \to A_\lambda$ *is* $K(\lambda)$*-quasiconformal with* $K(\lambda) \leq \frac{1+|\lambda|}{1-|\lambda|}$, $\lambda \in \Delta$.

We should mention that statement (4) is the new result proven here; statements (1), (2) and (3), due to Lempert, are being included for the sake of completeness. To obtain the optimal dilatation bound we turn to our second goal, to compare the different notions of Beltrami differentials in contact geometry and CR-manifolds, introduced respectively, by Koranyi and Reimann **[4]** and Lempert **[7]**. This with required preliminary material will be presented in the next section.

2. Inner actions and Beltrami differentials

It will be convenient start with a version of the Riemann mapping theorem in $\mathbf{C}^n$ due to Lempert (**[5]**, **[6]**), and use the formalism introduced by Semmes **[11]**. These Riemann mappings preserve the complex structure to some extent but are flexible enough to yield general existence results. In more precise terms, Lempert considered mappings $\rho : \bar{B} \to \bar{D}$ from the unit ball in $\mathbf{C}^n$ onto domains $D \subset \mathbf{C}^n$ containing the origin which satisfy the following three requirements:

(1) $\rho : \bar{B} \backslash \{0\} \to \bar{D} \backslash \{0\}$ is a smooth diffeomorphism and $\rho : \bar{B} \to \bar{D}$ is bilipshitz;
(2) ρ restricted to any complex line through the origin is holomorphic;
(3) ρ restricted to the boundary of any ball B_r centered at the origin and of radius $0 < r \leq 1$ is contact, *i.e.*

$$\rho_* H\partial B_r = H\partial D_r \qquad\qquad (D_r = \rho B_r).$$

For the last condition recall that when a domain is strictly pseudoconvex the horizontal tangent bundle $H\partial D = T\partial D \cap JT\partial D$, where J is the complex structure, defines a contact structure on the boundary.

In what follows a mapping with the above properties (1)-(3) will be called a Lempert mapping. The basic existence result is then:

***Theorem A* (Lempert).** — *Let D be a strictly linearly convex domain. Then there exists a Lempert mapping $\rho : \bar{B} \to \bar{D}$.*

A very nice exposition of the properties of the Lempert mappings was given by Semmes in **[11]**. The statement and proof of Theorem A, for example, may be found in **[11]** in the case of strictly convex domains and it is based essentially on the results in **[5]**. The proof of Theorem A in the case of strictly linearly convex domains follows exactly as in the strictly convex case (Theorem 5.2 in **[11]**) using the results in **[6]**.

It is important for our purposes to estimate the Beltrami differential of the Lempert mapping $\rho : \bar{B} \to \bar{D}$ in the sense [4], as a quasiconformal mapping on the level surfaces $\rho : \partial B_r \to \partial D_r$. Therefore, let first $r = 1$ and use the notation

$$\partial D = M$$

for the boundary of the stricly linearly convex domain $D \subset \mathbf{C}^2$. Then on the unit sphere

$$S^3 = \{(z,w) : |z|^2 + |w|^2 = 1\}$$

we consider the $(1,0)$-vectorfield $Z = \bar{z}\dfrac{\partial}{\partial w} - \bar{w}\dfrac{\partial}{\partial z}$. It is easy to check that Z is a section of $H^{1,0}S^3$ where

$$\mathbf{C} \otimes HS^3 = H^{1,0}S^3 \oplus H^{0,1}S^3$$

is the complexified horizontal bundle of the sphere and $H_p^{1,0}$ is the space of $(1,0)$-vectors, *i.e.* vectors of the form $Y - iJY$, $Y \in H_pM$. Since ρ is a contact mapping we can write

$$\rho_* Z = Y + \bar{\nu}\bar{Y} \text{ where } Y \in H^{1,0}_{\rho(z,w)}M. \tag{4}$$

We have used here the fact that $H^{1,0}_{\rho(z,w)}M$ is complex one-dimensional and therefore Y in (4) is uniquely determined.

In fact, in an invariant formulation one should think of ν in (4) not as a number but rather a complex-antilinear map $\nu : H_p^{1,0}M \to H_p^{1,0}M$ at each point $p \in M$ so that $\rho_* Z = Y + \overline{\nu(Y)}$. Once we have chosen a basis in $H_p^{1,0}S^3$ (or Z in (4)) ν becomes a complex number.

To study the variation of ν on M let us fix a point $(z,w) \in S^3$. Then $\zeta(z,w) \in S^3$ for $\zeta \in S^1$ and (4) gives us a function $\nu : S^1 \to \mathbf{C}$, $\nu = \nu(\zeta) = \nu_p(\zeta)$, $p = \rho(z,w)$. Moreover using property 3 of Lempert mappings we can do this consideration for any $\zeta \in \bar{\Delta} \setminus 0$ and we obtain a function ν on $\bar{\Delta}\setminus 0$.

The following simple statement (see also **[4]**) will be useful later in our note:

Lemma 2. — *Let us consider the Lempert mapping $\rho : \bar{B} \to \bar{D}$ of a strictly linearly convex domain D and the Beltrami differential ν as constructed above. Then we have the inequality: $|\nu| < 1$.*

Proof. — Let us denote by $u_0(z,w) = |z|^2 + |w|^2 - 1$ the defining function of B and $u_1(p) = u_0(\rho^{-1}(p))$, $p \in \bar{D}$. Then u_1 is a defining function of D that is smooth on $\bar{D}\setminus\{0\}$ (see **[11]**, lemma 2.9). Moreover if we consider the $(1,1)$-forms $\Omega_0 = \frac{1}{2}\partial\bar{\partial}u_0$ and $\Omega_1 = \frac{1}{2}\partial\bar{\partial}u_1$ by Lemma 2.6 in **[11]** we have

$$\Omega_0 = \rho^*\Omega_1. \tag{5}$$

We insert in the forms Ω_0 and Ω_1 vectors from the complexified horizontal spaces $\mathbf{C} \otimes HS^3$ and, respectively, $\mathbf{C} \otimes HM$. Relation (5) gives

$$\Omega_0(Z, \bar{Z}) = \Omega_1(Y + \bar{\nu}\bar{Y}, \bar{Y} + \nu Y). \tag{6}$$

Since $\Omega_1(Y, Y) = \Omega_1(\bar{Y}, \bar{Y}) = 0$ we have

$$0 < \Omega_0(Z, \bar{Z}) = (1 - |\nu|^2)\Omega_1(Y, \bar{Y}). \tag{7}$$

By strict pseudoconvexity of D we have $\Omega_1(Y, \bar{Y}) > 0$ and hence $|\nu| < 1$. $\square$

Furthermore, note that like in the classical case of planar quasiconformal mappings there is a relationship between the norm of the Beltrami coefficient and the quasiconformal dilatation of (2), given by

$$K(\rho) := \sup_{p \in M} K(p) = \frac{1 + ||\nu||_\infty}{1 - ||\nu||_\infty}, \tag{8}$$

see [**4**].

In his work [**7**] on embeddability of CR-structures Lempert studied inner actions on the boundary M of the strictly linearly convex domain D and defined for them another, apparently quite different Beltrami differential. Our next goal is to find a relationship between these two notions.

We begin by basic notation. Let $\{g_t\}, t \in \mathbf{R}$, be a smooth $\mathbf{R}$-action on M. It is called transverse if its infinitesimal generator $\{dg_t/dt\}_{t=0}$ is everywhere transverse to the contact plane field HM and a contact action $\{g_t\}$ in case the tangent maps g_{t*} preserve the contact distribution HM for all t. A consistent orientation of M is obtained by declaring the frame $X, JX, [JX, X]$ to be positive for a nonvanishing local section of HM. Then a transverse contact action is called positive if for one (or any) local section X of HM the frame X, JX and $\{dg_t/dt\}_{t=0}$ is positively oriented.

Let us now consider a positive contact action $\{g_\zeta\}, \zeta \in S^1$ of the unit circle. In general this action does not respect the complex structure J and one can measure this fact as follows. Let $p \in M$ and $X \in \mathbf{C} \otimes H_pM$ be such that X and $\bar{X}$ are linearly independent. Since the complex dimension of $\mathbf{C} \otimes H_pM$ is 2 this means that X and $\bar{X}$ span $\mathbf{C} \otimes H_pM$, so that there are complex numbers (not unique of course) a, b not both zero such that

$$a\bar{X} + bX \in H_p^{0,1}M. \tag{9}$$

Clearly $|a| \neq |b|$ for otherwise the $(0,1)$-vector in (9) and its conjugate would be dependent. We shall assume that X is chosen so that $|a| > |b|$ holds. In this case we say that X is a $(1,0)$-like vector.

If we fix $p \in M$ and X then the images $g_{\zeta*}X$ and $g_{\zeta*}\bar{X}$ span $\mathbf{C} \otimes H_{g_\zeta(p)}M$ as $g_{\zeta*}$ preserves the contact bundle HM. Therefore there are again complex numbers $a(\zeta)$,

$b(\zeta)$ not both zero, such that

$$a(\zeta)g_{\zeta*}\bar{X} + b(\zeta)g_{\zeta*}X \in H^{0,1}_{g_\zeta(p)}M. \tag{10}$$

Although $a(\zeta)$ and $b(\zeta)$ are not unique their quotient $\mu(\zeta) = b(\zeta)/a(\zeta) \in \mathbf{C} \cup \{\infty\}$ is uniquely determined and depends smoothly on ζ. Since $|\mu(\zeta)| \neq 1$ as before and $\mu(1) = b/a$ it follows that $|\mu(\zeta)| < 1$ for all $\zeta \in S^1$.

In this setting, Lempert calls μ the Beltrami differential associated to the trajectory of p. In fact μ depends on p itself and X but this dependence is just up to Moebius transformations of the unit disc ([**7**], proposition 3.1).

A positive contact action is now called *inner* if its Lempert-Beltrami differential μ, for each trajectory, extends to a smooth function on the closed unit disc $\bar{\Delta}$ in such a way that this extension is holomorphic in Δ.

The existence of the inner actions on boundaries of strictly linearly convex domains was proved by Lempert in [**7**]. We present here a different approach (in Theorem 3) that gives us the possibility to draw some precise consequences in the proof of Theorem 1. In fact, we are going to consider the action $g_\zeta : M \to M$ given by $g_\zeta(p) = \rho(\zeta(\rho^{-1}(p)))$ for $p \in M$, $\zeta \in S^1$. In other words, $g_\zeta : M \to M$ is the conjugate of the standard action $h_\zeta : S^3 \to S^3$, $h_\zeta(z,w) = \zeta(z,w)$ from the sphere S^3 to M by the Lempert mapping ρ.

It is in this setup that we can observe a connection between the two Beltrami differentials μ and ν. The observation is based on the fact that since the definition of μ depends on the choice of $p \in M$ and a vector $X \in \mathbf{C} \otimes H_pM$ one must do this selection carefully. After fixing $(z,w) \in S^3$ and $p = \rho(z,w) \in M$ (and hence also the function $\nu_p(\zeta)$) a consistent way is to take $X = \rho_* Z_{(z,w)} = Y + \bar{\nu}\bar{Y}$. We are allowed to do that since $|\nu| < 1$ by Lemma 2 and therefore X and $\bar{X}$ are independent. With these particular choices we then consider the Beltrami differential μ of g_ζ as given by (9) and (10).

Theorem 3. — *Let $D \subseteq \mathbf{C}^2$ be a strictly linearly convex domain with boundary M and let $\rho : \bar{B} \to \bar{D}$ be the Lempert mapping whose Beltrami coefficient, as a contact quasiconformal transformation, is ν. Then the action $g_\zeta : M \to M$ given by $g_\zeta(q) = \rho(\zeta(\rho^{-1}(q)))$ is an inner action. If the Lempert-Beltrami differential μ of this action g_ζ is defined as above we have then the relation*

$$\nu(\zeta) = -\Big(\frac{\zeta}{\bar{\zeta}}\Big)^2 \mu(\zeta) \quad \textit{for} \quad \zeta \in \bar{\Delta}\backslash\{0\}. \tag{11}$$

Proof. — We begin by proving the second statement for $\zeta \in S^1$. If we denote by $h_\zeta : S^3 \to S^3$ the standard action $h_\zeta(z,w) = \zeta \cdot (z,w)$ then it is easy to see that

$$h_{\zeta*}Z = \frac{\zeta}{\bar{\zeta}}Z, \ \zeta \in \bar{\Delta}. \tag{12}$$

Fix a point $(z,w) \in S^3$ and $p = \rho(z,w)$. For the value $\zeta = 1$ we have that $X = Y + \bar{\nu}\bar{Y}$ and $\bar{X} = \bar{Y} + \nu Y$ are linearly independent and there are numbers $a, b \in \mathbf{C}$ such that $a\bar{X} + bX \in H_p^{0,1}M$ or $(a\nu + b)Y + (a + b\bar{\nu})\bar{Y} \in H_p^{0,1}M$. Therefore $a\nu + b = 0$ and thus $\nu(1) = -b/a = -\mu(1)$. Furthermore $g_{\zeta *}X(p)$ and $g_{\zeta *}\bar{X}(p)$ are again independent and there are numbers $a(\zeta), b(\zeta) \in \mathbf{C}$ such that

$$a(\zeta)g_{\zeta *}\bar{X}(p) + b(\zeta)g_{\zeta *}X(p) \in H^{0,1}_{g_\zeta(p)}, \tag{13}$$

i.e. $\mu(\zeta) = b(\zeta)/a(\zeta)$.

On the other hand using (12)

$$\begin{aligned} g_{\zeta *}X(p) &= (\rho_* \circ h_{\zeta *} \circ \rho_*^{-1})X(p) = (\rho_* \circ h_{\zeta *})Z(z,w) = \rho_*\left(\frac{\zeta}{\bar{\zeta}}Z(\zeta(z,w))\right) \\ &= \frac{\zeta}{\bar{\zeta}}X(\rho(\zeta(z,w))) = \frac{\zeta}{\bar{\zeta}}(Y + \bar{\nu}(\zeta)\bar{Y}). \end{aligned}$$

Consequently (13) becomes

$$\left(a(\zeta)\frac{\bar{\zeta}}{\zeta}\nu(\zeta) + b(\zeta)\frac{\zeta}{\bar{\zeta}}\right)Y + \left(a(\zeta)\frac{\bar{\zeta}}{\zeta} + b(\zeta)\frac{\zeta}{\bar{\zeta}}\bar{\nu}(\zeta)\right)\bar{Y} \in H^{0,1}_{g_\zeta(p)}M, \tag{14}$$

which gives

$$a(\zeta)\frac{\bar{\zeta}}{\zeta}\nu(\zeta) + b(\zeta)\frac{\zeta}{\bar{\zeta}} = 0, \tag{15}$$

and therefore (11) follows for the points ζ on the unit circle S^1.

For the remaining part, the action $g_\zeta : M \to M$ was given by

$$g_\zeta = \rho \circ h_\zeta \circ \rho^{-1}, \ \zeta \in S^1. \tag{16}$$

It is clear that g_ζ is a contact action because ρ is a contact mapping by property 3 and h_ζ is a contact action on S^3. Furthermore g_ζ is transverse since $\rho|_{S^3} : S^3 \to M$ is a diffeomorphism and h_ζ is transverse on S^3. For the the first statement it therefore suffices to show that $\mu(\zeta)$ extends holomorphically to Δ. This follows along the same lines as in **[7]** but we present the argument for the convenience of the reader.

Let us recall that $\mu(\zeta)$ is given by

$$g_{\zeta *}\bar{X} + \mu(\zeta)g_{\zeta *}X \in H^{0,1}_{g_\zeta(p)}M, \ \zeta \in S^1 \tag{17}$$

where $X = Y + \bar{\nu}\bar{Y} \in \mathbf{C} \otimes H_pM$. Now $X = Y_1 + iY_2$ for some $Y_1, Y_2 \in H_pM$ where Y_1 and Y_2 are independent over $\mathbf{R}$.

Applying the projection map $\pi^{1,0} : \mathbf{C} \otimes H_{g_\zeta(p)}M \to H^{1,0}_{g_\zeta(p)}M$ to (17) we obtain

$$\pi^{1,0}g_{\zeta *}Y_1 - i\pi^{1,0}g_{\zeta *}Y_2 + \mu(\zeta)[\pi^{1,0}g_{\zeta *}Y_1 + i\pi^{1,0}g_{\zeta *}Y_2] = 0, \tag{18}$$

and thus (18) determines $\mu(\zeta)$ for $\zeta \in S^1$. To obtain a holomorphic extension we extend (18) to $\zeta \in \Delta$. Namely, we denote by $\xi(\zeta) = \pi^{1,0}g_{\zeta *}(Y_1)$ and $\eta(\zeta) = \pi^{1,0}g_{\zeta *}(Y_2)$ and we see as in **[7]** that ξ and η are holomorphic sections of the pullback bundle

$g_\zeta^* H^{1,0}M$. Our purpose is to show that (18) has then a solution for $\mu(\zeta)$ for $\zeta \in \Delta$ and the function $\mu : \Delta \to \mathbf{C}$ is holomorphic.

We can rewrite (18) in the form

$$\xi(\zeta) = i\frac{1-\mu(\zeta)}{1+\mu(\zeta)}\eta(\zeta), \quad \zeta \in \Delta. \tag{19}$$

If we denote $\alpha = i\frac{1-\mu}{1+\mu}$ we obtain $\xi(\zeta) = \alpha(\zeta)\eta(\zeta)$, $\zeta \in \Delta\backslash\{0\}$. Because ξ and η are holomorphic sections of $g_\zeta^* H^{1,0}M$ and the fibers of $H^{1,0}M$ have complex dimension 1 we conclude that (19) has a holomorphic solution $\alpha : \Delta\backslash\{0\} \to \mathbf{C}$. The point $\zeta = 0$ is troublesome as ξ and η vanish there.

On the other hand $\mu(\zeta) = \frac{i-\alpha(\zeta)}{i+\alpha(\zeta)}$ and thus we obtain that μ is a holomorphic function on $\Delta\backslash\{0\}$ if we show that $\alpha(\zeta) \neq -i$ for $\zeta \in \Delta\backslash\{0\}$. To see this let us assume $\alpha(\zeta_0) = -i$ for some $\zeta_0 \in \Delta\backslash\{0\}$. This gives

$$\pi^{1,0} g_{\zeta_0 *} X(p) = \pi^{1,0} g_{\zeta_0 *}(Y_1 + iY_2) = 0. \tag{20}$$

But then we use again (12) and write $g_{\zeta *}(X(p)) = \rho_*((\zeta/\bar{\zeta})Z(\zeta(z,w)))$. By definition $\rho : B_r \to D_r$ is a contact mapping for any $0 < r \leq 1$. Therefore

$$g_{\zeta *}(X(p)) = \frac{\zeta}{\bar{\zeta}}[Y(\rho(\zeta(z,w))) + \bar{\nu}(\zeta)\bar{Y}(\rho(\zeta(z,w)))] \tag{21}$$

with $Y(\rho(\zeta(z,w))) \in H^{1,0}_{\rho(\zeta(z,w))}M_{|\zeta|}$. Consequently (20) and (21) would imply

$$g_{\zeta_0 *}X(p) = 0$$

which is impossible, g_{ζ_0} being a diffeomorphism.

If we show that $|\mu| < 1$ on $\Delta\backslash\{0\}$ we can remove the isolated singularity $\xi = 0$ and conclude that μ is holomorphic on Δ. It is enough to show that $|\mu(\zeta)| \neq 1$ for $\zeta \in \Delta\backslash\{0\}$. Assuming $|\mu(\zeta_0)| = 1$ for some $\zeta_0 \in \Delta\backslash\{0\}$ we get $\alpha(\zeta_0) \in \mathbf{R}$ and $\xi(\zeta_0)$, $\eta(\zeta_0)$ are dependent over $\mathbf{R}$. But this is again impossible since $Y_1, Y_2 \in H_pM$ are independent and g_{ζ_0} is a diffeomorphism.

Finally, we note that the above arguments give now (11) for general ζ since μ satisfies (16) (and therefore (17)) in $\Delta\backslash\{0\}$. We can therefore use (21) and the same consideration as for $\zeta \in S^1$ to obtain

$$\nu(\zeta)\frac{\bar{\zeta}}{\zeta} + \mu(\zeta)\frac{\zeta}{\bar{\zeta}} = 0, \quad \zeta \in \bar{\Delta}\backslash\{0\}$$

which proves (11). □

Corollary 4. — *Consider $Y = \pi^{1,0}\rho_* Z$ as a vector field on M and define the action $g_\zeta(p) = \rho(\zeta\rho^{-1}(p))$ as above. Then*

$$g_{\zeta *}Y = Y' + \bar{\eta}(\zeta)\bar{Y}' \tag{22}$$

where η is holomorphic in Δ.

Proof. — Determining the Lempert-Beltrami coefficient μ of the trajectory of p by the vector $X_p = Y_p + \bar{\nu}(1)Y_p$ as above, we have from the argument in (17), (18) that

$$\pi^{1,0}[\mu(\zeta)g_{\zeta*}X + g_{\zeta*}\bar{X}] = 0$$

for all points $\zeta \in \Delta$. Since by assumption (22)

$$\pi^{1,0}[\mu(\zeta)g_{\zeta*}X + g_{\zeta*}\bar{X}] = (\mu(\zeta) + \mu(\zeta)\eta(\zeta)\bar{\nu}(1) + \eta(\zeta) + \nu(1))Y'_{g_\zeta(p)},$$

we must therefore have that

$$\eta(\zeta) = -\frac{\mu(\zeta) + \nu(1)}{1 + \bar{\nu}(1)\mu(\zeta)}.$$

As g_ζ is an inner action, it follows that $\eta(\zeta)$ is holomorphic. □

3. Holomorphic motions and Kobayashi indicatrix

The proof of Theorem 1 can now be obtained quickly from the results of the previous section. Indeed, the Lempert mapping gives a natural homeomorphism R from the unit ball B to the Kobayashi indicatrix $\bar{I}$ of the domain D,

$$R(w) = \frac{d}{d\zeta}\rho(\zeta w)_{|\zeta=0}.$$

The mapping R is **C**-homogenous of degree 1 on complex lines through the origin and as shown in **[5]**, **[11]** it is a contact transformation on ∂B. Therefore our holomorphic motion of the boundary $\partial\bar{I}$ of the indicatrix,

$$\Psi(\lambda, v) = \frac{f_v(\lambda)}{\lambda}, \qquad (\lambda, v) \in \Delta \times \partial\bar{I},$$

admits a natural factorization

$$\Psi(\lambda, v) = \frac{1}{\lambda}g_\lambda \circ \sigma(v),$$

where we have used the abreviation

$$\sigma = \rho|_{\partial B} \circ R^{-1}|_{\partial\bar{I}}.$$

Let $Y = \pi^{1,0}\rho_* Z$ be the vector field as above in Corollary 4. Since $\sigma : \partial\bar{I} \to \partial D$ is a smooth contact transformation we find a vector field W, $W(x) \in H^{1,0}_x\partial\bar{I}$ for all $x \in \partial\bar{I}$, such that

$$\sigma_* W = Y + \bar{\kappa}\bar{Y}$$

on $\partial\bar{I}$.

With these tools we can now estimate the quasiconformal dilatation of the holomorphic motions $\Psi(\lambda, \cdot)$. Namely by (8) one has to bound their Beltrami differentials and for this we have from Corollary 4 that

$$\Psi_*(\lambda,\cdot)W = \frac{1}{\lambda}(g_\lambda)_*(Y+\bar{\kappa}\bar{Y}) \tag{23}$$

$$= \frac{1}{\lambda}(Y' + \bar{\eta}(\lambda)\bar{Y}' + \bar{\kappa}\bar{Y}' + \eta(\lambda)\bar{\kappa}Y') \tag{24}$$

$$= Y'' + \frac{\bar{\lambda}}{\lambda}\cdot\frac{\bar{\eta}(\lambda)+\bar{\kappa}}{1+\bar{\eta}(\lambda)\kappa}\bar{Y}'',$$

where $Y'' = \lambda^{-1}(1+\eta(\lambda)\bar{\kappa})Y'$. Consequently, the absolute values satisfy $K(\Psi_*(\lambda,\cdot)) = (1+|\delta(\lambda)|)/(1-|\delta(\lambda)|)^{-1}$, where

$$\delta(\lambda) = \frac{\eta(\lambda)+\kappa}{1+\eta(\lambda)\bar{\kappa}}. \tag{25}$$

Since by Corollary 4, η is holomorphic in the unit disk with $||\eta||_\infty \le 1$, the same holds true for δ as well. On the other hand, as $\Psi_*(0,\cdot)$ is the identity mapping, $\delta(0)=0$. By Schwartz lemma we then get $|\delta(\lambda)| \le |\lambda|$ which completes the proof of Theorem 1.

Remark. — We would like to mention at the end that the results of Theorems 1 and 3 are true in more general situations than for strictly linearly convex domains. Whenever we have a Lempert mapping with the properties stated at the beginning the proofs carry over. In particular, the new work **[2]** finds the Lempert mappings for the larger class named circular-like domains; these are characterized by the fact that their Lempert invariants are bounded by 1.

Acknowledgements. — We would like to thank Ch. Leuenberger for useful conversations and I. Holopainen for reading the manuscript and making valuable comments.

Part of this paper has been written while the second author visited the University of Jyväskylä and part when the first author was visiting Institute of Advanced Study and University of Paris, Orsay. They are both thankful for the warm hospitality and excellent working conditions at these institutions. The second author is supported by a grant from the Swiss National Science Foundation.

References

[1] K. Astala, G. Martin, Holomorphic motions, *Preprint Univ. of Helsinki*, 1993.

[2] Z. Balogh, Ch. Leuenberger, Riemann mappings and circular representations of nonconvex domains in C^2, *Preprint Univ. Bern, 1996.*

[3] A. Douady, Prolongements de mouvements holomorphes, *Astérisque* **227**, 1995, pp. 1–12.

[4] A. Koranyi, H.M. Reimann, Quasiconformal mappings on CR manifolds, *Lecture Notes in Math.* **1422**, 1990, pp. 59–75.

[5] L. Lempert, La métrique de Kobayashi et la représentation des domaines sur la boule, *Bull. Soc. math. France* **109** (1981), pp. 427–474.

[6] L. Lempert, Intrinsic distances and holomorphic retracts, *Complex Analysis and Applications*, Sofia 1984, pp. 341–364.

[7] L. Lempert, On three-dimensional Cauchy-Riemann manifolds, *J. Amer. Math. Soc.* **5** (1992), pp. 923–969.

[8] L. Lempert, Holomorphic invariants, normal forms, and the moduli space of convex domains, *Ann. of Math.* **128** (1988), pp. 43–78.

[9] R. Mañé, P. Sad, D. Sullivan, On the dynamics of rational maps, Ann. Sci. Éc. Norm. Sup. **16** (1983) pp. 193–217.

[10] M. Y. Pang, Smoothness of the Kobayashi metric of non-convex domains *Internat. J. Math.* **4** (1993), pp. 953–987.

[11] S. Semmes, A Generalization of Riemann Mappings and Geometric Structures on a Space of Domains in C^n, *Memoirs of the AMS* **98** Nr. 472 (1992).

[12] Z. Slodkowski, Holomorphic motions and polynomial hulls, *Proc. Amer. Math. Soc.* **111** (1991), pp. 347–357.

K. Astala, University of Jyväskylä, Department of Mathematics, P.O. Box 35, FIN-40351 Jyväskylä, Finland • *E-mail :* `astala@jane.math.jyu.fi`

Z. Balogh, Universität Bern, Mathematisches Institut, Sidlerstrasse 5, CH-3012 Bern, Switzerland *E-mail :* `zoltan@math-stat.unibe.ch`

H.M. Reimann, Universität Bern, Mathematisches Institut, Sidlerstrasse 5, CH-3012 Bern, Switzerland • *E-mail :* `reimann@math-stat.unibe.ch`

MICHAEL BENEDICKS
LAI-SANG YOUNG

Markov extensions and decay of correlations for certain Hénon maps

Astérisque, tome 261 (2000), p. 13-56
<http://www.numdam.org/item?id=AST_2000__261__13_0>

Astérisque
261, 2000, p. 13–56

MARKOV EXTENSIONS AND DECAY OF CORRELATIONS FOR CERTAIN HÉNON MAPS

by

Michael Benedicks & Lai-Sang Young

Abstract. — Hénon maps for which the analysis in [BC2] applies are considered. Sets with good hyperbolic properties and nice return structures are constructed and their return time functions are shown to have exponentially decaying tails. This sets the stage for applying the results in [Y]. Statistical properties such as exponential decay of correlations and central limit theorem are proved.

0. Introduction and statements of results

Let $T_{a,b} : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T_{a,b}(x, y) = (1 - ax^2 + y, bx).$$

In **[BC2]**, Carleson and the first named author developed a machinery for analyzing the dynamics of $T_{a,b}$ for a positive measure set of parameters (a, b) with $a < 2$ and b small. For lack of a better word let us call these the "good" parameters. The machinery of **[BC2]** is used in **[BY]** to prove that for every "good" pair (a, b), $T = T_{a,b}$ admits a Sinai-Ruelle-Bowen measure ν. The significance of ν is that it describes the asymptotic orbit distribution for a positive Lebesgue measure set of points in the phase space, including most of the points in the vicinity of the attractor. The aim of the present paper is to show that (T, ν) has a natural "Markov extension" with an exponentially decaying "tail", and to obtain via this extension some results on stochastic processes of the form $\{\varphi \circ T^n\}_{n=0,1,2,...}$, where $\varphi : \mathbb{R}^2 \mapsto \mathbb{R}$ is a Hölder continuous random variable on the probability space $(\mathbb{R}^2, \nu)$.

1991 ***Mathematics Subject Classification.*** — 37A25, 37C05, 37C70, 37D25, 37D50, 37E20.
Key words and phrases. — Hénon map, Markov extension, Decay of correlation, Central Limit Theorem.

The research of the first author is partially supported by the Swedish Natural Science Research Council and the Göran Gustafsson Foundation.
The research of the second author is partially supported by the NSF.

Consider in general a map $f : M \circlearrowleft$ preserving a probability measure ν. By a *Markov extension* of (f, ν) we refer to a dynamical system $F : (\Delta, \tilde{\nu}) \circlearrowleft$ and a projection map $\pi : \Delta \mapsto M$; F is assumed to have a Markov partition (with possibly infinitely many states), F and π satisfy $\pi \circ F = f \circ \pi$, and $\pi_* \tilde{\nu} = \nu$. We do not require that π be 1-1 or onto.

Let (f, ν) be as in the last paragraph, and let X be a class of functions on M. We say that (f, ν) has *exponential decay of correlations* for functions in X if there is a number $\tau < 1$ such that for every pair $\varphi, \psi \in X$, there is a constant $C = C(\varphi, \psi)$ such that

$$\left| \int \varphi(\psi \circ f^n) d\nu - \int \varphi d\nu \int \psi d\nu \right| \leq C\tau^n \qquad \forall n \geq 0.$$

Also, we say that (f, ν) has a *central limit theorem* for φ with $\int \varphi d\nu = 0$ if the stochastic process $\varphi, \varphi \circ f, \varphi \circ f^2, \ldots$ satisfies the central limit theorem, *i.e.* if

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \varphi \circ f^i \xrightarrow{\text{dist}} \mathcal{N}(0, \sigma)$$

for some $\sigma \geq 0$. For $\sigma > 0$ this means that $\forall t \in \mathbb{R}$,

$$\nu \left\{ \frac{1}{\sqrt{n}} \sum_{0}^{n-1} \varphi \circ f^i < t \right\} \longrightarrow \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{t} e^{-u^2/2\sigma^2} du$$

as $n \to \infty$.

For $f = T_{a,b}$, (a, b) "good" parameters, we have the following results:

Theorem 1 **([BY]).** — *f admits an SRB measure ν.* (See Section 1.7 for the precise definition.)

Theorem 2 **([BY]).** — *ν is the unique SRB measure for f^n for every $n \geq 1$. This implies in particular that (f^n, ν) is ergodic $\forall n \geq 1$.*

By the general theory of SRB measures, the ergodicity of (f^n, ν) for all $n \geq 1$ is equivalent to (f, ν) having the mixing property, or that it is measure-theoretically isomorphic to a Bernoulli shift, see **[L]**.

For $\gamma > 0$, let $\mathcal{H}_\gamma$ be the space of Hölder continuous functions on $\mathbb{R}^2$ with Hölder exponent γ.

Theorem 3. — *(f, ν) has exponential decay of correlations for functions in $\mathcal{H}_\gamma$. The rate of decay, τ, may depend on γ.*

Theorem 4. — *(f, ν) has a central limit theorem for all $\varphi \in \mathcal{H}_\gamma$ with $\int \varphi d\nu = 0$; the standard deviation $\sigma > 0$ iff $\varphi \neq \psi \circ f - \psi$ for some $\psi \in L^2(\nu)$.*

Theorems 1 and 2 are proved in **[BY]**, while theorems 3 and 4 are new and are proved in this paper. But since an SRB measure is constructed in the process of proving Theorem 3, this paper also contains an independent proof of Theorem 1.

Questions of ergodicity or uniqueness of SRB measures, however, are of a different nature. We will *assume* Theorem 2 for purposes of the present paper.

As mentioned earlier on, our proof of theorems 3 and 4 are carried out using a Markov extension with certain special properties. The second named author has since extended this scheme of proof to a wider setting. We will refer to **[Y]** for certain facts not specific to the Hénon maps, but will otherwise keep the discussion here as self-contained as possible.

The following is a comprehensive summary of what is in this paper, section by section.

In Section 1 we recall from **[BC2]** and **[BY]** some pertinent facts about f.

The aim of Section 2 is to clean up the notion of distance to the "critical set" previously used in **[BC2]** and **[BY]**. We prove that the various distances used before are equivalent.

Section 3 is devoted to organizing the dynamics of f in a coherent fashion. We focus on a naturally defined Cantor set Λ with a product structure defined by local stable and unstable curves and with Λ intersecting each local unstable curve in a positive Lebesgue measure set. The dynamics on Λ is analogous to that of Smale's horseshoe, except that there are infinitely many branches with variable return times. A precise description of Λ is given in Propostion A in Section 3.1.

In Section 4 we study the return time function $R : \Lambda \to \mathbb{Z}^+$, *i.e.* $z \in \Lambda$ returns to Λ after $R(z)$ iterates in the representation above. (Note that $R(z)$ is not necessarily the first return time.) We prove that the measure of $\{R > n\}$ decays exponentially fast as $n \to \infty$. This estimate is stated in Lemma 5 in Section 4.1; it plays a crucial role in the subsequent analysis.

In Section 5 we consider the quotient space $\bar{\Lambda}$ obtained by collapsing Λ along W^s_{loc}-curves. We prove, modifying standard arguments for Axiom A systems where necessary, that $\bar{\Lambda}$ has a well defined metric structure and that the Jacobians of the induced quotient maps have a "Hölder"-type property. This step paves the way for the introduction of a Perron-Frobenius operator. The results are stated in Proposition B in Section 5.1.

Let $\overline{f^R} : \bar{\Lambda} \circlearrowleft$ denote the return map to Λ. In Section 6 we construct a tower map $F : \Delta \circlearrowleft$ over $f^R : \Lambda \circlearrowleft$ with height R (see Section 6.1). F is clearly an extension of f. A Perron-Frobenius operator is introduced for $\overline{F} : \overline{\Delta} \circlearrowleft$, the object obtained by collapsing W^s_{loc}-curves in Δ. At this point we appeal to a theorem in **[Y]** on the spectral properties of certain abstractly defined Perron-Frobenius operators. We explain briefly how a gap in the spectrum of this operator implies exponential decay of correlation for f, referring again to **[Y]** for the formal manipulations, and finish with a proof of the Central Limit Theorem.

Acknowledgements. — The authors are grateful to IHES, the University of Warwick and particularly MSRI, where part of this work was done. Benedicks wishes also to acknowledge the hospitality of UCLA.

1. Dynamics of certain Hénon maps

The purpose of this section is to review some of the basic ideas in [**BC2**] and [**BY**], and to set some notations at the same time. We would like to make the main ideas of this paper accessible to readers without a thorough knowledge of [**BC2**] and [**BY**], but will refer to these papers for technical information as needed. The summary in Section 1 of [**BY**] may be helpful.

1.1. General description of attractors. — In this paper we are interested in the parameter range $a < 2$ and near 2, $b > 0$ and small. The facts in Section 1.1 are elementary and hold for $f = T_{a,b}$ for an open set of parameters (a, b).

There is a fixed point located at approximately $(\frac{1}{2}, \frac{1}{2}b)$; it is hyperbolic and its unstable manifold, which we will call W, lies in a bounded region of $\mathbb{R}^2$. Let Ω be the closure of W. Then Ω is an attractor in the sense that there is an open neighborhood U of Ω with the property that $\forall\, z \in U$, $f^n z \to \Omega$ as $n \to \infty$.

Away from the y-axis, f has some hyperbolic properties. For example, let $\delta \gg b$ and let $s(v)$ denote the slope of a vector v. Then

(i) on $\{|x| \geq \delta\}$, Df preserves the cones $\{|s(v)| \leq \delta\}$;
(ii) $\exists\, M_0 \in \mathbb{Z}^+$ and $c_0 > 0$ such that if $z, fz, \ldots, f^{M-1}z \in \{|x| \geq \delta\}$ and $M \geq M_0$, then
$$\left|Df_z^M v\right| \geq e^{c_0 M}|v| \quad \forall\, v \text{ with } |s(v)| \leq \delta.$$
It is easy to show, however, that Ω is not an Axiom A attractor.

In contrast to Section 1.1, the statements in Section 1.2–1.6 hold only for a positive measure set of parameters. For the rest of this paper we fix a pair of "good" parameters (a, b) and write $f = T_{a,b}$.

1.2. The critical set. — A subset $\mathcal{C} \subset W$, called the *critical set*, is designated to play the role of critical points for 1-dimensional maps. Points in $\mathcal{C}$ have x-coordinates ≈ 0; they lie on $C^2(b)$ segments of W (a curve is called $C^2(b)$ if it is the graph of a function $y = \varphi(x)$ with $|\varphi'|$, $|\varphi''| \leq 10b$); and they have "homoclinic" behaviour in the sense that if τ denotes a unit tangent vector to W, then for $z \in \mathcal{C}$, $\left|Df_z^j \tau\right| \leq (5b)^j\ \forall\, j \geq 0$.

Other important properties of $z \in \mathcal{C}$ are that $\forall\, n \geq 1$:

(i) $\left|Df_z^n \binom{0}{1}\right| \geq e^{c(n-1)}$ for some $c \approx \log 2$;
(ii) "dist"$(f^n z, \mathcal{C}) > e^{-\alpha n}$ for some small $\alpha > 0$. (The precise meaning of "dist" will be given shortly.)

The idea in **[BC2]**, roughly speaking, is that when an orbit of $z_0 \in \mathcal{C}$ comes near $\mathcal{C}$, there is a near-interchange of stable and unstable directions (hence a setback in hyperbolicity); but then the orbit of z_0 follows that of some $\widetilde{z} \in \mathcal{C}$ for some time, regaining some hyperbolicity on account of (i). To arrange for (i), it is necessary to keep the orbits of $\mathcal{C}$ from switching stable and unstable directions too drastically too soon; hence (ii).

We now give the precise meaning of "dist"$(\cdot, \mathcal{C})$. Consider $z \in \mathcal{C}$ and let $n_1 > 0$ be the first time its orbit returns to $(-\delta, \delta) \times \mathbb{R}$. It is arranged that there is $\widetilde{z}_1 \in \mathcal{C}$ of an earlier generation (see below) with respect to which $f^{n_1}z$ is in *tangential position*, *i.e.* $\widetilde{z}$ lies in a $C^2(b)$ segment of W extending $> 4\,|f^{n_1}z - \widetilde{z}_1|$ to each side of $\widetilde{z}_1$, and the vertical distance between $f^{n_1}z$ and this segment is $< |f^{n_1}z - \widetilde{z}_1|^4$; see e.g. **[BY]**, Subsection 1.4.1. Here "dist"$(f^{n_1}, \mathcal{C})$ means $|f^{n_1}z - \widetilde{z}_1|$.

We say that $f^{n_1}z$ is "bound" to $\widetilde{z}_1 \in \mathcal{C}$ for the next p_1 iterates, where p_1 is the smallest j s.t. $\left|f^{n_1+j}z - f^j\widetilde{z}_1\right| > e^{-\beta j}$ for some fixed $\beta > \alpha$. At time $n_1 + p_1$, we say that the orbit of z is "free", and it remains free until the first $n_2 \geq n_1 + p_1$ when it returns again to $(-\delta, \delta) \times \mathbb{R}$. The binding procedure above is then repeated, with bound period p_2 etc.

It is convenient to modify slightly the above definitions of p_i so that the bound periods become "nested", *i.e.* if a bound period is initiated in the middle of another one, it also expires before the first one does. (See Section 6.2 of **[BC2]**.)

We return to the notion of "generations" to which we referred a few paragraphs back. There is a unique $z_0 \in \mathcal{C}$ lying in the roughly horizontal segment of W containing our fixed point. The part of W between f^2z_0 and fz_0 is denoted by W_1 and called the leaf of generation 1. Leaves of higher generations are defined inductively by $W_n \equiv f^{n-1}W_1 - W_{n-1}$, and a critical point is of generation n if it is in W_n.

1.3. Dynamics on W. — In Proposition 1 of **[BY]**, it is shown that the orbit of every $z \in W$ can be controlled using those of $\mathcal{C}$. More precisely, consider z in a local unstable manifold of our fixed point, and let $n_1 > 0$ be the first time its orbit goes into $(-\delta, \delta) \times \mathbb{R}$. It is shown that there is a "suitable" $\widetilde{z}_1 \in \mathcal{C}$ to which we will regard $f^{n_1}z$ as bound for some period of time. "Suitable" here means that (1) $f^{n_1}z$ is in *generalized tangential position* wrt $\widetilde{z}_1$ (generalized tangential positions are slight generalizations of tangential positions; see Section 1.6 of **[BY]**); and (2) the angle between $\tau(f^{n_1}z)$, the tangent vector to W at $f^{n_1}z$ and a certain vector field about $\widetilde{z}_1$ is "correct"; this will be explained in Section 1.5. After a bound period as defined in Section 1.2, the orbit of z then becomes free until it gets into $(-\delta, \delta) \times \mathbb{R}$ again, finds another suitable point $\widetilde{z}_2 \in \mathcal{C}$ to bind with, and the story repeats itself.

Not only do suitable binding points always exist (**[BC2]**, Section 7.2), it is shown in **[BY]**, Lemma 7, that one could systematically assign to each maximal free segment γ intesecting $(-\delta, \delta) \times \mathbb{R}$ a critical point $\widetilde{z}(\gamma)$ that is suitable for binding for all $z \in \gamma$. The picture is as follows:

(i) If γ contains a critical point $\tilde{z}$, then $\tilde{z}(\gamma) = \tilde{z}$; this is always the case if neither end point of γ lies in $(-\delta, \delta) \times \mathbb{R}$.

(ii) If only one end point of γ lies in $(-\delta, \delta) \times \mathbb{R}$, say the left end point γ_-, and γ does not contain a critical point, then $\tilde{z}(\gamma)$ is taken to be the binding point of γ_- (note that γ_- is also in bound state); $\tilde{z}(\gamma)$ always lies to the left of γ_-, away from γ.

(iii) If both end points $\gamma_\pm$ of γ are in $(-\delta, \delta) \times \mathbb{R}$ and γ does not contain a critical point then the binding point of at least one of $\gamma_\pm$ lies on the opposite side of $\gamma_\pm$ as γ and can be taken to be $\tilde{z}(\gamma)$.

We state some estimates for $|Df_z^n \tau|$, $z \in W$, that are consequences of the behaviour of the critical set and the binding process above. Unless otherwise referenced, these estimates are proved in Corollary 1 of **[BY]**:

(I) *Free period estimates.*

(i) Every free segment γ has slope $< 2b/\delta$, and $\gamma \cap (-\delta, \delta) \times \mathbb{R}$ is a $C^2(b)$ curve (Lemmas 1 and 2, **[BY]**).

(ii) There is $M_0 \in \mathbb{Z}^+$ and $c_0 > 0$ s.t. if z is free and $z, fz, \ldots, f^{M-1}z \notin (-\delta, \delta) \times \mathbb{R}$ for $M \geq M_0$, then $|Df_z^M \tau| \geq e^{c_0 M}$.

(II) *Bound period estimates.*

The following hold for some $c \approx \log 2$: if $z \in (-\delta, \delta) \times \mathbb{R}$ is free and is bound at this time to $\tilde{z} \in \mathcal{C}$ with bound period p, then

(i) if $e^{-\nu-1} \leq |z - \tilde{z}| \leq e^{-\nu}$, then $\frac{1}{2}\nu \leq p \leq 5\nu$;

(ii) $|Df_z^j \tau| \geq |z - \tilde{z}|\, e^{cj}$ for $0 < j < p$;

(iii) $|Df_z^p \tau| \geq e^{cp/3}$.

(III) *Orbits ending in free states.*

There exists $c_1 > \frac{1}{3}\log 2$ s.t. if $z \in W \cap (-\delta, \delta) \times \mathbb{R}$ is in a free state, then

$$|Df_z^{-j} \tau| \leq e^{-c_1 j} \qquad \forall j \geq 0$$

(Lemma 3, **[BY]**).

1.4. Bookkeeping, derivative and distortion estimates. — Let $\mathcal{P}$ be the following partition of the interval $(-\delta, \delta)$: first we write $(-\delta, \delta)$ as the disjoint union $\bigcup\{I_\nu : |\nu| \geq \text{ some } \nu_0\}$ where $I_\nu = (e^{-(\nu+1)}, e^{-\nu})$ for $\nu > 0$ and $I_\nu = -I_\nu$ for $\nu < 0$; then each I_ν is further subdivided into ν^2 intervals $\{I_{\nu,j}\}$ of equal length.

For $x_0 \in \mathbb{R}$, we let $\mathcal{P}_{[x_0]}$ denote a copy of $\mathcal{P}$ with 0 "moved" to x_0. Similarly, if γ is a roughly horizontal curve in $\mathbb{R}^2$ and $z_0 \in \gamma$, we let $\mathcal{P}_{[z_0]}$ denote the obvious partition on γ. Once γ and z_0 are specified, we will use $I_{\nu,j}$ to denote the corresponding subsegment of γ. Also, if $J \in \mathcal{P}_{[\cdot]}$, we let nJ denote the segment n times the length of J centered at J.

The following derivative estimate is very similar to the derivative estimates in the proof of Lemma 7.2 in **[BC2]**.

Derivative estimate. — *Suppose that the point z belongs to a free segment of W and satisfies* $\operatorname{dist}(f^j z, \mathcal{C}) \geq \delta e^{-\alpha j}$ $\forall\, j < n$ *for some integer n. Then there is a constant $c_2 > 0$ such that*

$$|Df_z^n \tau| \geq \delta e^{c_2 n}. \tag{1.1}$$

Since the proof follows step by step the proof of Lemma 7.2 in **[BC2]**, it is omitted. The only difference is that in the present situation the allowed approach rate to the critical set is much slower than that in Lemma 7.2 of **[BC2]**: $\operatorname{dist}(f^n z, \mathcal{C}) \geq \delta e^{-\alpha n}$ $\forall\, n \geq 0$, versus $\operatorname{dist}(f^n z, \mathcal{C}) \geq e^{-36n}$ $\forall\, n \geq 0$. This leads to the the expansion estimate of (1.1).

The following is proved in Proposition 2 of **[BY]**:

Distortion estimate. — *Let $\gamma \subset (-\delta, \delta) \times \mathbb{R}$ be a segment of W. We assume that the entire segment has the same itinerary up to time N in the sense that*

(i) *all $z \in \gamma$ are bound or free simultaneously at any one moment in time;*

(ii) *if $0 = t_0 < t_1 < \cdots < t_q$ are the consecutive free return times before N, then $\forall\, k \leq q$ the entire segment $f^{t_k}\gamma$ has a common binding point $z^{(k)} \in \mathcal{C}$ and $f^{t_k}\gamma \subset 5J_k$ for some $J_k \in \mathcal{P}_{[z^{(k)}]}$.*

Then $\exists\, C_1$ independent of γ or N s.t. $\forall\, z_1, z_2 \in \gamma$,

$$\frac{|Df_{z_1}^N \tau|}{|Df_{z_2}^N \tau|} \leq C_1.$$

1.5. Fields of contracted directions. — First we state a general perturbation lemma for matrices. Given $A_1, A_2, \ldots$, we write $A^n := A_n \cdots A_1$. The following is a slight paraphrasing of Lemma 5.5 and Corollary 5.7 in **[BC2]**. All the matrices below are assumed to have $|\det| = b$.

Matrix perturbation lemma. — *Given $\kappa \gg b$, $\exists\, \lambda$ with $b \ll \lambda < \min(1, \kappa)$ s.t. if $A_1, \ldots, A_n$, $A'_1, \ldots, A'_n \in GL(2, \mathbb{R})$ and $v \in \mathbb{R}^2$ satisfy*

$$|A^i v| \geq \frac{1}{2}\kappa^i \qquad \textit{and} \qquad \|A_i - A'_i\| < \lambda^i \qquad \forall\, i \leq n,$$

then we have, for all $i \leq n$:

(i) $\left|A'^i v\right| \geq \frac{1}{2}\kappa^i$;

(ii) $\sphericalangle(A^i v, A'^i v) \leq \lambda^{i/4}$.

If $A \in G(2, \mathbb{R})$ is s.t. $|Av|/|v| \neq$ const, let $e(A)$ denote one of the two unit vectors most contracted by A. We will write $e_n(z) := e(Df_z^n)$ wherever it makes sense. From the perturbation lemma above, it follows that if $|Df_{z_0}^j v| \geq \kappa^j$, $0 \leq j \leq n$, for some κ and some v, then there is a ball B_n of radius $(\lambda/5)^n$ about z_0 on which e_n is defined and has the property that $|Df^n e_n| \leq 2(b/\kappa)^n$.

Assuming that κ is fixed and e_n is defined in a neighborhood of z_0 as above, the following hold (Section 5, **[BC2]**):

(i) e_1 is defined everywhere and has slope $= 2ax + \mathcal{O}(b)$;
(ii) $|e_n - e_m| \le \mathcal{O}(b^m)$ for $m < n$;
(iii) for (x_1, y_1), $(x_2, y_2) \in$ some B_n with $|y_1 - y_2| \le |x_1 - x_2|$,
$$|e_n(x_1, y_1) - e_n(x_2, y_2)| = (2a + \mathcal{O}(b))\,|x_1 - x_2|\,.$$

The perturbation lemma above applies in particular to critical points; see Section 1.2. Indeed, every $z_0 \in \mathcal{C}$ is constructed as the limit of a sequence $\{z_n\}$ where z_n is the unique point in the $C^2(b)$ segment of W containing z_0 with $\tau(z_n) = e_n(z_n)$. Going back to the notion of "suitability" of binding points at the beginning of Section 1.3, a formulation of requirement (2) could be that

$$3\,|\tilde{z}_1 - f^{n_1} z| \le \sphericalangle\left(\tau\left(f^{n_1} z\right), e_{\ell_1}\left(f^{n_1} z\right)\right) \le 5\,|\tilde{z}_1 - f^{n_1} z|$$

where $\ell_1 \approx -\varepsilon \cdot \log|f^{n_1} z - \tilde{z}_1|$ is small enough that e_{ℓ_1} is defined on a neighborhood of $\tilde{z}_1$ containing $f^{n_1} z$.

1.6. More on the geometry of the critical set. — The following facts about the relative locations of critical points are used in sections 2 and 3 of this paper.

***Fact 1* (Lemma 5, [BY]).** — *Let $\tilde{z} \in \mathcal{C}$ be contained in a $C^2(b)$ curve $\gamma \subset W$. Assume that γ extends to $> 2d$ on each side of $\tilde{z}$, and let $\zeta \in \gamma$ be s.t. $|\tilde{z} - \zeta| = d$. Then there are no critical points z with $|z - \zeta| < d^2$.*

***Fact 2* (Existence of critical points, [BC2] Section 6.2, [BY] Subsection 1.3.1)**

There is a number ρ, $b \ll \rho \ll 1$, s.t. the following holds: if $z = (x, y)$ lies in a $C^2(b)$ segment $\gamma \subset W$ of generation n with γ extending $> 2\rho^n$ to each side of z, and there is a critical point $\tilde{z} = (\tilde{x}, \tilde{y}) \in \mathcal{C}$ s.t.

(i) *$x = \tilde{x}$,*
(ii) *$\tilde{z}$ is of generation $< n$,*
(iii) *$|z - \tilde{z}| < b^{n/540}$,*

then there is a unique critical point $\hat{z} = (\hat{x}, \hat{y}) \in \gamma$ with $|x - \hat{x}| < |y - \hat{y}|^{1/2}$.

One way to get a sense of the relative location of a point to the critical set is to do the "capture" procedure introduced in **[BC2]**, sections 6.4 and 7.2. This procedure guarantees that near every free $z \in W$ there are many long $C^2(b)$ segments of W some of which will contain critical points. The picture is as follows (for a precise statement see **[BY]** Subsection 2.2.2):

If $z \in W$ is free, then there is a family of $C^2(b)$ subsegments of W labeled $\{\gamma_i\}_{i=1,2,\ldots,i(z)}$, where $i(z)$ is the last integer i with $3^{i+1} < \mathrm{gen}(z)$, s.t.

(i) $m \le$ generation of $\gamma_i \le 3m$, $m = 3^i$,
(ii) γ_i is centered at $\approx z$, and has length $\approx 10\rho^m$,
(iii) $\mathrm{dist}(z, \gamma_i) < (Cb)^m$.

(iv) if $z^{(i)}$ is the point on γ_i with the same x-coordinate as z then $|\tau(z)-\tau(z^{(i)})| \leq (Cb)^{m/6}$.

There are, in fact, two such families, one above and one below z.

One may assume that γ_1 contains a critical point. If this critical point is sufficiently near the middle of γ_1, then by Fact 2, γ_2 would also contain a critical point. This may continue all the way down the stack, or there may exist an i s.t. $\tilde{z}_i \in \mathcal{C} \cap \gamma_i$ is so far to one side that no critical point lies in γ_{i+1}. If this happens z is in tangential position with respect to $\tilde{z}_i$.

1.7. SRB measures. — In this article (and also in **[BY]**), an f-invariant Borel probability measure ν is called an *SRB measure* if f has a positive Lyapunov exponent ν-a.e. and the conditional measures of ν on unstable manifolds are absolutely continuous with respect to the Riemannian measures on these leaves. The following are proved in **[BY]**:

(i) f admits an SRB measure ν;
(ii) ν is unique (*i.e.* f admits no other SRB measure); hence (f,ν) is ergodic;
(iii) (ii) is in fact true for f^n for all $n \geq 1$.

It follows from general nonuniform hyperbolic theory that (iii) is equivalent to (f,ν) having the mixing property, the K-property, and in fact to its being isomorphic to a Bernoulli shift (see e.g. **[L]**).

2. Preliminaries: cleaning up the notion of $\mathrm{dist}(\cdot,\mathcal{C})$ in [BC2] and [BY2]

In this section as in the rest of the paper, it is assumed that $f = T_{a,b}$ where (a,b) are "good parameters" as discussed in Section 1.

Two notions of the distance to the critical set for a point z on a free segment of W have been used in **[BC2]** and **[BY]**. The first is a pointwise definition, in which we think of $\mathrm{dist}(z,\mathcal{C})$ as $|z-\tilde{z}(z)|$, where $\tilde{z}$ is a certain critical point captured by z. We will call this distance $d_{\mathrm{cap}}(z,\mathcal{C})$. The second notion is more globally defined. It is shown in **[BY]** that one could systematically assign a critical point $\tilde{z}(\tilde{\gamma})$ to every maximal free segment $\tilde{\gamma}$. Let us define $d_\gamma(z,\mathcal{C})$ to be $|\tilde{z}(\tilde{\gamma})-z|$, where $\tilde{\gamma}$ is the maximal free segment containing γ. Precise definitions of $d_{\mathrm{cap}}(\cdot,\mathcal{C})$ and $d_\gamma(\cdot,\mathcal{C})$ are given below. The main purpose of this section is to prove

Lemma 1. — *For each point belonging to a free segment of W, we have*

$$\frac{d_{\mathrm{cap}}(z,\mathcal{C})}{d_\gamma(z,\mathcal{C})} = 1 + \mathcal{O}(\max(b,d^2)),$$

where $d = \min(d_{\mathrm{cap}}(z,\mathcal{C}), d_\gamma(z,\mathcal{C}))$.

We state also a related fact which is, in some ways, more basic:

Lemma 1'. — *Suppose that z is a point that is in tangential position to two different critical points z_1 and z_2. Let $d_1 = |z - z_1|$, $d_2 = |z - z_2|$ and $d = \min(d_1, d_2)$. Then*

$$\frac{d_1}{d_2} = 1 + \mathcal{O}(\max(b, d^2)).$$

We now begin to justify these claims. The following technical sublemma along with Lemma 5 in **[BY]** (see Fact 1 Section 1.6) will be used repeatedly to rule out the presence of critical points in certain regions. Part (b) has independent interest; it plays an important role, for instance, in the proof of Lemma 1.

Sublemma 1. — *Let γ_0 and γ be two free $C^2(b)$ segments of W. Suppose that γ_0 contains a critical point z_0, and that there exist two points $\zeta_0 \in \gamma_0$ and $\zeta \in \gamma$ with the same x-coordinate. Let $d_0 = |\zeta_0 - z_0|$.*

(a) *If $|\zeta - \zeta_0| < d_0^4$ and $|\tau(\zeta) - \tau(\zeta_0)| < d_0^2$, then for all $z \in \gamma$ with $|z - z_0| = d \geq d_0$, there can be no critical point at a distance $< d^2$ from z.*

(b) *The assumptions in part (a) are satisfied if ζ is (say) the left end point of γ, it is in a bound state, and its binding point z_0 lies to the left of γ.*

In the situation of part (b), Sublemma 1 allows us to essentially regard γ as a continuation of γ_0 (which may not be very long compared to γ).

Proof. — The proof of (a) is a slight modification of that of Lemma 5 in **[BY]** and will be omitted.

To prove (b) let us first briefly review the binding procedure. For a detailed account see **[BC2]**, sections 6 and 7 and **[BY]**, subsections 1.6.2 and 2.2.2. Let n be the generation of γ, and assume that attached to the left endpoint ζ of γ is a bound segment B. Recall that there is a hierarchy of bindings associated with B. We let $\tilde{z}_0$ be the critical point with the property that at this time, *i.e.* at time n, B is bound to $\tilde{z}_m = f^m \tilde{z}_0$ and $\tilde{z}_m$ is free. Let $\tilde{z}^*$ denote the new binding point acquired by $\tilde{z}_m$ at this time. Then $\tilde{z}^*$ is located on a segment of generation $m_1 < m$. The capture procedure resulting in $\tilde{z}^*$ calls for $\tilde{z}_{m-m_1}$ to be in a favorable position (in particular out of all fold periods); $\tilde{z}_{m-m_1}$ then draws in a segment γ' of W_1 and $\tilde{z}^*$ lies on $f^{m_1}\gamma'$.

We claim that $f^{-m_1}\zeta$ is outside of all fold periods. First, it cannot be in a fold period initiated in the time interval $[n - m, n]$, since bindings and the corresponding fold periods initiated in this time interval are the same as those of $\tilde{z}_0$. Suppose then $f^{-m_1}\zeta$ is in a fold period initiated before time $n - m$. The corresponding bound period in this case would have to last $> (C \log(1/b))m$ iterates beyond time $n - m_1$, contradicting our assumption that ζ is free.

Having established that $f^{-m_1}\zeta$ lies in a segment of W sufficiently parallell to W_1, the estimates in (a) follow immediately from capture arguments and the Matrix Perturbation Lemma in Section 1.5. □

Definition of $d_{\rm cap}(z, \mathcal{C})$. — Let $\{\gamma_i\}_{i=1,\dots,k}$ be a stack of leaves captured by z. We let i_* be the largest integer i such that γ_i contains a critical point.

Case 1: $i_* < k$. — Let $\tilde{z}(z)$ be the critical point on γ_{i_*}.

Case 2: $i_* = k$. — We will show in this case that there is a critical point on $\tilde{\gamma}$, the maximal free segment containing γ. This critical point is unique (this follows e.g. from Lemma 5 in **[BY]**); it will be our $\tilde{z}(z)$. The existence of $\tilde{z}(z)$ follows readily from Fact 2, Section 1.6, once we verify that γ extends $\geq 2\rho^g$ on both sides of z and z_*, g being the generation of z and z_* the critical point on γ_{i_*}. We leave this as an exercise.

Definition of $d_\gamma(z, \mathcal{C})$. — Let γ be a maximal free segment. In Lemma 7 of **[BY]**, we established a rule for assigning a critical point $\tilde{z}(\gamma)$ to each γ. See Section 1.3 for what is proved. Given that the binding points of all critical orbits are selected and fixed, the only situation for which there might be some ambiguity in the choice of $\tilde{z}(\gamma)$ is when both end points of γ are in $(-\delta, \delta) \times \mathbb{R}$, *i.e.* case (iii) in Section 1.3. Figure 1 shows all possible configurations of the locations of the binding points $\tilde{z}_+$ (resp. $\tilde{z}_-$) relative to γ_+ (resp. γ_-).

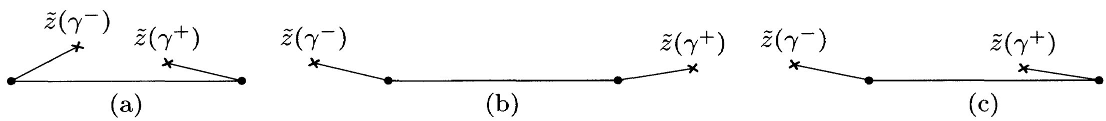


FIGURE 1

In **[BY]** we ruled out (a); (b) will be eliminated in Sublemma 2 below. What is left is (c) (and its mirror image). If (c) occurs , $\tilde{z}(\gamma) = \tilde{z}_-$.

Proof of Lemma 1. — Let $z \in \gamma$, where γ is a maximal free segment of W. The idea of our proof is as follows. Look at the contractive fields centered at $\tilde{z}(\gamma)$ and $\tilde{z}(z)$. We will show that for a suitable choice of m, z lies in the domains of the e_m-fields induced by both points. Since the angle between $e_m(z)$ and $\tau(z)$ is supposed to reflect the distance between z and the respective critical points, we must have $|\tilde{z}(\gamma) - z| \approx |\tilde{z}(z) - z|$.

We consider the case where $\tilde{z}(\gamma) \notin \gamma$. (The proof is slightly simpler when γ contains a critical point.) Let $d_\gamma = |\tilde{z} - z|$, $d_c = |\tilde{z}(z) - z|$. First we observe the weaker estimate $d_\gamma^2 \leq d_c \leq d_\gamma^{1/2}$; to see that $d_c \geq d_\gamma^2$, use Sublemma 1 (both (a) and (b)); to see that $d_\gamma \geq d_c^2$, use Lemma 5 in **[BY]**. Let m_γ and m_c be defined by

$$\begin{cases} \left(\dfrac{\lambda}{5}\right)^{2m_c} \leq d_c \leq \left(\dfrac{\lambda}{5}\right)^{2m_c - 1}, \\ \left(\dfrac{\lambda}{5}\right)^{2m_\gamma} \leq d_\gamma \leq \left(\dfrac{\lambda}{5}\right)^{2m_\gamma - 1}, \end{cases}$$

and let $m = \min(m_\gamma, m_c)$. Then the above relation between d_γ and d_c implies that

$$\frac{1}{2} \leq \frac{m}{m_\gamma}, \frac{m}{m_c} \leq 1.$$

Thus z lies well inside the balls $B_{(\lambda/5)^m}(\tilde{z}(z))$ and $B_{(\lambda/5)^m}(\tilde{z}(\gamma))$, the domains of e_m around these points.

Let $\hat{z}$ be the point on $\tilde{\gamma}$, the $C^2(b)$-segment containing $\tilde{z}(z)$, having the same x-coordinate as z. Since $|\tau(\tilde{z}(z)) - e_m(\tilde{z}(z))| = \mathcal{O}(b^m)$ and $\tilde{\gamma}$ is $C^2(b)$, it follows that

$$|\tau(\hat{z}) - e_m(\hat{z})| = (2a + \mathcal{O}(b))d_c. \tag{2.1}$$

We would like to duplicate the estimate in the last paragraph with z playing the role of $\hat{z}$ and $\tilde{z}(\gamma)$ instead of $\tilde{z}(z)$, except that z does not lie on γ_0, the $C^2(b)$ segment containing $\tilde{z}(\gamma)$, and in any case we do not know how long γ_0 is. To get around this, note that

$$\sphericalangle(\tau(\tilde{z}(\gamma)), \tau(z)) \leq \sphericalangle(\tau(\tilde{z}(\gamma)), \tau(\zeta_0)) + \sphericalangle(\tau(\zeta_0), \tau(\zeta)) + \sphericalangle(\tau(\zeta), \tau(z)),$$

where ζ is the end point of γ closer to $\tilde{z}(\gamma)$ and ζ_0 is the point on γ_0 with the same x-coordinate as ζ. Part (b) of Sublemma 1 then gives

$$|\tau(z) - e_m(z)| = \left(2a + \max\left(\mathcal{O}(b), \mathcal{O}\left(|\zeta - \tilde{z}(\gamma)|^2\right)\right)\right) \cdot d_\gamma. \tag{2.2}$$

Finally, since $\tilde{\gamma}$ is obtained by capturing we have $|\tau(\hat{z}) - \tau(z)| \ll d_c^8$, say. Also, $|e_m(\hat{z}) - e_m(z)| \leq 10 d_c^4$ (apply Property (iii) of Section 1.5 twice). These together with (2.1) and (2.2) give the desired result. □

We omit the proof of Lemma 1', which is very similar to that of Lemma 1.

Remark 1. — This proof shows that the intuitive definition of the distance to the critical set for a point $z \in W$ really ought to be the angle between $\tau(z)$ and $e_m(z)$, where e_m is the contractive field of a "suitable" order.

In order to make the definition of $d_\gamma(z, \mathcal{C})$ unambiguous it remains to prove

Sublemma 2. — *The configuration in Figure 1(b) does not occur.*

Proof of Sublemma 2. — The proof is based on the same ideas as that of Lemma 1. Fix an arbitrary $z \in \gamma$. Then Sublemma 1 applied to $\tilde{z}_\pm$ tells us that

$$|z - \tilde{z}_+|^2 \leq |z - \tilde{z}_-| \leq |z - \tilde{z}_+|^{1/2}.$$

Let $d = \max(|z - \tilde{z}_+|, |z - \tilde{z}_-|)$ and m an integer defined by $(\lambda/5)^{2m} \approx d$ as in the proof of Lemma 1. Then z lies well inside the balls $B_{(\lambda/5)^m}(\tilde{z}_-)$ and $B_{(\lambda/5)^m}(\tilde{z}_-)$, the domains of e_m around these points, and we obtain a contradiction since the field arround $\tilde{z}_-$ says $e_m(z)$ must have positive slope and the one around $\tilde{z}_+$ says that the slope is negative. □

3. Construction of a "horseshoe" with positive measure

3.1. Goal of this section. — "Horseshoes" are well known to be building blocks of uniformly hyperbolic systems. We will show in this section that f can be viewed as the discrete time version of a special flow built over a "horseshoe". In order to have positive SRB measure, the "horseshoe" here must necessarily have infinitely many branches with unbounded return times. This picture will be made precise in the statement of Propostion A.

We begin with some formal definitions.

Definition 1

(a) Let Γ^u and Γ^s be two families of C^1 curves in $\mathbb{R}^2$ such that
  (i) the curves in Γ^u, respectively Γ^s, are pairwise disjoint;
  (ii) every $\gamma^u \in \Gamma^u$ meets every $\gamma^s \in \Gamma^s$ in exactly one point; and
  (iii) there is a minimum angle between γ^u and γ^s at the point of intersection.
Then the set
$$\Lambda: \ = \{\gamma^u \cap \gamma^s : \gamma^u \in \Gamma^u, \gamma^s \in \Gamma^s\}$$
is called the *lattice* defined by Γ^u and Γ^s.

(b) Let Λ and Λ' be lattices. We say that Λ' is a *u-sublattice* of Λ if Λ' and Λ have a common defining family Γ^s and the defining family Γ^u of Λ contains that of Λ'; *s-sublattices* are defined similarly.

(c) Given a lattice Λ, $Q \subset \mathbb{R}^2$ is called the *rectangle spanned by* Λ if $\Lambda \subset Q$ and ∂Q is made up of two curves from Γ^u and two from Γ^s.

In Propostion A we will assert the existence of two lattices Λ^+ and Λ^- with essentially identical properties. For notational simplicity let us agree to the following convention: statements about "Λ" will apply to both Λ^+ and Λ^-. For example, "let Γ^u and Γ^s be the defining families of Λ" means there are four families of curves; the families $(\Gamma^u)^+$ and $(\Gamma^s)^+$ define Λ^+ while $(\Gamma^u)^-$ and $(\Gamma^s)^-$ define Λ^-.

Proposition A. — *There are two lattices Λ^+ and Λ^- in $\mathbb{R}^2$ with the following properties. Let Γ^u and Γ^s be the defining families of Λ; for $z \in \Lambda$, let $\gamma^u(z)$ denote the γ^u-curve in Γ^u containing z. Then:*

(1) (Topological structure) Λ *is the disjoint union of s-sublattices $\Lambda_i, i = 1, 2, \ldots,$ where for each i, $\exists\, R_i \in \mathbb{Z}^+$ s.t. $f^{R_i}\Lambda_i$ is a u-sublattice of Λ^+ or Λ^-.*

(2) (Hyperbolic estimates)
  (i) *Every $\gamma^u \in \Gamma^u$ is a $C^2(b)$ curve; and $\exists\, \lambda_1 > 1$ s.t.*
$$|Df_z^{R_i}\tau| \geq \lambda_1^{R_i}$$
$\forall\, z \in \gamma^u \cap Q_i$, τ being a unit tangent vector to γ^u at z and Q_i being the rectangle spanned by Λ_i.

(ii) $\forall\, z \in \Lambda$ *and* $\forall\, \zeta \in \gamma^s(z)$ *we have*

$$d\left(f^j z, f^j \zeta\right) < C b^j \qquad \forall\, j \geq 1.$$

(3) $\mathrm{Leb}(\Lambda \cap \gamma^u) > 0 \quad \forall\, \gamma^u \in \Gamma^u$.

(4) (Return time estimates) *Let* $R\colon \Lambda \to \mathbb{Z}^+$ *be defined by* $R(z) = R_i$ *for* $z \in \Lambda_i$. *Then* $\exists\, C_0 > 0$ *and* $\theta_0 < 1$ *s.t. on every* γ^u,

$$\mathrm{Leb}\,\{z \in \gamma^u \colon R(z) \geq n\} \leq C_0 \theta_0^n \qquad \forall\, n \geq 1.$$

The rest of this section is devoted to proofs of Assertions (1) and (2) in Propostion A; Assertions (3) and (4) are proved in Section 4. There are slight (and totally harmless) inaccuracies in the above formulation of Propostion A. They are noted in Remarks 2 and 4 in Section 3.4.

3.2. Some preliminary constructions. — First we assume f is a 1-dimensional map and construct for f a Cantor set that would play the role of Λ in Propostion A. Then we carry this construction over to W_1, the top leaf of W (see Section 1.2). We will address certain technical problems in 2-d that are not present in 1-d, and conclude that the two Cantor sets we have constructed have identical geometric estimates.

Temporarily then, we think of f as a map of $[-1, 1]$ given by $f(x) = 1 - ax^2$ for some $a \lesssim 2$, and let Ω_0 be one of the two outermost intervals in the partition $\mathcal{P}$ defined in Section 1.4. We define inductively $\Omega_0 \supset \Omega_1 \supset \Omega_2 \supset \ldots$ as follows. Let ω be a connected component of Ω_{n-1}. First we delete from ω the interval $f^{-n}\left(-\delta e^{-\alpha n}, \delta e^{-\alpha n}\right)$; and if the f^n-image of a component of what is left of ω does not contain some $I_{\nu j}$, then we delete that also. What remains goes into Ω_n. Our desired Cantor set is $\Omega_\infty \equiv \bigcap_n \Omega_n$.

We assume the following is true: if M_1 is the minimum time it takes for $x \in (-\delta, \delta)$ to return to $(-\delta, \delta)$, then $e^{\alpha M_1} \geq 10$. We assume also the corresponding fact for our 2-d map. This is easily arranged since α is fixed before we choose a or δ.

Returning to 2-d, we let Ω_0 be the corresponding segment in W_1 and try to construct Ω_n using the same rules and same notations as in 1-d. Let ω be a connected component of Ω_{n-1}. We assume for the moment the following geometric fact:

$(*)$ if part of $f^n\omega$ is bound and part is free, then the bound part lies at one or both ends of $f^n\omega$.

If all of $f^n\omega$ is in the bound state, or if $f^n\omega \cap (-\delta, \delta) \times \mathbb{R} = \varnothing$, do nothing; *i.e.* put $\omega \subset \Omega_n$. If not, let γ be the free part of $f^n\omega$, and let $\tilde{\gamma}$ be the maximal free segment containing γ. We will use as binding point $\tilde{z}(\tilde{\gamma})$, where $\tilde{z}(\cdot)$ is as defined in Section 2. Deletions are then made with respect to this binding point, and $\Omega_\infty = \bigcap \Omega_n$ as before.

To justify $(*)$, consider the function t defined on $f^n\omega$ where $t(z)$ is the time to expiration of all bound periods at z (counting only the ones initiated before this step). We take as our induction hypotheses not only $(*)$ but that $t|f^n\omega$ has the following profile: it is either decreasing (by which we mean non-increasing) or it decreases from one end to its minimum and increases from there to the other end. It is easy to check that this type of profile is maintained on each component of Ω_n even if new bindings are imposed.

We note also that the bound part at each end of $f^n\omega$ (if it exists) is small relative to $|I_{\nu j}|$, where $I_{\nu j}$ is the element of the partition determined by $\tilde{z}(\tilde{\gamma})$ that meets it. We know from 1-d or **[BC2]** that this is true wrt the partition determined by some binding point and Lemma 1 assures us that all binding points are essentially the same. This reasoning also gives us that no bound part is ever deleted.

3.3. Stable curves. — The purpose of this subsection is to construct Γ^s, which will consist of a family of local stable manifolds through $\Omega_0 \subset W_1$. We noted in Section 1.4 that $\exists\, c_2 > 0$ s.t. $|Df_z^n\tau| \geq \delta e^{c_2 n}$ $\forall\, z \in \Omega_n$. From Section 1.5 then it follows that e_n, the field of most contracted directions of Df^n, is defined in a neighborhood of every $z \in \Omega_n$. To construct Γ^s, however, we need to know that the domain of e_n is larger than this.

As noted in Section 1.5, e_1 is defined everywhere. We integrate e_1, and let $Q_0 = \bigcup_{z\in\tilde{\Omega}_0} \gamma_1(z)$ where $\gamma_1(z)$ is the integral curve segment of length $10b$ centered at z, and $\tilde{\Omega}_0$ is the (Cb)-neighborhood of Ω_0 in W_1. We will not need this for some time, but the γ_1-curve in Q_0 have slopes $\approx \pm 2a\delta$ depending on whether we are working with Ω_0^+ or Ω_0^-.

Suppose that at step n, corresponding to every connected component ω of Ω_{n-1} we have a strip Q_ω foliated by integral curves of e_n. More precisely, $Q_\omega = \bigcup_{z\in\tilde{\omega}} \gamma_n(z)$ where $\gamma_n(z)$ is the integral curve segment of length $10b$ centered at z and $\tilde{\omega}$ is the $(Cb)^n$-neighborhood of ω in W_1. We think of γ_n as temporary stable manifolds of order n.

Let $\omega' \subset \omega$ be a component of Ω_n. We want to show that $Q_{\omega'}$ is well defined and is contained in Q_ω. For $z \in \omega'$, let $U_n(z)$ be the $(Cb)^n$-neighborhood of $\gamma_n(z)$ in $\mathbb{R}^2$. First we claim that e_{n+1} is defined on all of $U_n(z)$. Since $\left|Df_z^j\tau\right| \geq \kappa^j$, $\forall\, j \leq n+1$, it suffices, by Section 1.5, to check that $d(f^j\zeta, f^j z) < \lambda^j$, $\forall\, j \leq n+1$, $\forall\, \zeta \in U_n(z)$. Let ζ' be the point in $\gamma_n(z)$ nearest to ζ. Then

$$\begin{aligned} d(f^j\zeta, f^j z) &\leq d(f^j\zeta, f^j\zeta') + d(f^j\zeta', f^j z) \\ &< (Cb)^n \cdot 5^j + Cb^j < \lambda^j. \end{aligned}$$

Next we claim that $\gamma_{n+1}(z)$ is well defined and lies inside $U_n(z)$. We see this in two steps: first we use the Lipschitzness of e_n (Property (iii), Section 1.5) and a Gronwall type inequality to see that $e_n|U_n(z)$ can be mapped diffeomorphically onto $\partial/\partial y$ on $\mathbb{R}^2$ via a diffeomorphism Φ_n with $||D\Phi_n|| \leq e^{5\cdot 10b}$; then use $|e_{n+1} - e_n| < (Cb)^n$ and the "straightened out" coordinates of e_n to conclude that the Hausdorff distance between $\gamma_{n+1}(z)$ and $\gamma_n(z)$ is $\lesssim 10b \cdot (Cb)^n < \frac{1}{2}(Cb)^n$. Finally, observe that these arguments are easily extended to γ_{n+1}-curves through points in $(Cb)^{n+1}$-neighborhoods of ω', proving $Q_{\omega'} \subset Q_\omega$.

Taking limits of these "temporary stable manifolds", we obtain genuine stable manifolds for points in Ω_∞.

Lemma 2

(1) $\forall\, z \in \Omega_\infty$, *there is a* C^1 *curve* $\gamma_\infty(z)$ *of length* $10b$ *and centered at* z *s.t.* $\forall\, \zeta \in \gamma_\infty(z)$,
$$d(f^j\zeta, f^j z) < Cb^j \quad \forall\, j \geq 1;$$
(2) $\forall\, z, z' \in \Omega_\infty$, $z \neq z' \Longrightarrow \gamma_\infty(z) \cap \gamma_\infty(z') = \varnothing$;
(3) *if* $f^n\gamma_\infty(z) \cap \gamma_\infty(z') \neq \varnothing$, *then* $f^n\gamma_\infty(z) \subset \gamma_\infty(z')$.

Proof of (1). — Let z_n be the right end point of ω_{n-1}, the component of Ω_{n-1} containing z. Since $\bigcap_n \omega_n = \{z\}$ (reason: $|Df^n\tau| \geq \delta e^{c_2 n}$ on ω_{n-1} and $f^n\omega_{n-1}$ has length < 2), it follows from the estimates above that as $n \to \infty$, $\gamma_n(z_n)$ converges uniformly to a curve which we will call $\gamma_\infty(z)$. Because the e_n's have a uniform Lipschitz constant (Section 1.5, property (iii)), $\gamma_n(z_n)$ in fact converges in the C^1 sense to $\gamma_\infty(z)$. Thus the contractive estimates for $f^j \mid \gamma_n(z_n)$ carry over to $f^j \mid \gamma_\infty(z)$. $\square$

Before proving (2) and (3) we need to do some preparatory work. Consider a $C^2(b)$ curve γ lying in Q_0 and joining $\partial^s Q_0$, the two boundary components of Q_0 that are not part of W. For each connected component ω of Ω_{n-1}, we let γ_ω denote $\gamma \cap Q_\omega$. Note that every point in γ_ω is connected to some point in $\widetilde{\omega}$ by an integral curve γ_n, and also that if $\omega' \subset \Omega_n$ is contained in ω, then $\gamma_{\omega'} \subset \gamma_\omega$. For $z \in \gamma$ we will use $\tau(f^j z)$ to denote the unit tangent vector to $f^j\gamma$ at $f^j z$.

Sublemma 3. — *Let* γ *be as above. Then if* $\omega \subset \Omega_{n-1}$ *is s.t. part of* $f^n\omega$ *is free and intersects* $(-\delta, \delta) \times \mathbb{R}$, *then the binding point* $\widehat{z}$ *for* $f^n\omega$ *selected earlier is also suitable for* $f^n\gamma_{\omega'}$ *where* $\omega' = \omega \cap \Omega_n$ *(see Section 1.3 and Section 1.5 for the meaning of "suitable"). It follows from this that for all* $\omega \in \Omega_{n-1}$, $f^j \mid \gamma_\omega$, $j \leq n$, *has the bound/free estimates expressed in Section 1.3 and the distortion estimate in Section 1.4.*

Proof of Sublemma 3. — We fix $\zeta \in \gamma_\omega$ and investigate the suitability of $\widehat{z}$ as a binding point for $(f^n\zeta, Df^n_\zeta\tau(\zeta))$. Let $z \in \widetilde{\omega}$ be s.t. $\zeta \in \gamma_n(z)$. Then $d(f^n z, f^n\zeta) < Cb^n$. This cannot jeopardize the generalized tangential position part of the requirement since Cb^n is totally insignificant compared to $d(f^n z, \widehat{z})$, which is $> e^{-\alpha n}$. As to the angle part of the requirement, write
$$\sphericalangle(Df^n_z\tau(z), Df^n_\zeta\tau(\zeta)) \leq \sphericalangle(Df^n_z\tau(z), Df^n_z\tau(\zeta)) + \sphericalangle(Df^n_z\tau(\zeta), Df^n_\zeta\tau(\zeta)).$$
The first term is $< 20b \cdot Cb^n$ because both $\tau(z)$ and $\tau(\zeta)$ have slopes $< 10b$, $\|Df^n_z\| > 1$, and both $\tau(z)$ and $\tau(\zeta)$ make angles $\approx 2a\delta$ with $e_n(z)$. The second term is $< Cb^{n/4}$ by the matrix perturbation lemma in Section 1.5. The difference between $\tau(f^n z)$ and $\tau(f^n\zeta)$, therefore, are insignificant relative to $(2a \pm 1) \cdot d(f^n z, \widehat{z})$, the size of the angle they are supposed to make with the relevant contracting field about $\widehat{z}$.

The estimates in Section 1.3 depend on the pair (ζ, τ), $\zeta \in \gamma_\omega$ being "controlled". (For the precise definition see 1.4.2 and 1.5.1 of **[BY]**.) The distortion estimate in Section 1.4 holds for $C^2(b)$ segments all of whose points and tangent vectors are controlled. □

Proof of Lemma 2 (continued). — We prove (3); the proof of (2) is similar. We will try to derive a contradiction assuming $f^n\gamma_\infty(z) \not\subset \gamma_\infty(z')$. Let N be a sufficiently large number to be specified. Let η and η' be points in $f^n\gamma_\infty(z)$ and $\gamma_\infty(z')$ respectively s.t.

(i) η and η' are joined by a horizontal line segment $\gamma \subset Q_0$ and

(ii) $\eta \in \partial Q_\omega$ where ω is the component of Ω_{N-1} containing z'. Since η and η' lie in some γ_ω, Sublemma 3 tells us that if $f^N\omega$ is free (our 1st requirement on N), then

$$\left|f^N\eta - f^N\eta'\right| \geq e^{cN}\left|\eta - \eta'\right| \gtrsim e^{cN} \cdot \frac{1}{2}(Cb)^N.$$

On the other hand, if $q \in f^n\gamma_\infty(z) \cap \gamma_\infty(z')$, then

$$\begin{aligned}\left|f^N\eta - f^N\eta'\right| &\leq \left|f^N\eta - f^Nq\right| + \left|f^Nq - f^N\eta'\right| \\ &< \frac{Cb^{N+n}}{(b/5)^n} + Cb^N = (5^n+1)Cb^N.\end{aligned}$$

These two estimates of $\left|f^N\eta - f^N\eta'\right|$ are clearly incompatible for $N \gg n$. □

3.4. Definition of Λ and return times. — We now specify the two families Γ^u and Γ^s that define Λ.

Definition of Γ^s. — We let $\Gamma^s = \{\gamma_\infty(z) \colon z \in \Omega_\infty\}$ where $\gamma_\infty(\cdot)$ is as in Lemma 2.

Definition of Γ^u. — We let $\widetilde{\Gamma}^u = \{\gamma \subset W : \gamma$ is a $C^2(b)$ segment connecting the two components of $\partial^s Q_0\}$, and let $\Gamma^u = \{\gamma \colon \gamma$ is the pointwise limit of a sequence in $\widetilde{\Gamma}^u\}$.

Remark 2. — (1) We have not proved that the curves in Γ^u are pairwise disjoint. However, since every $\gamma \in \Gamma^u$ is the monotone limit of curves in $\widetilde{\Gamma}^u$, there are at most countably many pairs that intersect. It is easy to see that they play no role.

(2) Without further analysis, we also cannot conclude that $\gamma \in \Gamma^u$ is better than $C^{1+1}(b)$, since they are uniform limits of $C^2(b)$ curves. This also is inconsequential.

Having completed the definition of Λ, we now proceed to define the s-sublattices that make up Λ and their return times R. Let us remind ourselves again that in actuality we are interested in the set $\Lambda^+ \cup \Lambda^-$, where $\Lambda^\pm$ correspond to the lattices we have constructed near $\Omega_0^\pm \times [-b, b]$, Ω_0^+ and Ω_0^- being the two outermost intervals in the partition $\mathcal{P}$ introduced in Section 1.4. When we speak about return times, we are referring to return times from the set $\Lambda^+ \cup \Lambda^-$ to itself, *i.e.* a point in Λ^+ may return to Λ^+ or Λ^-. To keep the notations simple we will continue to write just "Λ".

We stipulate ahead of time that $\forall z \in \Lambda$, $R(z) = R(z') \ \forall z' \in \gamma^s(z)$, so R need only be defined on $\Lambda \cap \Omega_0$. We will construct partitions on subsets of Ω_0 and use 1-dimensional language. For example, $f^n x = y$ for $x, y \in \Lambda \cap \Omega_0$ means that $f^n x \in \gamma^s(y)$. Similarly, for subsegments $\omega, \omega' \subset \Omega_0$, $f^n\omega = \omega'$ means that $f^n\omega \cap \Lambda$, when slid along γ^s-curves to Ω_0, gives exactly $\omega' \cap \Lambda$. (We caution that "$f^n\omega = \omega'$" does not imply $f^n(\omega \cap \Lambda) = \omega' \cap \Lambda$!) For $\omega \subset \Omega_{n-1}$, $\mathcal{P} \mid f^n\omega$ refers to $\mathcal{P}_{[\tilde{z}]}$ where $\tilde{z}$ is the binding point for $f^n\omega$ selected earlier.

We will construct below sets $\widetilde{\Omega}_n \subset \Omega_n$ and partitions $\widetilde{\mathcal{P}}_n$ on $\widetilde{\Omega}_n$ so that $\widetilde{\Omega}_0 \supset \widetilde{\Omega}_1 \supset \widetilde{\Omega}_2 \supset \dots$ and $z \in \widetilde{\Omega}_{n-1} - \widetilde{\Omega}_n$ iff $R(z) = n$. As usual, we think of points belonging to the same element of $\widetilde{\mathcal{P}}_n$ as having indistinguishable trajectories up to time n. We augment $\mathcal{P}$ defined in Section 1.4 to $\mathcal{P} = \{\omega \in \text{ original } \mathcal{P}\} \cup \{[-1, -\delta), (\delta, 1]\}$, and let $\widehat{\mathcal{P}}$ be the partition on $\Omega_0 - \Omega_\infty$ dividing this set into connected components. The symbol "$\vee$" refers to the join of two partitions, *i.e.* $\mathcal{A} \vee \mathcal{B} \equiv \{A \cap B \colon A \in \mathcal{A}, B \in \mathcal{B}\}$.

An interval $\omega \subset \Omega_n$ is said to make a *regular return* to Ω_0 at time n if

(i) all of $f^n\omega$ is free;
(ii) $f^n\omega \supset 3\Omega_0$.

Rules for defining $\widetilde{\Omega}_n, \widetilde{\mathcal{P}}_n$ *and* R:

(0) $\widetilde{\Omega}_0 = \Omega_0$, $\widetilde{\mathcal{P}}_0 = \{\widetilde{\Omega}_0\}$.
Consider $\omega \in \widetilde{\mathcal{P}}_{n-1}$.

(1) If ω does not make a regular return to Ω_0 at time n, put $\omega \cap \Omega_n$ into $\widetilde{\Omega}_n$, and let
$$\widetilde{\mathcal{P}}_n \left|(\omega \cap \Omega_n) = (f^{-n}\mathcal{P})\right|(\omega \cap \Omega_n)$$
with the usual adjoining of end intervals (this is always done with or without our saying so explicitly).

(2) If ω makes a regular return at time n, we put $\omega' \equiv (\omega - f^{-n}\Omega_\infty) \cap \Omega_n$ in $\widetilde{\Omega}_n$, and let $\widetilde{\mathcal{P}}_n \left|\widetilde{\omega} = (f^{-n}\mathcal{P} \vee f^{-n}\widehat{\mathcal{P}})\right|\omega'$. For $z \in \omega$ s.t. $f^n z \in \Omega_\infty$, we define $R(z) = n$.

(3) We require that $R \geq n_0$ for some n_0 to be specified in Section 6.1. To comply with this, if $\omega \in \widetilde{\mathcal{P}}_{n-1}$ makes a regular return at time n with $n < n_0$, then we treat ω according to Rule (1) and not Rule (2).

(4) For $z \in \bigcap_n \widetilde{\Omega}_n$, set $R(z) = \infty$.

Remark 3. — We digress to make the following adjustments in our definitions of $\widetilde{\mathcal{P}}_n$; they will simplify the proofs in Section 4. Let us say that a $C^2(b)$ segment $\gamma \subset (-\delta, \delta)$ is of "full length" if $\exists \, \nu, j$ s.t. $\gamma \cap I_{\nu j} \neq \varnothing$ and $\ell(\gamma) \approx \ell(I_{\nu j})$. Recall that in Section 3.2 we made sure that when something is deleted from $f^n\omega$, $\omega \in \Omega_{n-1}$, no "short" segment is left behind. We wish to do the same for $f^n\omega$ for every $\omega \in \widetilde{\mathcal{P}}_{n-1}$. For definiteness suppose that $\omega \in \widetilde{\mathcal{P}}_{n-1}$ was created at step $k \leq n-1$, and that not all of ω will remain in Ω_n. We distinguish between the cases where $f^k\omega \approx$ some $I_{\nu j}$ and where $f^k\omega$ is a gap of Λ.

If $f^k\omega$ is a gap of Λ, then $\mathrm{dist}(f^n(\partial\omega), \mathcal{C}) \geq \delta e^{-\alpha(n-k)} \geq 10\delta e^{-\alpha n}$, so the end points of $f^n\omega$ straddle the forbidden interval $(-\delta e^{-\alpha n}, \delta e^{-\alpha n})$ by wide margins and no problem will arise.

Suppose $f^k\omega \approx$ some $I_{\nu j}$. Since deletions occur only at free returns, we have $\ell(f^n\omega) \gg \delta e^{-\alpha n}$. The only problematic scenario is when a tiny part of $f^n\omega$ sticks out, say, to the left of $(-\delta e^{-\alpha n}, \delta e^{-\alpha n})$. If this awkward bit remains in Ω_n, then there must be something in Ω_{n-1} that is mapped by f^n to the left of it. The reasoning of the last paragraph rules out the possibility that the left end point of $f^n\omega$ is a limit of infinitely many small segments coming from the gaps of Λ. Thus $\exists\, \omega' \in \widetilde{\mathcal{P}}_{n-1}$ that shares this relevant end point with ω. Hence $\exists\, \omega'' \in \widetilde{\mathcal{P}}_k$ with this property and $f^k\omega'' \approx$ some $I_{\nu' j'}$. We now retroactively move the boundary between ω and ω'' so that the awkward bit in question belongs to the image of ω' or ω''. It is easy to see that no boundary is moved more than once, for a gap beween the adjacent elements appear immediately thereafter.

It is now clear what the sublattices in Propostion A (1) are: each Λ_i is an s-sublattice corresponding to a subset of $\Omega_0 \cap \Lambda$ of the form $f^{-n}\Omega_\infty \cap \Lambda \cap \omega$, $\omega \in \widetilde{\mathcal{P}}_{n-1}$ making a regular return at time n. Note that if Λ_i is one of the s-sublattices, and $\gamma = Q_i \cap \gamma^u$, $\gamma^u \in \Gamma^u$, $Q_i =$ the rectangle spanned by Λ_i, then $f^n\gamma \in \Gamma^u$. To see this, first assume $\gamma \subset W$. Then $f^n\gamma$ is $C^2(b)$ because it is free; hence it is in $\widetilde{\Gamma}^u$. This property clearly passes on to curves in Γ^u, proving $f^n\Lambda_i \subset \Lambda$. Note also that the hyperbolic estimates in Propostion A are simply Estimate III in Section 1.3 and Lemma 2 (1).

To complete our objective of proving Assertions (1) and (2) in Propostion A then, it remains only to show that $f^{R_i}\Lambda_i$ is a u-sublattice. This requires proving that the Cantor set $f^{R_i}\Lambda_i$ somehow matches completely with Λ in the horizontal direction. We claim that this is a consequence of our construction but defer the proof to Section 3.5.

Remark 4. — We have not proved that $R(z) < \infty$ for every $z \in \Lambda$. Indeed, the assertion in Propostion A (1) that $\Lambda = \bigcup \Lambda_i$ is inaccurate and should be ammended to read "for every $\gamma^u \in \Gamma^u$, $\mathrm{Leb}((\Lambda - \bigcup \Lambda_i) \cap \gamma^u) = 0$". That $R < \infty$ a.e. on $\Lambda \cap \gamma^u$ will follow from the Main Lemma in Section 4.

3.5. Matching of Cantor sets. — To complete the proof of Propostion A (1), we need to show that whenever Rule (2) in the previous subsection is applied,

$$f^n(\omega \cap \Omega_\infty) \supset \Omega_\infty.$$

We formulate this as

Lemma 3. — Let $\omega \in \Omega_{n-1}$ be s.t. $f^n\omega$ crosses Q_0 completely. Then $\forall\, z \in \Lambda$, $\exists\, z' \in \omega \cap \Lambda$ s.t. $f^n z' \in \gamma^s(z)$.

Let us first explain the central idea of the proof assuming that f is a 1-dimensional map. Given $z \in \Omega_\infty$, there is (by hypothesis) $z' \in \omega$ with $f^n z' = z$; what is at issue is whether $z' \in \Omega_\infty$. First, $z' \in \Omega_n$ because $\omega \subset \Omega_{n-1}$ and $f^n z' \in \Omega_0$. It suffices therefore to show that $|f^{j+n}z'| > 2\delta e^{-(j+n)\alpha}$ $\forall j > 0$. This is true because $|f^{j+n}z'| = |f^j z| > \delta e^{-\alpha j}$, which is $> 10\delta e^{-\alpha(j+n)}$ since $e^{\alpha n} \geq 10$.

For the 2-dimensional situation at hand, what complicates matters is that different layers of W require different binding points, and that the binding point at step $n+j$ for $f^{n+j}z'$ may not be vertically aligned with the binding point at step j for $f^j z$. Our aim in this subsection is to dispel with these technicalities so that the 1-d argument prevails.

Lemma 3'. — *Let ω be a connected component of Ω_{n-1}, and suppose that $f^n Q_\omega$ crosses Q_0 completely in the horizontal direction. Then $\forall j \geq 1$, if ω_j is a component of Ω_j, then there is a component ω_{n+j} of Ω_{n+j} s.t. $Q_{\omega_j} \cap f^n Q_\omega \subset f^n Q_{\omega_{n+j}}$.*

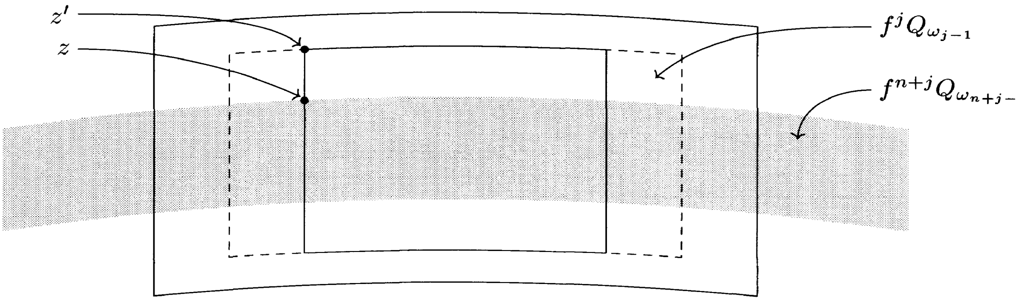


FIGURE 2

In this subsection we will regard Q_ω as foliated by temporary stable curves through ω, ignoring the slight discrepancies between the temporary curves of various generations or their slightly different domains of definition. Those matters were dealt with in Section 3.3. We remark that if Lemma 3$'$ holds, then it will follow that $(\bigcup_{\omega_j} Q_{\omega_j}) \cap f^n Q_\omega \subset \bigcup_{\omega_{n+j}} f^n Q_{\omega_{n+j}}$ for all $j \geq 1$. Taking the limit as $j \to \infty$, we will obtain $(\bigcup_{z \in \Omega_\infty} \gamma^s(z)) \cap f^n Q_\omega \subset \bigcup_{z \in \Omega_\infty} f^n \gamma^s(z)$, which gives Lemma 3.

Proof of Lemma 3'. — Let n and ω be fixed, and assume the conclusion of Lemma 3$'$ for all components of Ω_{j-1}. We pick one ω_{j-1}, and let ω_{n+j-1} be as in the lemma, see Figure 2. We will examine what is deleted from $f^j Q_{\omega_{j-1}}$ at step j versus what is deleted from $f^{n+j} Q_{\omega_{n+j-1}}$ at step $n+j$.

Let $z \in f^{n+j}\omega_{n+j-1} \cap f^j Q_{\omega_{j-1}}$ be such that z is deleted at step $n+j$, and let $z' \in \gamma_j^s(z) \cap f^j \omega_{j-1}$. We will show that z' is deleted at step j. The notations and results of Section 2 will be used heavily in the next few lines. First if $d_\gamma(z', \mathcal{C}) < \delta e^{-\alpha j}$, we are done. Suppose not. Then since $|z - z'| < (Cb)^j$,

$$|z - \tilde{z}(z')| \geq |z' - \tilde{z}(z')| - |z - z'| \geq \frac{9}{10}\delta e^{-\alpha j},$$

and z is in tangential position wrt $\tilde{z}(z')$. But we also have

$$|z - \tilde{z}(z)| \leq 2\delta e^{-(n+j)\alpha} < \frac{2}{10}\delta e^{-\alpha j},$$

and this is incompatible with our estimate on $|z - \tilde{z}(z')|$. □

4. Return time estimates

The goal of this section is to prove Assertions (3) and (4) in Propostion A.

4.1. Statement of lemmas and ideas of proofs. — Let $\Omega_0 \subset W_1$ be as in Section 3, and recall that there are sets $\Omega_0 = \widetilde{\Omega}_0 \supset \widetilde{\Omega}_1 \supset \widetilde{\Omega}_2 \supset \dots$ and partitions $\widetilde{\mathcal{P}}_n$ on $\widetilde{\Omega}_n$ so that for all $z \in \Omega_\infty$, $z \in \widetilde{\Omega}_n$ iff $R(z) > n$, and points in the same element of $\widetilde{\mathcal{P}}_n$ are viewed as having the same itinerary up to time n. (See Section 3.2 and Section 3.4.) Let $|\cdot|$ denote the Lebesgue measure on γ^u-curves.

Lemma 4. — $|\Omega_\infty| > 0$.

Proof. — This is a 1-d argument using estimates in **[BC1]**. In the construction of $\{\Omega_n\}$, let $\omega \subset \Omega_k$ be a component that is formed at step k, and suppose that some part of it will be deleted at step n. Since $f^k\omega \supset$ some $I_{\nu j}$, we are guaranteed that $|f^n\omega| \geq \delta^{3\beta}e^{-3\alpha\beta k} \geq \delta^{3\beta}e^{-3\alpha\beta n}$. But the subsegment of $f^n\omega$ to be deleted has length $\leq 4\delta e^{-\alpha n}$. Taking distorsion into consideration when pulling back to Ω_0, we have that

$$\frac{|\Omega_{n-1} - \Omega_n|}{|\Omega_{n-1}|} \leq C_1\delta^{(1-3\beta)}e^{-\alpha(1-3\beta)n}$$

and the statement of the lemma follows since

$$\prod_{n=M_1}^{\infty}\left(1 - C_1\delta^{(1-3\beta)}e^{-\alpha(1-3\beta)n}\right) > 0,$$

where M_1 is the minimum time for a point in $(-\delta, \delta)$ to return to $(-\delta, \delta)$; see Section 3.2. □

***Lemma 5* (Main Lemma).** — $\exists\, C_0 > 0$ *and* $\theta_0 < 1$ *s.t.*

$$\left|\widetilde{\Omega}_n\right| < C_0\theta_0^n \qquad \forall\, n \geq 1.$$

Remark 5. — We have stated lemmas 4 and 5 for Ω_n on W_1, but the corresponding statements are true for every $\gamma^u \in \Gamma^u$ with uniform estimates (independent of γ^u). The proofs are in fact identical through the use of Lemma 2. A related fact that will not be needed till later is in fact the absolute continuity of $\{\gamma^s\}$ as a "foliation"; see Sublemma 10 in Section 5. It says in particular that $\exists\, C > 0$ s.t. for all γ, $\gamma' \in \Gamma^u$, if $\Psi : \gamma \cap \Lambda \mapsto \gamma'$ is defined by $\Psi(z) = \gamma(z) \cap \gamma'$, then $|\Psi(A)| \leq C|A|$ for every Borel subset A of $\Lambda \cap \gamma$.

We postpone the proof of Lemma 5 for later, but use instead the remainder of this subsection to discuss the main ideas behind this tail estimate for R. Consider a segment $\omega \subset \Omega_k$.

(1) It is easy to see that once $f^n\omega$ becomes sufficiently long, then it will make a regular return to Ω_0 within a finite number of iterates. Our situation is as follows: as we iterate f, ω grows in length — except when it comes near $\mathcal{C}$, at which time it may lose a piece in the middle and it may get subdivided into $I_{\nu j}$'s for distortion control; these components are then iterated individually. An unfortunate component of ω may get cut faster than it has the chance to grow, but our contention is that because of the estimates in Section 1.3 the *general tendency* is for a component to grow long.

(2) When a regular return occurs, small pieces corresponding to the gaps of Λ are created, and these small pieces are handled individually as they move on. We must therefore carry out the large deviation estimate in (1) simultaneously for the entire collection of gaps; such an estimate will involve the distribution of gap sizes.

(3) As has already been suggested in (2), it is not quite the end of the story when a component of ω grows long, for at regular returns only a (fixed) percentage of the long segment gets absorbed into the Cantor set Λ. To estimate distribution of return times we must estimate the frequencies with which the components containing typical points makes regular returns.

These ideas are made rigorous in Sections 4.2, 4.3 and 4.4. We remark also that (1) is essentially dealt with in **[BC2]** in the slightly different context of parameter exclusions, and that we learned some of the estimates for (2) and (3) from **[C]**.

4.2. Growth of components of a segment to a fixed size: a large deviation estimate. — In this and the next subsections we will be studying the time evolution of a curve γ which is contained in $f^j\widetilde{\Omega}_j$ for some $j \geq 0$. It is convenient to think of points as being in γ at time 0, so let us introduce the following notations: $\Omega_k^{(j)} \equiv \{z \in \gamma \colon f^{-j}z \in \Omega_{j+k}\}$, similarly for $\widetilde{\Omega}_k^{(j)}$, and $\widetilde{\mathcal{P}}_k^{(j)}(z) \equiv \{\eta \in \gamma \colon f^{-j}\eta \in \widetilde{\mathcal{P}}_{j+k}(f^{-j}z)\}$.

We will also use the following language. For $z \in \gamma$, we say that z makes an *essential free return* ($\equiv$ e.f.r.) to $(-\delta, \delta) \times \mathbb{R}$ at time k if $z \in \widetilde{\Omega}_{k-1}^{(j)}$ and $f^k\widetilde{\mathcal{P}}_{k-1}^{(j)}(z)$ is free and contains some $I_{\nu j}$. We say z makes a *regular return* to Ω_0 at time k if $z \in \widetilde{\Omega}_{k-1}^{(j)}$ and $f^k\widetilde{\mathcal{P}}_{k-1}^{(j)}(z)$ makes a regular return. We define the stopping time

$$E(z) \equiv \text{the smallest } k \in \mathbb{Z}^+ \text{ s.t. either } z \notin \Omega_k^{(j)} \text{ or } z \text{ makes a regular return at time } k$$

and let

$$\gamma_n = \{z \in \gamma : E(z) > n\}.$$

***Sublemma 4* (*cf.* Section 2, [BC2]).** — *There exist $D_1' > 0$ and $\theta_1' < 1$ for which the following holds. Let $\gamma = f^j\omega$ for some $\omega \in \widetilde{\mathcal{P}}_j$, and let $n \in \mathbb{Z}^+$. We assume that γ is free and is $\approx$ some $I_{r,\ell}$ with $|r| \le n/6$. Then*

$$|\gamma_n| \le D_1' \theta_1'^{\,n} |\gamma|.$$

Proof. — We will prove that

$$(*) \qquad |\gamma_n| \le D_1' e^{-\frac{1}{6}n + \frac{9}{10}|r|} |\gamma|$$

for some D_1' independent of γ or n. This implies the Sublemma immediately: for $|r| \le n/6$, $D_1' e^{-\frac{1}{6}n+\frac{9}{10}|r|} \le D_1' e^{-\frac{1}{6}n+\frac{9}{10}\frac{n}{6}} = D_1' e^{-\frac{1}{60}n}$, so it suffices to take $\theta_1' = e^{-1/60}$.

Consider $z \in \gamma$ with $E(z) > n$, and suppose that z makes exactly s e.f.r.'s in the first n iterates, at times $0 = t_0 < t_1 < \cdots < t_s \le n$. It follows from Remark 2 in Section 3.4 that $f^{t_i}\widetilde{\mathcal{P}}_{t_i}^{(j)}(z) \approx$ some $I_{r_i\ell_i}$ for each i. A slightly extended version of our estimates in Section 1.3 gives for all $i < s$:

$$t_{i+1} - t_i \le 4\,|r_i| \qquad \text{and} \qquad \left|f^{t_{i+1}-t_i} I_{r_i\ell_i}\right| \ge e^{-3\beta|r_i|}.$$

This second inequality can be used to estimate the fraction

$$\varphi(r_1, \ldots, r_s) \equiv \frac{1}{|\gamma|} \cdot |\{z \in \gamma \colon \mathcal{F}(z) = (r_1, \ldots, r_s)\}|$$

where $\mathcal{F}(z)$ denotes the r_i-locations of the e.f.r.'s of z. Letting C_1 be the distortion constant in Section 1.4 and writing $r_0 = r$, we have

$$\begin{aligned}\varphi(r_1, \ldots, r_s) &< C_1^s \prod_{i=1}^{s} \exp\{-|r_i| + 3\beta\,|r_{i-1}|\} \\ &< C_1^s \exp\Big\{-\frac{7}{8}\sum_{i=1}^{s}|r_i| + 3\beta|r|\Big\}.\end{aligned}$$

Next we make, for fixed s and R, the purely combinatorial estimate on the number of all possible s-tuples $(r_1, \ldots, r_s)$, $r_i \in \mathbb{Z}$, with $\sum_{i=1}^s |r_i| = R$. This is clearly $< 2^s\binom{R+s-1}{s-1}$. For us, since the time between consecutive e.f.r.'s is $> \Delta \equiv \log(1/\delta)$, the number of feasible $(r_1, \ldots, r_s)$ as locations of e.f.r.'s is in fact $\le 2^{R/\Delta} \cdot \binom{R+(R/\Delta)-1}{R/\Delta}$, which by Sterling's formula is $< 2^{R/\Delta}(1+\sigma(\delta))^R$ with $\sigma(\delta) \to 0$ as $\delta \to 0$.

Let

$$A_{s,R}^n \equiv \big\{z \in \gamma_n \colon z \text{ makes exactly } s \text{ e.f.r.'s up to time } n \text{ and } \textstyle\sum_{i=1}^s |r_i| = R\big\}.$$

We may then estimate $|\gamma_n|$ by

$$|\gamma_n| = \sum_{\substack{\text{all relevant} \\ s,R}} |A_{s,R}^n| \le \sum_{R=(n/4)-|r|}^{\infty} \sum_{s=1}^{R/\Delta} \begin{pmatrix} \#(r_1, \ldots, r_s) \\ \text{with } \sum_1^s |r_i| = R \end{pmatrix} \cdot \varphi(r_1, \ldots, r_s) \cdot |\gamma|.$$

The lower limit of summation for R comes from the fact that $|r| + 4R$ must be $> n$, otherwise the $(s+1)^{st}$ e.f.r. or a regular return, whichever happens first, would have taken place by time n. We do not need to concern ourselves with $s = 0$ because γ must make an e.f.r. by time $n/2$ (because $e^{-n/6}e^{c_1 n/2} \gg 1$). Note that this is an overestimate also in the sense that some of the $(r_1, \ldots, r_s)$-configurations are forbidden due to the $I_{r_i \ell_i}$'s being too close to $\mathcal{C}$.

Plugging our earlier estimates into this last inequality, we obtain

$$|\gamma_n| \leq C \sum_{R=(n/4)-|r|}^{\infty} 2^{R/\Delta}(1+\sigma(\delta))^R \cdot C_1^{R/\Delta} e^{-\frac{7}{8}R+3\beta|r|} \cdot |\gamma|,$$

which is less than the right side of $(*)$ provided δ is sufficiently small. □

We now re-state Sublemma 4 in anticipation of how it will be used.

Corollary to Sublemma 4. — *There exist $D_1 > 0$ and $\theta_1 < 1$ for which the following holds. Let γ be contained in a free $(C^2(b))$ segment of W. We assume that either*

(i) *$\gamma = f^j\omega$ for some $\omega \in \widetilde{\mathcal{P}}_j$, γ need not be of "full length"; or*

(ii) *$\gamma = \bigcup_i f^j\omega_i$ where for each i, $\omega_i \in \widetilde{\mathcal{P}}_j$ and $f^j\omega_i \approx$ some $I_{r_i\ell_i}$.*

Then

$$|\gamma_n| \leq D_1\theta_1^n \qquad \forall n \geq 1.$$

Proof. — First we prove (ii). For fixed n, we have by Sublemma 4 that

$$\left|\{z \in \gamma - \left(-e^{-n/6}, e^{-n/6}\right) : E(z) > n\}\right| \leq D_1'\theta_1'^n|\gamma| \leq D_1'\theta_1'^n,$$

and also that

$$\left|\gamma \cap \left(-e^{-n/6}, e^{-n/6}\right)\right| \leq 2e^{-n/6}.$$

To prove (i), fix n and observe as above that we may assume $|\gamma| \geq e^{-n/6}$, and that γ makes an e.f.r. at time $j_0 < n/2$. Suppose that this is not a regular return, and let $\gamma' = f^{j_0}\gamma$. Then $|\gamma_n| \leq |f^{j_0}\gamma_n| \leq |\{z \in \gamma' : E(z) > n/2\}|$, and this last quantity is estimated as in the proof of (ii). □

4.3. Growth of "gaps" to a fixed size. — First we prove a sublemma about the distribution of gap sizes.

Sublemma 5. — *There exist $C > 0$ and $\sigma > 0$ s.t. for all $\gamma^u \in \Gamma^u$, if $\mathcal{G} = \{$components of $\gamma^u - \Lambda\}$, then*

$$\sum_{\gamma \in \mathcal{G}:\ |\gamma| \leq \ell} |\gamma| \leq C\ell^{\sigma}.$$

Proof. — In view of Sublemma 3, it suffices to consider $\gamma^u = \Omega_0$.

Observe first that all the gaps of Ω_∞ created at step n have length

$$> C_1{}^{-1}4\delta e^{-\alpha n}e^{-c_1 n}.$$

(If $f^n\omega$ partially crosses $(-\delta e^{-\alpha n}, \delta e^{-\alpha n})$, then the part deleted is attached to an earlier gap, making it even bigger.) Given ℓ, let N_0 be s.t. $\ell \approx e^{-(\alpha+c_1)N_0}$. Then

$$\sum_{\substack{\gamma\in\mathcal{G}\\|\gamma|\le\ell}} |\gamma| \le \sum_{n\ge N_0} \sum_{\substack{\omega=\text{comp.}\\ \text{of }\Omega_{n-1}}} C_1 4e^{-(1-3\beta)\alpha n}|\omega|$$

(*cf.* Lemma 4). Thus

$$\sum_{\substack{\gamma\in\mathcal{G}\\|\gamma|\le\ell}} |\gamma| \le Ce^{-\alpha N_0(1-3\beta)} \le C\ell^{\sigma},$$

some $\sigma > 0$. □

Sublemma 6. — *Let $\Lambda^c \equiv \Omega_0 - \Lambda$, and for $z \in \Lambda^c$, define $E(z)$ as in Section 4.2 with $\gamma =$ the component of Λ^c containing z. Let $\Lambda^c_n \equiv \{z \in \Lambda^c\colon E(z) > n\}$. Then $\exists\, D_2 > 0$ and $\theta_2 < 1$ s.t.*

$$|\Lambda^c_n| \le D_2\theta_2^n \qquad \forall\, n \ge 1.$$

Proof. — Let $\mathcal{G}$ be the set of all components of Λ^c. For each $n \in \mathbb{Z}^+$, let $\mathcal{G}'_n \equiv \{\omega \in \mathcal{G}\colon |\omega| \le D_1\theta_1^n\}$ where D_1 and θ_1 are as in the Corollary to Sublemma 4, and let $\mathcal{G}''_n = \mathcal{G} - \mathcal{G}'_n$.

By Sublemma 5,

$$\sum_{\omega\in\mathcal{G}'_n} |\omega| \le C(D_1\theta_1^n)^{\sigma} \equiv D'_2\theta_1^{\sigma n}.$$

For $\omega \in \mathcal{G}''_n$, we know from the Corollary to Sublemma 4 that $\omega_n := \{z \in \omega : E(z) > n\}$ has length $|\omega_n| \le D_1\theta_1^n$, and from Sublemma 5 that if

$$N_k = \#\{\omega\colon D_1\theta_1^k \le |\omega| < D_1\theta_1^{k-1}\},$$

then

$$N_k \le \frac{C(D_1\theta_1^{k-1})^{\sigma}}{D_1\theta_1^k} \equiv C'\theta_1^{(k-1)\sigma-k}.$$

Thus

$$\sum_{\omega\in\mathcal{G}''_n} |\omega_n| \le D_1\theta_1^n \cdot \sum_{k\le n} N_k \le D''_2\theta_1^{\sigma n}.$$

□

4.4. Frequencies of regular returns and Proof of Lemma 5. — We define a sequence of stopping times $T_0 < T_1 < \cdots$ on subsets of Ω as follows. Let $T_0 \equiv 0$, and assuming that $T_{k-1}(z)$ is defined, let $T_k(z)$ be the smallest $j > T_{k-1}$ s.t. $\widetilde{\mathcal{P}}_{j-1}(z)$ makes a regular return to Ω_0 at time j. Let $\Theta_k = \{z \in \Omega_0 : T_k(z) \text{ is defined}\}$. It follows from the Corollary to Sublemma 4 that $\Theta_k \supset \Omega_\infty$ a.e. for each k. Observe that Θ_k is the disjoint union of a countable number of segments $\{\omega\}$ with the property that each ω is an element of some $\widetilde{\mathcal{P}}_{j-1}$, $T_k|\omega \equiv j$, and a certain proportion of ω is absorbed back into Λ at time j. This is to say, $\exists\, \varepsilon_0 > 0$ such that for all $\omega \subset \Theta_k$ as above,

$$\frac{|\omega \cap \{R = T_k\}|}{|\omega|} \geq \varepsilon_0.$$

This implies inductively that for every k,

$$|\{z \in \Theta_k : R(z) > T_k\}| \leq (1 - \varepsilon_0)^k.$$

Let $\varepsilon_1 > 0$ be a small number to be determined. Then for all n,

$$\widetilde{\Omega}_n \subset \{z \in \widetilde{\Omega}_n : T_{[\varepsilon_1 n]}(z) > n\} \cup \{z \in \Theta_{[\varepsilon_1 n]} : R(z) > T_{[\varepsilon_1 n]}(z)\}.$$

The measure of the second set on the right has already been estimated. It remains therefore to prove

Sublemma 7. — $\exists\, D_3 > 0$, $\theta_3 < 1$, *and* $\varepsilon_1 > 0$ *such that*

$$|\{z \in \widetilde{\Omega}_n : T_{[\varepsilon_1 n]}(z) > n\}| < D_3 \theta_3^n \qquad \forall\, n \geq 1.$$

Proof. — Let $1 \leq n_1 < n_2 < \cdots < n_\ell \leq n$ be fixed for the time being. For $k \leq n$, we define $A_k \equiv A_k(n_1, \ldots, n_\ell)$ to be

$$A_k \equiv \{z \in \widetilde{\Omega}_k : \text{the regular return times of } z \text{ up to time } k \text{ are exactly those } n_i\text{'s with } n_i \leq k\},$$

and we estimate $|A_n|$ following these steps:

(i) $|A_{n_1 - 1}| \leq D_1 \theta_1^{n_1 - 1}$ by Sublemma 4 applied to $\gamma = \Omega_0$.

(ii) Note that A_{n_1-1} is a union of elements of $\widetilde{\mathcal{P}}_{n_1-1}$, and that A_{n_1-1} could be seen as

$$A_{n_1} = \{\omega = \omega' \cup \omega'' : \omega \in \widetilde{\mathcal{P}}_{n_1-1} \mid A_{n_1-1},\ \omega \text{ making a regular return at time } n_1\},$$

where $\omega' = \omega \cap f^{-n_1}\Lambda^c$ and $\omega'' = (\omega - f^{-n_1}\Omega_0) \cap \Omega_{n_1}$.

(iii) Using Sublemma 4 to deal with ω' and Sublemma 6 to deal with ω'' we obtain

$$\frac{|A_{n_2-1}|}{|A_{n_1-1}|} \leq \frac{D_3' \theta_3'^{\,n_2 - n_1 - 1}}{|\Omega_0|}$$

for some D_3' and θ_3' independent of the n_i's.

(iv) Proceeding inductively, we obtain

$$\begin{aligned} |A_n| &= \frac{|A_n|}{|A_{n_\ell - 1}|} \cdot \frac{|A_{n_\ell - 1}|}{|A_{n_{\ell-1} - 1}|} \cdots \frac{|A_{n_2-1}|}{|A_{n_1-1}|} \cdot |A_{n_1-1}| \\ &\leq \left(\frac{D_3'}{|\Omega_0|\theta_3'}\right)^{\ell} \theta_3'^{\,n}. \end{aligned}$$

We may now choose $\varepsilon_1 > 0$ small enough that

$$\left(\frac{D'_3}{|\Omega_0|\theta'_3}\right)^{\varepsilon_1} \cdot \theta'_3 \equiv \theta''_3 < 1,$$

and conclude that

$$\begin{aligned}\left|\{z \in \widetilde{\Omega}_n \colon T_{[\varepsilon_1 n]} > n\}\right| &= \sum_{\ell=0}^{[\varepsilon_1 n]} \sum_{\substack{(n_1,\dots,n_\ell)\colon \\ 1\le n_1<\cdots<n_\ell\le n}} |A_n(n_1,\dots,n_\ell)| \\ &< \sum_{\ell=0}^{[\varepsilon_1 n]} \binom{n}{\ell} (\theta''_3)^n \\ &< D_3\theta_3^n\end{aligned}$$

provided ε_1 is sufficiently small. □

5. Reduction to expanding maps

5.1. Purpose of this section. — From Assertion (1) in Propostion A, we know that $f^R \colon \Lambda \circlearrowleft$ sends γ^s-fibers to γ^s-fibers, so that *topologically* a quotient map is well defined. More precisely, let $\bar{\Lambda} = \Lambda/\approx$ where $\approx$ is the equivalence relation defined by $z \approx z'$ iff $z' \in \gamma^s(z)$. Then $\overline{f^R} \colon \bar{\Lambda} \circlearrowleft$ makes sense, and with $\bar{\Lambda}_i$ having the obvious meaning, $\overline{f^R}$ maps each one of the Cantor sets $\bar{\Lambda}_i$ homeomorphically onto $\bar{\Lambda}$.

The aim of this section is to study the *differential* properties of $\overline{f^R} \colon \bar{\Lambda} \circlearrowleft$ in the sense of the Jacobian of $\overline{f^R}$ with respect to a certain reference measure. Let $T\colon (X_1, m_1) \to (X_2, m_2)$ be a measureable bijection between two measure spaces. We say that T is *nonsingular* if T maps sets of m_1-measure 0 to sets of m_2-measure 0. For a nonsingular transformation T, we define the Jacobian of T with respect to m_1 and m_2, written $J_{m_1,m_2}(T)$ or simply JT, to be the Radon-Nikodym derivative $d(m_2 \circ T)/dm_1$.

Proposition B. — *There is a measureable family of reference measures $\{m_\gamma, \gamma \in \Gamma^u\}$ with the following properties:*

(1) *Each m_γ is supported on $\gamma\cap\Lambda$; it is a finite measure equivalent to the restriction of 1-dimensional Lebesgue measure on γ to $\gamma \cap \Lambda$.*
(2) *m_γ is invariant under sliding along γ^s,* i.e. *if $\theta\colon \gamma \cap \Lambda \to \gamma' \cap \Lambda$ is defined by $\{\theta(z)\} = \gamma' \cap \gamma^s(z)$, then for $E \subset \gamma \cap \Lambda$, $m_{\gamma'}(\theta E) = m_\gamma(E)$.*
(3) *For $z \in \gamma \cap \Lambda_i$, let $Jf^R(z)$ denote the Jacobian of $f^R \mid (\gamma \cap \Lambda_i)$ at z with respect to our reference measures on the respective γ^u-curves (we know that $f^R \mid (\gamma \cap \Lambda_i)$ is nonsingular on account of (1)). Then*

$$Jf^R(z) = Jf^R(z')$$

for all $z' \in \gamma^s(z)$.
(4) *$\exists\, \lambda > 1$ s.t. $Jf^R(z) \ge \lambda^R$ a.e.*

(5) *Restricted to each* $\gamma \cap \Lambda_i$, $(\log Jf^R) \circ (f^R)^{-1}$ *is "Hölder" in the sense to be made precise in Section 5.4, with uniform estimates independent of* γ *or* i.

Property (2) above tells us that $\{m_\gamma\}$ defines a reference measure $\bar{m}$ on our quotient space $\bar{\Lambda}$. Property (3) says that $\overline{f^R}\colon \bar{\Lambda} \circlearrowleft$ is nonsingular w.r.t. $\bar{m}$; we will call its Jacobian $J\overline{f^R}$. Properties (4) and (5) allow us to view $\overline{f^R}\colon (\bar{\Lambda}, \bar{m}) \circlearrowleft$ as a piecewise uniformly expanding map whose derivative has a certain "Hölder" property. We use " " for "Hölder" because it is not the usual Hölder condition; the relevant condition here is dynamically defined and will be explained in Section 5.4.

Our proof of Proposition B is essentially an adaption of some ideas used in the construction of Gibbs states. See e.g. **[B]** for an exposition.

5.2. The reference measures. — In this subsection we define $\{m_\gamma, \gamma \in \Gamma^u\}$ and prove Properties (1) and (2) in Proposition B. For simplicity of notation we will write m instead of m_γ when there is no ambiguity about γ. The Jacobian w.r.t. m will be denoted $J(\cdot)$, while the one w.r.t. 1-dimensional Lebesgue measure on γ^u-curves will be denoted $(\cdot)'$, *i.e.* $f'(z) = |Df_z\tau(z)|$ for $z \in \gamma^u$.

We pick and fix an arbitrary γ^u-curve in the definition of Λ and call it $\widehat{\gamma}$. For $z \in \Lambda$, let $\widehat{z}$ denote the point in $\widehat{\gamma} \cap \gamma^s(z)$, and let $\varphi(z) = \log f'(z)$. We define for $n = 1, 2, \ldots$

$$u_n(z) \equiv \sum_{i=0}^{n-1} \left(\varphi(f^i z) - \varphi(f^i \widehat{z})\right).$$

Sublemma 8. — $\exists\, C' > 0$ *and* b' *with* $b < b' \ll 1$ *s.t.* $\forall\, n > k \geq 0$,

$$\sum_{i=k}^{n} \left(\varphi(f^i z) - \varphi(f^i \widehat{z})\right) \leq C'(b')^k$$

Proof. — First we write

$$\varphi(f^i z) - \varphi(f^i \widehat{z}) = \log \frac{f'(f'z)}{f'(f^i \widehat{z})} \leq \frac{|f'(f^i z) - f'(f^i \widehat{z})|}{f'(f^i \widehat{z})}.$$

Then letting $\tau_i = \tau(f^i z)$ and $\widehat{\tau}_i = \tau(f^i \widehat{z})$, we have

$$|f'(f^i z) - f'(f^i \widehat{z})| \leq |Df_{f^i z}\tau_i - Df_{f^i \widehat{z}}\tau_i| + |Df_{f^i \widehat{z}}\tau_i - Df_{f^i \widehat{z}}\widehat{\tau}_i|.$$

The first term above is clearly $\leq C'b^i$ since $d(f^i z, f^i \widehat{z}) \leq Cb^i$ (Lemma 2 (1)). The second term is $\leq 5|\tau_i - \widehat{\tau}_i|$, which we estimate by

$$\begin{aligned} \sphericalangle(\tau_i, \widehat{\tau}_i) &\leq \sphericalangle(Df^i_z\tau_0, Df^i_{\widehat{z}}\tau_0) + \sphericalangle(Df^i_{\widehat{z}}\tau_0, Df^i_{\widehat{z}}\widehat{\tau}_0) \\ &\leq Cb^{i/4} + Cb^i, \end{aligned} \tag{5.1}$$

the first because of the Matrix Pertubation Lemma in Section 1.5 and the second because $Df^i_{\widehat{z}}$ is hyperbolic and τ_0 and $\widehat{\tau}_0$ are bounded away from their most contracted direction.

To complete the proof, observe that $f'(f^j\hat{z}) \geq b/5\ \forall j$, and that $f'(f^j\hat{z}) \geq \delta$ for the first few j's. The desired conclusion follows easily with, say, $b' > b^{1/8}$. □

It follows from Sublemma 8 that $u = \lim_n u_n$ exists for all $z \in \Lambda$. We know in fact that $|u|$ can be made arbitrarily small for b small. On each γ, let m be the measure whose density w.r.t. Lebesgue measure on γ is $\chi_{\Lambda\cap\gamma} \cdot e^u$, $\chi_{(\cdot)}$ being the characteristic function. Property (1) of Proposition B is immediate. Next follows a lemma, which gives a Lipschitz estimate for the tangential derivative at free return times: $(f^n)'z$. This estimate is used in several places: in the proofs of the Hölder regularity of the Jacobians (Section 5.4), and the invariance of the reference measures m_γ (Assertion (2) of Proposition B).

For $z_1, z_2 \in \gamma$, let $[z_1, z_2]$ denote the segment of γ between z_1 and z_2.

Sublemma 9. — *$\exists\, C_2'$ depending on δ s.t. the following holds for each γ^u and every $n \geq 0$. Let $\omega \subset \gamma^u$ be a segment in Ω_n, and suppose that*

(i) *for each $i \in n$, $f^i\omega \subset 3I_{\mu j}$ for some μ, j;* *and*

(ii) *$f^n\omega$ is free.*

Then $\forall\, z_1, z_2 \in \omega$ we have

(1) $\log \dfrac{(f^n)'z_1}{(f^n)'z_2} \leq C_2' e^{\alpha n}\, |f^n[z_1, z_2]|$;

(2) $\log \dfrac{(f^n)'z_1}{(f^n)'z_2} \leq C_2'\, |f^n[z_1, z_2]|$ *if $f^n\omega \supset \Omega_0$.*

Proof. — We follow the proof of Proposition 2 in **[BY]** p. 562–564, but make an improvement in the estimates. As in this proof we obtain

$$T \stackrel{def}{=} \log \frac{(f^n)'z_1}{(f^n)'z_2} \leq C \sum_{k=0}^{q} \frac{|f^{t_k}[z_1, z_2]|}{e^{-\nu_k}}$$

where $\{t_k\}_{k=0}^{q}$ are the free return times, $t_q = n$, and $f^{t_k}\omega \overset{\subset}{\approx} I_{\nu_k}$. We then define

$$m(\nu) = \max\{t_k : \nu_k = \nu\},$$

and using the fact that $|f^{t_{k+1}}[z_1, z_2]| \geq 2|f^{t_k}[z_1, z_2]|$ we have

$$T \leq C' \sum_{\nu \in S} \frac{|f^{m(\nu)}[z_1, z_2]|}{e^{-\nu}}$$

where S is the set of ν_k's not counted with multiplicity. Since $f^i\omega$ lies $> \delta e^{-\alpha i}$ from the critical set for each i, and the $m(\nu)$'s are distinct for different ν's, we obtain

$$\sum_{\nu \in S} \frac{|f^{m(\nu)}[z_1, z_2]|}{e^{-\nu}} \leq e^{\alpha n} \left(\sum_{k=0}^{q} 2^{-k} \right) |f^n[z_1, z_2]|$$

proving (1). To prove (2) we have as in the proof of Proposition 2 in **[BY]**

$$T \le C' \sum_{\nu \in S} \frac{|f^{m(\nu)}[z_1, z_2]|}{e^{-\nu}} \le C' \sum_{\nu} \frac{1}{\nu^2} = C_1,$$

where C_1 is the usual distortion constant (see Section 1.4). Now for each ν apply this to points in $f^{m(\nu)}\omega$ for the time interval $[m(\nu), n]$ to obtain

$$\frac{|f^{m(\nu)}[z_1, z_2]|}{|f^{m(\nu)}\omega|} \le C_1 \frac{|f^n[z_1, z_2]|}{|f^n\omega|},$$

and conclude that

$$\begin{aligned} T &\le C' \sum_{\nu \in S} \frac{|f^{m(\nu)}\omega|}{e^{-\nu}} \cdot \frac{|f^{m(\nu)}[z_1, z_2]|}{|f^{m(\nu)}\omega|} \\ &\le C' \cdot \left(\sum_{\nu \in S} \frac{|f^{m(\nu)}\omega|}{e^{-\nu}} \right) \cdot C_1 \frac{|f^n[z_1, z_2]|}{|f^n\omega|} \\ &\le \frac{C_1^2}{|\Omega_0|} |f^n[z_1, z_2]|. \quad \square \end{aligned}$$

□

Let γ and γ' be arbitrary curves in Γ^n, and let $\theta : \gamma \cap \Lambda \to \gamma'$ be defined by $\theta(z) \in \gamma^s(z) \cap \gamma'$. Property (2) of Proposition B follows from the following sublemma:

Sublemma 10. — *Temporarily let μ_γ and $\mu_{\gamma'}$ denote the Lebesgue measures on γ and γ' respectively. Then $\theta_*^{-1}\mu_{\gamma'}$ is absolutely continuous wrt μ_γ, written $\theta_*^{-1}\mu_{\gamma'} \prec\!\prec \mu_\gamma$, and*

$$\frac{d\theta_*^{-1}\mu_{\gamma'}}{d\mu_\gamma}(z) = e^{u(z)-u(\theta z)} \qquad \textit{for} \qquad \mu_\gamma \textit{ a.e. } z \in \gamma \cap \Lambda.$$

Absolutely continuity arguments are well known in dynamical systems (see e.g. **[PS]**), but since our setting is a little nonstandard let us include a proof. Observe that the m-measures are designed precisely so that θ takes m-measures to m-measures.

Proof. — Let $\omega \subset \widehat{\omega}$ be subsegments of γ with the property that $\widehat{\omega}$ makes a regular return to Ω_0 at time $\widehat{k}$, ω is free at time $k > \widehat{k}$, and all points in ω have the same itinerary up to time k in the usual sense. We require that $0 \ll \widehat{k} \ll k$ and that $|f^{\widehat{k}}\omega| \gg (Cb)^{\widehat{k}/4}$. (The second condition requires that k not be too much larger than $\widehat{k}$; it is not a serious imposition.) Let $\omega' \subset \widehat{\omega}'$ be the corresponding subsegments of γ'. In what follows "$a \approx b$" means that a/b is very near 1 and tends to 1 as all the "$\gg$" tend to ∞.

Let z be an arbitrary point in $\omega \cap \Lambda$, and let $z' = \theta z$. We claim that

$$|\omega'| \approx \frac{|f^{\widehat{k}}\omega'|}{(f^{\widehat{k}})'z'} = \frac{|f^{\widehat{k}}\omega'|}{|f^{\widehat{k}}\omega|} \cdot \frac{|f^{\widehat{k}}\omega|}{(f^{\widehat{k}})'z} \cdot \frac{(f^{\widehat{k}})'z}{(f^{\widehat{k}})'z'} \approx |\omega| \cdot e^{u(z)-u(z')}.$$

For the first "$\approx$" use the fact that $(f^{\widehat{k}})'$ is roughly identical at all points in ω'. This is true by Sublemma 9 provided that $k - \widehat{k}$ is sufficiently large. The same argument is used for ω in the second "$\approx$". Additionally we need the fact that $|f^{\widehat{k}}\omega'| \approx |f^{\widehat{k}}\omega|$, which is true because the two curves are so short they can be regarded as straight lines, and their lengths are $\gg (Cb)^{\widehat{k}/4}$ while their slopes are $< (Cb)^{\widehat{k}/4}$ apart (by the Matrix Perturbation Lemma in Section 1.5) In the third "$\approx$" we use Sublemma 8 and the fact that $\widehat{k}$ is large.

Let $\widetilde{\Lambda} = \{z \in \Lambda : z \text{ makes infinitely returns to } \Lambda\}$. We leave it as an excercise for the reader to verify that there is a cover $\mathcal{U}$ of $\widetilde{\Lambda}$ by pairwise disjoint sets of the type ω above with M arbitrarily large. To prove $\theta_*^{-1}\mu_{\gamma'} \prec\prec \mu_\gamma$, let A be a closed subset of $\gamma \cap \widetilde{\Lambda}$. Choosing a subcover $\{w_i\}$ of $\mathcal{U}$ s.t. $\mu_\gamma(\bigcup \omega_i) < \mu_\gamma(A) + \varepsilon$, we have that $\mu_{\gamma'}(\theta A) \lesssim e^{2\max|u|} \sum_i \mu_{\gamma'}(\omega_i) < e^{2\max|u|}(\mu_\gamma(A) + \varepsilon)$. To prove the statement on Radon-Nikodym derivatives, consider a Lebesgue density point z of $\gamma \cap \widetilde{\Lambda}$ and choose ω containing z with $|\omega \cap \Lambda| \approx |\omega|$, $|\omega' \cap \Lambda| \approx |\omega'|$. □

5.3. The Jacobians. — For a.e. $z \in \Lambda \cap \gamma$, $\gamma \in \Gamma^u$, we have

$$J(f^R)(z) = (f^R)'z \cdot \frac{e^{u(f^R z)}}{e^{u(z)}}.$$

Proof of Property (3) in Proposition B. — We will verify that $J(f^R)(z)$ depends only on $\widehat{z}$ and not on z:

$$\begin{aligned}\log J(f^R)z &= \sum_{i=0}^{R-1} \varphi(f^i z) \;+ \sum_{i=0}^{\infty} \left(\varphi\left(f^i\left(f^R z\right) - \varphi\left(f^i\left(\widehat{f^R z}\right)\right)\right)\right) \\ &\qquad - \sum_{i=0}^{\infty} \left(\varphi\left(f^i z\right) - \varphi\left(f^i \widehat{z}\right)\right) \\ &= \sum_{i=0}^{R-1} \varphi\left(f^i \widehat{z}\right) + \sum_{i=0}^{\infty} \left(\varphi\left(f^i\left(f^R \widehat{z}\right)\right) - \varphi(f^i(\widehat{f^R z}))\right).\end{aligned}$$

□

Proof of Property (4) in Proposition B. — As observed earlier on, $|u|$ can be made arbitrarily small; it is in fact of order b. Our Jacobian $J(f^R)z$ is therefore a small pertubation of $(f^R)', z$, which is $\geq e^{c_1 R}$ for some $c_1 > \frac{1}{3}\log 2$ (see Section 1.3). □

5.4. Regularity of the Jacobian. — Having established that $\bar{m}$ and $J(\overline{f^R})$ make sense on $\bar{\Lambda}$, we now introduce a dynamically defined notion of "Hölderness" satisfied by $J(\overline{f^R}) \circ (\overline{f^R})^{-1}$.

For $z_1, z_2 \in \bar{\Lambda}$, define their *separation time* $s(z_1, z_2)$ to be the smallest n s.t. $f^n z_1$, $f^n z_2$ do not lie in three contiguous $I_{\nu j}$'s. Here we have taken the liberty to confuse $z \in \bar{\Lambda}$ with the representative on some $\gamma \in \Gamma^u$, and to include $[-1, -\delta)$ and $(\delta, 1]$ when we speak about $I_{\nu j}$'s.

Definition 2. — A function $\psi\colon \bar\Lambda \to \mathbb{R}$ is said to be *Hölder with respect to the separation time* $s(\cdot,\cdot)$ if $\exists\, C > 0$ and $\beta < 1$ s.t. for m-a.e. $z_1, z_2 \in \bar\Lambda$, $|\psi z_1 - \psi z_2| \le C\beta^{s(z_1,z_2)}$.

The following lemma gives the precise statement of Property (5) in Proposition B.

Lemma 6. — *$\exists\, C_2 > 0$ and $\beta < 1$ s.t. for every i and $\forall\, z_1, z_2 \in \bar\Lambda_i$,*

$$\left| \frac{J(\overline{f^R})(z_1)}{J(\overline{f^R})(z_2)} - 1 \right| \le C_2 \beta^{s(\overline{f^R}z_1, \overline{f^R}z_2)}.$$

This can be rephrased as follows. For each i let $(\overline{f^R})_i^{-1}\colon \bar\Lambda \to \bar\Lambda_i$ be the inverse of $\overline{f^R} \mid \bar\Lambda_i$. Then $z \to J(\overline{f^R}) \circ (\overline{f^R})_i^{-1}(z)$ is Hölder w.r.t. $s(\cdot,\cdot)$ above with uniform C_2 and β independent of i.

Remark 6. — Some explanations are probably in order here.

(1) *Why would the regularity of the Jacobian involve separation times?* If we were working with a 1-d map $f\colon [-1,1] \circlearrowleft$, then $x \mapsto \log f'(x)$ is Hölder in the usual sense — provided that near the critical point 0, we compare two points only if they are much closer to each other than to 0, e.g. if they lie in 3 contiguous $I_{\nu j}$'s. In particular, two points on opposite sides of 0 cannot be compared. In the present situation $\bar\Lambda$ is obtained by collapsing W^s_{loc}-curves, so that points in $\bar\Lambda$ represent not points in $\mathbb{R}^2$ but *futures* of orbits, and $J(\overline{f^R})(z)$ has incorporated into it information on the entire orbit of the point $z \in \Lambda$. Now two points $z_1, z_2 \in \gamma^u$ could be arbitrarily near each other, and be mapped at some future time to opposite sides of the critical set. The sooner this takes place, the less one could expect $J(\overline{f^R})(z_1)$ and $J(\overline{f^R})(z_2)$ to be comparable. Hence separation time enters.

(2) *Why $C\beta^{s(\overline{f^R}z_1, \overline{f^R}z_2)}$?* Built into this formulation is the assumption that the map f is expanding on average. Consider for simplicity a C^2 uniformly expanding map g with $g' \approx \lambda > 1$. Fix $\delta > 0$ small enough that $d(gx, gy) \approx \lambda\, d(x,y)$ whenever $d(gx, gy) < \delta$. Consider x, y and n s.t. $d(g^i x, g^i y) < \delta\ \forall i \le n$. The following estimate is standard:

$$\begin{aligned} |\log(g^n)'x - \log(g^n)'y| &\le \sum_{i=0}^{n-1} \left|\log g'(g^i x) - \log g'(g^i y)\right| \\ &\le C \sum_{0}^{n-1} d(g^i x, g^i y) \le C' d(g^n x, g^n y). \end{aligned}$$

Now if $s(x', y')$ is the first time $d(g^s x', g^s y') > \delta$, then $d(g^n x, g^n y) \approx C\lambda^{-s(g^n x, g^n y)}$, so that

$$|\log(g^n)'x - \log(g^n)'y| \le C'(\lambda^{-1})^{s(g^n x, g^n y)}.$$

Here β plays the rôle of λ^{-1}, and separation may occur long before two orbits move $> \delta$ apart.

We now proceed with the proof of Lemma 6.

Sublemma 11. — *$\exists C_2'' > 0$ and $\beta < 1$ s.t. $\forall z_1, z_2 \in \gamma^u \cap \Lambda$, $|u(z_1) - u(z_2)| \le C_2'' \beta^{s(z_1,z_2)}$.*

Proof. — Let $n = s(z_1, z_2)$, and pick $k \in [\frac{n}{3}, \frac{n}{2}]$ s.t. $f^k[z_1, z_2]$ is free. We know that k exists because if $f^{n/3}[z_1, z_2]$ is in bound state, then it was $> e^{-\frac{n}{3}\alpha}$ from the critical set when the last (total) bound period was initiated, which means that this bound period must expire before time $\frac{1}{3}n + 4\alpha\frac{1}{3}n < \frac{1}{2}n$ (see Section 1.2). Write

$$u(z_1) - u(z_2) = \sum_{i=0}^{\infty} \left\{ \left(\varphi(f^i z_1) - \varphi(f^i \widehat{z}_1)\right) - \left(\varphi(f^i z_2) - \varphi(f^i \widehat{z}_2)\right) \right\}.$$

The part $\sum_{i=0}^{k-1} \{\cdot\}$ is estimated by

$$\left| \sum_{i=0}^{k-1} \{\cdot\} \right| \le \left| \log \frac{(f^k)' z_1}{(f^k)' z_2} \right| + \left| \log \frac{(f^k)' \widehat{z}_1}{(f^k)' \widehat{z}_2} \right|.$$

The first term on the right is $\le C_2' e^{\alpha k} |f^k[z_1, z_2]|$ by Sublemma 9 (1). Observing that separation can occur only when $f^n[z_1, z_2]$ is free and using the estimates for orbits ending in free states in Section 1.3, we have that $|f^k[z_1, z_2]| \le e^{-c_1(n-k)} |f^n[z_1, z_2]|$ for some $c_1 \gtrsim \frac{1}{3}\log 2$. Altogether this first term contributes $\le C_2' e^{\alpha\frac{n}{2} - c_1\frac{n}{2}} \le C_2' \beta^n$ for some $\beta < 1$. The corresponding term for $\widehat{z}_i$, $i = 1, 2$, is handled similarly.

For $\sum_{i \ge k} \{\cdot\}$ we have, by Sublemma 8,

$$\begin{aligned} \left| \sum_k^{\infty} \{\cdot\} \right| &\le \left| \sum_k^{\infty} \left(\varphi(f^i z_1) - \varphi(f^i \widehat{z}_1)\right) \right| + \left| \sum_k^{\infty} \left(\varphi(f^i z_2) - \varphi(f^i \widehat{z}_2)\right) \right| \\ &\le 2C'(b')^k \le 2C' b^{n/8}. \end{aligned}$$

□

Proof of Lemma 6. — It suffices to work with one γ^u-curve. We consider $z_1, z_2 \in \gamma^u \cap \Lambda_i$ and let $n = s(f^R z_1, f^R z_2)$. We noted in Section 5.3 that

$$\begin{aligned} \log \frac{J(f^R) z_1}{J(f^R) z_2} &= \log \frac{(f^R)' z_1}{(f^R)' z_2} + \left(u(f^R z_1) - u(f^R z_2)\right) - \left(u(z_1) - u(z_2)\right) \\ &\stackrel{\text{def}}{=} \text{(I)} + \text{(II)} + \text{(III)}. \end{aligned}$$

Since z_1 and z_2 lie in a segment that makes a regular return to Ω_0 at time R, we have by Sublemma 9 (2) that (I) $\le C_2' |f^R[z_1, z_2]|$. Using Section 1.3 again we see that $|f^R[z_1, z_2]| \le e^{-c_1 n} |f^n[f^R z_1, f^R z_2]| \le e^{-c_1 n}$. Also, (II) $\le C_2'' \beta^n$ by Sublemma 11, and (III) $\le C_2'' \beta^{s(z_1,z_2)}$ where $s(z_1, z_2)$ is obviously $> n$. □

6. Proofs of Theorems

To study the rate of mixing of f it is not sufficient to consider $f^R : \Lambda \circlearrowleft$ alone: the return time function R also plays an important role. In this section we construct a tower $\Delta = \bigcup_{\ell=0}^{\infty} \Delta_\ell$ the bottom level of which is Λ and construct a map $F : \Delta \circlearrowleft$ in a way analogous to that of building a special flow over $f^R : \Lambda \circlearrowleft$ under the function R. This allows us to consider the Perron-Frobenius operator or transfer operator associated with $\overline{F} : \overline{\Delta} \circlearrowleft$, the quotient map of $F : \Delta \circlearrowleft$ obtained by collapsing along local stable leaves. Spectral properties of this operator are summed up in Proposition C in Section 6.3. We refer the reader to **[Y]** for a proof of this Proposition, and derive from it the results of this paper.

6.1. Construction of a tower. — Let

$$\Delta \stackrel{\text{def}}{=} \{(z, \ell) : z \in \Lambda, \ell = 0, 1, 2, \ldots, R(z) - 1\}.$$

We introduce $F : \Delta \circlearrowleft$ defined by

$$F(z, \ell) = \begin{cases} (z, \ell + 1) & \text{if } \ell + 1 < R(z) \\ (f^R z, 0) & \text{if } \ell + 1 = R(z). \end{cases}$$

It is clear that there is a projection $\pi : \Delta \to \mathbb{R}^2$ s.t. $\pi|\Delta_0$ is the identity map on Λ and $f \circ \pi = \pi \circ F$.

An equivalent but less formal way of looking at Δ is to view it as the disjoint union $\bigcup_{\ell=0}^{\infty} \Delta_\ell$ where $\Delta_\ell \stackrel{\text{def}}{=} \{z \in \Lambda : R(z) > \ell\}$ denotes the ℓ^{th} level of the tower. Next we subdivide each level into components $\Delta_\ell = \bigcup_i \Delta_{\ell,i}$ in such a way that F has a Markov type property with respect to the partition $\{\Delta_{\ell,i}\}$. Again using 1-d language, a natural subdivision of Δ_ℓ might be the restriction of the partition $\widetilde{\mathcal{P}}_\ell$ constructed in Section 3.4; but this partition is too "big". We introduce instead a sequence of partitions $\mathcal{P}_n$ on $\widetilde{\Omega}_n$ so that $\mathcal{P}_n$ is coarser than $\widetilde{\mathcal{P}}_n$ and each element of $\mathcal{P}_n$ contains no more than finitely many elements of $\mathcal{P}_{n+1}$. This is easily done by following the algorithm in the construction of $\widetilde{\mathcal{P}}_n$, except that when $f^n\omega$ is a regular return, no subdivisions are made on that part of ω that gets mapped onto $\Omega_0 - \Omega_\infty$. (Elements of $\mathcal{P}_n$ are not necessarily intervals; they may have "holes" due to absorption into Ω_∞.) We say that $z_1, z_2 \in \Delta_\ell$ are in the same component $\Delta_{\ell,i}$ if they both lie in the same element of $\mathcal{P}_\ell$.

We summarize the topological properties of $F : \Delta \circlearrowleft$:

(I) Δ is the disjoint union $\bigcup_{\ell=0}^{\infty} \Delta_\ell$ where the ℓ^{th} level Δ_ℓ is a copy of $\{z \in \Lambda : R(z) > \ell\}$; each Δ_ℓ is further subdivided into a finite number of "components" $\Delta_{\ell,i}$ each one of which is a copy of an s-subrectangle of Λ.

(II) Under F, each $\Delta_{\ell,i}$ is mapped onto the union of finitely many components of $\Delta_{\ell+1}$ and possibly a u-subrectangle of $\Delta_0^{\pm}$. Let $\Delta_{\ell,i}^* = \Delta_{\ell,i} \cap F^{-1}\Delta_0$. We think of points in $\bigcup \Delta_{\ell,i}^*$ as "returning to the bottom level" under F, while other points "move upward" to the next level.

From the description above it is clear that the quotient map $\overline{F} : \overline{\Delta} \circlearrowleft$ obtained by collapsing γ^s-curves as in Section 5 is well defined. Let $\bar{m}$ be the reference measure on $\overline{\Lambda}$ or $\overline{\Delta}_0$. Since each $\overline{\Delta}_{\ell,i}$ is a copy of a subset of $\overline{\Delta}_0$, $\bar{m}$ is defined on $\overline{\Delta}_{\ell,i}$ via the natural identification. Let $J(\overline{F})$ denote the Jacobian of $\overline{F}$ with respect to $\bar{m}$; more precisely, if $z \in \overline{\Delta}_{\ell,i}$ and $\overline{F}z \in \overline{\Delta}_{\ell+1,i'}$, then $J\overline{F}(z)$ is the Jacobian of the map $\overline{F} \mid \left(\overline{\Delta}_{\ell,i} \cap \overline{F}^{-1}\left(\overline{\Delta}_{\ell+1,i'}\right)\right)$. Also, given the present setting it is natural to define the separation time of $z_1, z_2 \in \Delta$ to be

$$s(z_1, z_2) \equiv \text{ the smallest } n \geq 0 \text{ s.t. } f^n z_1 \text{ and } f^n z_2 \text{ lie in different } \Delta_{\ell,i}\text{'s}.$$

This definition of $s(\cdot,\cdot)$ will permanently replace the one in Section 5.4. Observe that under the present definition, $z_1, z_2 \in \Delta_0$ separate faster than under the old one. Hence the distortion estimate in Lemma 4 is all the more valid. Again we summarize:

(III) There is a reference measure $\bar{m}$ on $\bar{\Delta}$ uniformly equivalent to the restriction of Lebesgue measure on $\gamma^u \cap \Lambda$ for every γ^u such that with respect to $\bar{m}$, the Jacobian $J\overline{F}$ of $\overline{F}$ satisfies:
  (i) $J\overline{F}(z) = 1 \;\forall\, z \notin \bigcup \bar{\Delta}^*_{\ell,i}$;
  (ii) $\exists\, C_2 > 0$ and $\beta < 1$ s.t. $\forall\, z_1, z_2 \in \overline{F}^{-1}\left(\bar{\Delta}_0^{\pm}\right)$,

$$\left| \frac{J\overline{F}(z_1)}{J\overline{F}(z_2)} - 1 \right| \leq C_2 \beta^{s(\overline{F}z_1, \overline{F}z_2)}.$$

Here $\bar{\Delta}_0^+$ and $\bar{\Delta}_0^-$ are the two components of $\bar{\Delta}_0$. Let

$$C_3 = \frac{\bar{m}\left(\bar{\Delta}_0\right)}{\min\left(\bar{m}\left(\bar{\Delta}_0^+\right), \bar{m}\left(\bar{\Delta}_0^-\right)\right)}.$$

We state for the record the following very important tail estimate on the height of the tower $\bar{\Delta}$ (or equivalently the return times to $\bar{\Delta}_0$).

(IV) The height function $R : \bar{\Delta}_0 \to \mathbb{Z}^+$ has the following properties:
  (i) $R \geq N$ where N is chosen so that $C_2 e^{C_2} C_3 \beta^N \leq 1/100$;
  (ii) $\exists\, C_0 > 0$ and $\theta_0 < 1$ s.t.

$$\bar{m}\{R > n\} \leq C_0 \theta_0^n \qquad \forall\, n \geq 0.$$

The lower bound for R in (i) is for purposes of guaranteeing a definite amount of contraction for the Perron-Frobenius operator between consecutive returns of an orbit to the base. The feasibility of such a bound was arranged in Section 3.4. Note the order in which the constants in (III) and (IV) are chosen: C_2, C_3 and β, and hence N can be chosen to depend only on the derivatives of f and not on the construction of the tower; whereas C_0 and θ_0 depend on f as well as on N. The tail estimate in (ii) is a slight reformulation of Propostion A (IV)(ii).

6.2. SRB measures: Proof of Theorem 1. — We construct in this subsection an SRB measure ν for f with $\nu(\Lambda) > 0$. This gives an alternate proof of Theorem 1 to that in **[BY]**.

Let $f^R : \Lambda \circlearrowleft$ be the mapping with $f^R|\Lambda_i = f^{R_i}|\Lambda_i$ for each i, and let μ_0 denote the restriction of 1-dimensional Lebesgue measure on Ω_0 to $\Lambda \cap \Omega_0$. For $n = 1, 2, \ldots$, let

$$\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} \left(f^R\right)_*^i \mu_0.$$

Then μ_n is supported on a countable number of γ^u-curves on each one of which it has a density ρ_n. Clearly, $\rho_n = 0$ on $\gamma^u - \Lambda$. The distortion estimate in Sublemma 9 tells us that

$$\frac{\rho_n(x)}{\rho_n(y)} \leq C_1 \qquad \text{for a.e. } x, y \in \Lambda \cap \gamma^u.$$

From this and from the absolute continuity of the curves in Γ^s and the boundedness of the Radon-Nikodym derivatives (see Sublemma 10), it follows that if $\widetilde{\rho}_n \mid \gamma^u := \rho_n/\mu_n(\gamma^u)$, then $M^{-1} \leq \widetilde{\rho}_n \leq M$ a.e. on $\gamma^u \cap \Lambda$ for some M independent of n and γ^u.

Letting $n \to \infty$, a subsequence μ_{n_k} converges weakly to μ_∞. We have immediately that μ_∞ is f^R-invariant and that it is supported on Λ.

Let $\{\mu_\infty^\gamma\}$ be the conditional measures of μ_∞ on γ^u-curves, and let Q_ω be an s-subrectangle of Q corresponding to an arbitrary subsegment ω of some γ^u. Then for a.e. $\gamma \in \Gamma^u$, we have

$$M^{-1} \min_{\gamma^u \in \Gamma^u} |Q_\omega \cap \gamma^u| \leq \mu_\infty^\gamma(Q_\omega \cap \gamma) \leq M \max_{\gamma^u \in \Gamma^u} |Q_\omega \cap \gamma^u|,$$

proving (again using the absolute continuity of Γ^s) that μ_∞^γ is uniformly equivalent to the arclength measure $s \mid (\gamma \cap \Lambda)$.

To extract from μ_∞ an f-invariant measure, simply let

$$\nu = \sum_{i=0}^{\infty} f_*^i \left(\mu_\infty \mid \{R > i\}\right).$$

That ν is a finite measure follows from Propostion A (4) (ii); we may therefore normalize and assume $\nu\left(\mathbb{R}^2\right) = 1$. It is clear that ν satisfies the definition of an SRB measure as defined in Section 1.7.

By the same token, we could view μ_∞ as a measure on Δ_0, and construct as above an F-invariant measure $\widetilde{\nu}$ on Δ with $\pi_*\widetilde{\nu} = \nu$. It is also clear from the discussion above that $\bar{\nu}$, the measure on $\bar{\Delta}$ that is the quotient of $\widetilde{\nu}$, is uniformly equivalent to our reference measure $\bar{m}$.

6.3. Definition and properties of the Perron-Frobenius operator associated with $\overline{F} : \bar{\Delta} \circlearrowleft$. — First we introduce the function space on which our operator acts. Fix $\varepsilon > 0$ with the following two properties:

(i) $e^{2\varepsilon}\theta_0 < 1$ where θ_0 is as in Section 6.1 (IV) (ii);

(ii) $\frac{1}{\bar{m}\bar{\Delta}_0} \sum_{\ell,i} \bar{m}\bar{\Delta}^*_{\ell,i} e^{\ell\varepsilon} \leq 2$.

Note that property (ii) is consistent with

$$\frac{1}{\bar{m}\bar{\Delta}_0} \sum_{\ell,i} \bar{m}\bar{\Delta}^*_{\ell,i} = 1.$$

We remark for future reference the relative sizes of β and $e^{-\varepsilon}$: from Section 6.1 (IV)(i) we have that β^N times various constants is $\leq 1/100$, while (ii) above implies that $e^{\varepsilon N} \leq 2$. Thus β should be thought of as $< e^{-\varepsilon}$.

Our function space X will consist of those $\overline{\varphi} : \bar{\Delta} \to \mathbb{C}$ with $\|\overline{\varphi}\| < \infty$, where $\|\cdot\|$ is a weighted L^∞+ Hölder norm defined as follows: let $\overline{\varphi}_{\ell,i} = \overline{\varphi} \mid \bar{\Delta}_{\ell,i}$, and let $|\cdot|_p$ denote the L^p-norm $(1 \leq p \leq \infty)$ wrt the reference measure $\bar{m}$. We define

$$\left\|\overline{\varphi}_{\ell,i}\right\|_\infty = \left|\overline{\varphi}_{\ell,i}\right|_\infty e^{-\ell\varepsilon}$$

and

$$\left\|\overline{\varphi}_{\ell,i}\right\|_h = \left(\operatorname*{ess\,sup}_{z_1,z_2\in\Delta_{\ell,i}} \frac{|\overline{\varphi}z_1 - \overline{\varphi}z_2|}{\beta^{s(z_1,z_2)}} \right) \cdot e^{-\ell\varepsilon}$$

where ε is as above. Finally let

$$\|\overline{\varphi}\| = \|\overline{\varphi}\|_\infty + \|\overline{\varphi}\|_h$$

where

$$\|\overline{\varphi}\|_\infty = \sup_{\ell,i} \left\|\overline{\varphi}_{\ell,i}\right\|_\infty \qquad \text{and} \qquad \|\overline{\varphi}\|_h = \sup_{\ell,i} \left\|\overline{\varphi}_{\ell,i}\right\|_h .$$

The Perron-Frobenius operator or transfer operator associated with the dynamical system $\overline{F} : \bar{\Delta} \circlearrowleft$ is defined by

$$\mathcal{P}\left(\overline{\varphi}\right)(z) = \sum_{w:\overline{F}w=z} \frac{\overline{\varphi}(w)}{J\overline{F}(w)}.$$

Our choice of $(X, \|\cdot\|)$ was to ensure that $\mathcal{P}$ has the following spectral properties:

Proposition C

(1) *$\mathcal{P} : X \to X$ is a bounded linear operator; its spectrum $\sigma(\mathcal{P})$ has the following properties:*
   - $\sigma(\mathcal{P}) \subset \{|\lambda| \leq 1\}$
   - $\exists\, \tau_0 < 1$ *s.t.* $\sigma(\mathcal{P}) \cap \{|\lambda| \geq \tau_0\}$ *consists of a finite number of points the eigenspaces corresponding to which are all finite dimensional.*

(2) $1 \in \sigma(\mathcal{P})$, *and $\bar{\rho} \in X$ is an eigenfunction corresponding to the eigenvalue* 1. *Moreover, if the greatest common divisor (gcd) of $\{R(z) : z \in \Delta_0\} = 1$, then* 1 *is the only element of $\sigma(\mathcal{P})$ with $|\lambda| = 1$ and its eigenspace is* 1*-dimensional.*

See [**Y**] for a proof of Proposition C.

6.4. Decay of correlations: Proof of Theorem 3. — Recall that we are concerned with a "good" Hénon map $f : \mathbb{R}^2 \circlearrowleft$ which we know has an attractor Σ and a unique SRB measure ν. *We assume for the rest of this paper that* (f^n, ν) *is ergodic for all* $n \geq 1$. The following conventions will be used: if φ is a function on $\mathbb{R}^2$ or on Σ, then the lift of φ to Δ will be called $\widetilde{\varphi}$; and if $\widetilde{\varphi} : \Delta \to \mathbb{C}$ is constant on γ^s-curves then we will sometimes confuse it with the obvious function on $\bar{\Delta}$ called $\overline{\varphi}$.

For $\gamma > 0$, let $\mathcal{H}_\gamma = \mathcal{H}_\gamma(\Sigma)$ denote the class of Hölder continuous functions on Σ with exponent γ, *i.e.*

$$\mathcal{H}_\gamma = \{\varphi : \Sigma \to \mathbb{R} \mid \exists C = C_\varphi \text{ s.t. } \forall x, y \in \Sigma, |\varphi x - \varphi y| \leq C|x - y|^\gamma\}.$$

We will use as shorthand for the correlation between φ and $\psi \circ f^n$ with respect to ν the notation $D_n(\varphi, \psi; \nu)$, *i.e.*

$$D_n(\varphi, \psi; \nu) = \left| \int (\psi \circ f^n)\, \varphi d\nu - \int \varphi d\nu \int \psi d\nu \right|.$$

We outline below the steps needed to derive from Proposition C an exponentially small bound for $D_n(\varphi, \psi; \nu)$. Since the derivation is completely formal and (with the exception of one small geometric fact noted below) has nothing to do with the present setting, we will refer the reader to **[Y]** for details of the proofs.

Here are the main steps of the argument:

(1) Observe that $D_n(\varphi, \psi; \nu) = D_n\left(\widetilde{\varphi}, \widetilde{\psi}; \widetilde{\nu}\right)$, and that by considering a power of f if necessary, we may assume $\gcd\{R(z) : z \in \Delta_0\} = 1$.

(2) In preparation for using the Perron-Frobenius operator, we maneuver $D_n(\widetilde{\varphi}, \widetilde{\psi}; \widetilde{\nu})$ into an object describable purely in terms of functions on $\bar{\Delta}$. This is done in two steps:

   (i) Fix $\kappa \in \left(0, \frac{1}{2}\right)$, and let $k = \kappa n$. Let $\mathcal{M}$ be the partition of Δ into $\{\Delta_{\ell,j}\}$, and let $\overline{\psi}_k : \bar{\Delta} \to \mathbb{R}$ be the function constant on elements η of $\mathcal{M}_{2k} := \bigvee_{i=0}^{2k-1} F^{-i}\mathcal{M}$ with $\overline{\psi}_k \mid \eta := \widetilde{\psi} \circ F^k$ (some selected point in η). Verify that

$$D_n\left(\widetilde{\varphi}, \widetilde{\psi}; \widetilde{\nu}\right) = D_{n-k}\left(\widetilde{\varphi}, \widetilde{\psi} \circ F^k; \widetilde{\nu}\right) \approx D_{n-k}\left(\widetilde{\varphi}, \overline{\psi}_k; \widetilde{\nu}\right).$$

   (ii) Let $\overline{\varphi}_k$ be defined as above, and let $\widetilde{\varphi}_k := d\left(F_*^k\left(\overline{\varphi}_k \widetilde{\nu}\right)\right)/d\widetilde{\nu}$. Verify that

$$D_{n-k}\left(\widetilde{\varphi}, \overline{\psi}_k; \widetilde{\nu}\right) \approx D_{n-k}\left(\widetilde{\varphi}_k, \overline{\psi}_k; \widetilde{\nu}\right)$$

   and observe that

$$D_{n-k}\left(\widetilde{\varphi}_k, \overline{\psi}_k; \widetilde{\nu}\right) = \left| \int \overline{\psi}_k \mathcal{P}^n\left(\overline{\varphi}_k \bar{\rho}\right) d\bar{m} - \int \overline{\psi}_k \bar{\rho} d\bar{m} \int \overline{\varphi}_k \bar{\rho} d\bar{m} \right|.$$

(3) Use Proposition C to prove that, for some $\tau_1 < 1$,

$$\left\| \mathcal{P}^n\left(\overline{\varphi}_k \bar{\rho}\right) - \left(\int \overline{\varphi}_k \bar{\rho} d\bar{m}\right) \bar{\rho} \right\| \leq \text{const} \cdot \tau_1^{n-2k}.$$

In Step (2) above, it is neccessary to translate the Hölder property for $\varphi, \psi \in \mathcal{H}_\gamma$ to a Hölder type condition for $\widetilde{\varphi}$ and $\widetilde{\psi}$. The following geometric fact is used:

Sublemma 12. — $\forall z \in \Delta$, $\operatorname{diam}\left(\pi F^k\left(\mathcal{M}_{2k}(z)\right)\right) \leq 2C\alpha^k$.

We leave the proof as an easy exercise.

6.5. Proof of the Central Limit Theorem. — Theorem 4 (the CLT) is also proved in **[Y]** but we prefer to give another proof here. As in **[Y]** this proof is based on a theorem of Gordin **[G]** but we apply Gordin's theorem to test functions in the Banach space X and use an L^2 approximation argument (which in fact uses the decay of correlation) to prove the theorem for Hölder test functions on Δ.

The version of Gordin's theorem we need may be stated as follows:

Theorem 5 (**[G]**). — *Let $(\Omega, \mathcal{F}, \nu)$ be a probability space, let $T : \Omega \circlearrowleft$ be a non-invertible measure-preserving transformation, and let $\varphi \in L^2(\nu)$ be s.t. $E\varphi = 0$. Suppose that*

$$(*) \qquad \sum_{j\geq 0} |E(\varphi|T^{-j}\mathcal{F})|_2 < \infty$$

Then

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \varphi \circ T^i \overset{\text{dist}}{\longmapsto} \mathcal{N}(0, \sigma)$$

where

$$\sigma^2 = \lim_{n\to\infty} \frac{1}{n} \int \left(\sum_{i=0}^{n-1} \varphi \circ T^i\right)^2 d\nu.$$

Exponential decay of correlations alone is not sufficient to conclude that $(*)$ holds. Suppose, however, that there is a reference measure m with respect to which T is non-singular, and suppose that $d\nu = \rho\, dm$ for some $\rho \geq c > 0$. Then we have a well-defined Perron-Frobenius operator given by $\mathcal{P}(\varphi) = \psi$, where ψ is the density of $T_*(\varphi m)$, and a gap in the spectrum of $\mathcal{P}$ (with respect to a suitable function space) is sufficient to conclude that $(*)$ converges exponentially. In fact, we have

$$(**) \qquad \int |E(\varphi|T^{-j}\mathcal{M})|^2\, d\nu \leq |\varphi|_\infty \int |\mathcal{P}^j(\varphi\rho)|\, dm$$

see **[K]** or **[Y]**. See also Ruelle's earlier work **[R]**.

For $\varphi \in \mathcal{H}_\gamma$ let $\widetilde{\varphi}(z) = \varphi \circ \pi(z)$ be the lift of φ to Δ. For $k \in \mathbb{Z}^+$, we use $\widetilde{\varphi}^{(k)}$ to denote $\widetilde{\varphi} \circ F^k$ and define $\overline{\varphi}_k$ by

$$\overline{\varphi}_k | A = \frac{1}{\widetilde{\nu}(A)} \int_A \widetilde{\varphi} \circ F^k\, d\widetilde{\nu},$$

where the A's are the elements of the partition $\mathcal{M}_{2k}$ defined in Section 6.4.

As in Section 6.3, for functions on $\overline{\Delta}$ we use $|\cdot|_p$ to denote the L^p-norm, $1 \leq p \leq \infty$, with respect to the reference measure $\overline{m}$ and $||\cdot||$ to denote the norm on the space X. Let $||\cdot||_{\mathcal{H}_\gamma}$ denote the usual γ-Hölder norm for functions $\varphi : \mathbb{R}^2 \to \mathbb{C}$, *i.e.*

$$||\varphi||_{\mathcal{H}_\gamma} = \sup_{x \in \mathbb{R}^2} |\varphi(x)| + \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{|x-y|^\gamma}.$$

The following estimates are used in several places.

Sublemma 13. — *There is a constant $C = C(f)$ such that for all functions $\overline{\varphi} \in X$ we have*

$$|\overline{\varphi}|_1 \leq C||\overline{\varphi}||.$$

This is an easy exercise (see **[Y]**, Section 3.2).

Sublemma 14. — *For $\varphi \in \mathcal{H}_\gamma$*

$$\sup |\overline{\varphi}_k - \widetilde{\varphi} \circ F^k| \leq C(f)||\varphi||_{\mathcal{H}_\gamma} \lambda^k \qquad \forall\, k \geq 0,$$

where $\lambda = \lambda(f, \gamma)$ satisfies $0 < \lambda < 1$.

This sublemma is a direct consequence of Sublemma 12.

Let us introduce the notation

$$S_n(\widetilde{\varphi}) = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \widetilde{\varphi} \circ F^j.$$

We now fix $\psi \in \mathcal{H}_\gamma$ with $\int \psi\, d\nu = 0$ and prove the CLT for this function. Observe first that there is a constant $K_0 = K_0(\psi)$ such that for all $\varepsilon > 0$, if $N(\varepsilon) := [K_0 \log(1/\varepsilon)]$, then

$$\sup |\overline{\psi}_{N(\varepsilon)} - \widetilde{\psi} \circ F^{N(\varepsilon)}| \leq \varepsilon.$$

This follows immediately from Sublemma 14.

Lemma 7. — *There is a function $r(\varepsilon)$ with $r(\varepsilon) \to 0$ as $\varepsilon \to 0$ such that for each $\varepsilon > 0$, if*

$$\begin{cases} T_n = S_n(\widetilde{\psi} \circ F^{N(\varepsilon)}) \\ U_n = S_n(\overline{\psi}_{N(\varepsilon)}) \end{cases}$$

then

$$\sup_{n \geq 0} ||T_n - U_n||_{L^2(\widetilde{\nu})} \leq r(\varepsilon).$$

For the proof of this lemma we need the following estimates:

Sublemma 15. — *There exist constants $C > 0$ and $0 < \beta_0, \tau < 1$, β_0 and τ depending only on f and $C = C(f, \psi)$, such that for all $j, N \in \mathbb{Z}^+$ the following hold:*

(i) $\left|\int \overline{\psi}_N \overline{\psi}_N \circ F^j d\overline{\nu}\right| \leq C\beta_0^{-2N} \tau^j$;

(ii) $\left|\int \overline{\psi}_N \widetilde{\psi}^{(N)} \circ F^j d\widetilde{\nu}\right| \leq C\beta_0^{-2N} \tau^j$;

(iii) $\left|\int \widetilde{\psi}^{(N)} \overline{\psi}_N \circ F^j d\widetilde{\nu}\right| \le C\beta_0^{-2N}\tau^j$;

(iv) $\left|\int \widetilde{\psi}^{(N)} \widetilde{\psi}^{(N)} \circ F^j d\widetilde{\nu}\right| \le C\beta_0^{-2N}\tau^j$.

Since $\overline{\psi}_N$ can also be viewed as a function on Δ, that the left side of (i)–(iv) above is $\le C\tau^j$ for some $C = C(f, \psi, N)$ follows from the decay of correlations proof outlined in Section 6.4. The proof of this sublemma consists of re-doing the estimates there and making transparent the dependence on N.. We carry this out for (i) and (ii) and leave the rest as exercises.

Proof of Sublemma 15(i) and (ii). — For $\overline{\varphi}, \overline{\eta} \in X$ we have

$$\begin{aligned} D_j(\overline{\varphi}, \overline{\eta}; \widetilde{\nu}) &= \left|\int \overline{\eta} \circ F^j \overline{\varphi}\, d\widetilde{\nu} - \left(\int \overline{\varphi}\, d\widetilde{\nu}\right)\left(\int \overline{\eta}\, d\widetilde{\nu}\right)\right| \\ &= \left|\int \overline{\eta} \left[\mathcal{P}^j(\overline{\varphi}\,\overline{\rho}) - \left(\int \overline{\varphi}\,\overline{\rho}\, d\overline{m}\right)\overline{\rho}\right] d\overline{m}\right| \\ &\le |\overline{\eta}|_\infty \left|\mathcal{P}^j(\overline{\varphi}\,\overline{\rho}) - \left(\int \overline{\varphi}\,\overline{\rho}\, d\overline{m}\right)\overline{\rho}\right|_1 . \end{aligned}$$

From the spectral information of the Perron-Frobenius operator and Sublemma 13 we conclude that

$$\begin{aligned} D_j(\overline{\varphi}, \overline{\eta}; \widetilde{\nu}) &\le C|\overline{\eta}|_\infty \,||\overline{\varphi}\,\overline{\rho}||\, \tau_0^j \\ &\le C|\overline{\eta}|_\infty \max\{||\overline{\rho}||, |\overline{\rho}|_\infty\} \max\{||\overline{\varphi}||, |\overline{\varphi}|_\infty\}\tau_0^j. \end{aligned}$$

Now replace both $\overline{\varphi}$ and $\overline{\eta}$ by $\overline{\psi}_N$ and note that $||\overline{\psi}_N|| \le C(\psi)\beta^{-N}$. It follows that (i) holds.

As for the proof of (ii) let $\widetilde{\eta}$ be the lift of a function $\eta \in \mathcal{H}_\gamma$, let $\overline{\varphi} \in X$ and consider

$$\begin{aligned} D_j(\overline{\varphi}, \widetilde{\eta}; \widetilde{\nu}) &= \left|\int (\widetilde{\eta} \circ F^j)\overline{\varphi}\, d\widetilde{\nu}\right| = \left|\int (\widetilde{\eta} \circ F^k) \circ F^{j-k}\overline{\varphi}\, d\widetilde{\nu}\right| \\ &\le \left|\int [\widetilde{\eta} \circ F^k - \overline{\eta}_k] \circ F^{j-k}\overline{\varphi}\, d\widetilde{\nu}\right| + \left|\int \overline{\eta}_k \circ F^{j-k}\overline{\varphi}\, d\widetilde{\nu}\right| \\ &\le |\overline{\varphi}|_\infty \,||\widetilde{\eta} \circ F^k - \overline{\eta}_k||_{L^1(\widetilde{\nu})} + \left|\int \overline{\eta}_k \mathcal{P}^{j-k}(\overline{\rho}\,\overline{\varphi})\, d\overline{m}\right| \\ &\le C|\overline{\varphi}|_\infty \,||\eta||_{\mathcal{H}_\gamma}\lambda^k + C|\overline{\eta}_k|_\infty \,||\overline{\varphi}||\tau_0^{j-k}. \end{aligned}$$

We have here used Sublemma 14 to estimate $||\widetilde{\eta} \circ F^k - \overline{\eta}_k||_{L^1(\widetilde{\nu})}$.

Finally we choose $k = [\kappa j]$ for a suitable small κ and substitute $\overline{\psi}_N$ for $\overline{\varphi}$ and $\widetilde{\psi}^{(N)}$ for $\widetilde{\eta}$ above. The conclusion of (ii) with suitable choices of β_0 and τ then follows from the estimates $||\psi \circ f^N||_{\mathcal{H}_\gamma} \le K_1^N ||\psi||_{\mathcal{H}_\gamma}$, $K_1 = K_1(f, \gamma)$, and $||\overline{\psi}_N|| \le C(\psi)\beta^{-N}$. □

Proof of Lemma 7. — We have

$$\|T_n - U_n\|_2^2 = \frac{1}{n}\left[n \left\| \widetilde{\psi} \circ F^N - \overline{\psi}_N \right\|_2^2 + \sum_{j=1}^{n-1} (n-j) \int \left(\widetilde{\psi}^{(N)} - \overline{\psi}_N\right)\left(\widetilde{\psi}^{(N)} - \overline{\psi}_N\right) \circ F^j d\widetilde{\nu} \right].$$

We pick $j_0 = j_0(\varepsilon)$. The exact choice of j_0 will be made later. For $1 \le j \le j_0$ we estimate the covariances by Cauchy's inequality

$$\left| \int \left(\widetilde{\psi}^{(N)} - \overline{\psi}_N\right)\left(\widetilde{\psi}^{(N)} - \overline{\psi}_N\right) \circ F^j d\widetilde{\nu} \right| \le \left\| \widetilde{\psi}^{(N)} - \overline{\psi}_N \right\|_2 \left\| \widetilde{\psi}^{(N)} - \overline{\psi}_N \right\|_2 \le \varepsilon \cdot \varepsilon = \varepsilon^2.$$

For j in the range $j_0 < j \le n-1$ we use the estimates of Sublemma 15. Combining these estimates we obtain

$$\|U_n - T_n\|_2^2 \le \varepsilon^2(1 + j_0) + \frac{4C}{\beta_0^{2N}} \sum_{j=j_0+1}^{n-1} \frac{1}{n}(n-j)\, e^{-\theta j}$$

with $\tau = e^{-\theta}$.

The last sum is estimated as

$$\frac{1}{n} \sum_{j=j_0+1}^{n-1} (n-j) e^{-\theta j} \le e^{-\theta j_0}.$$

Hence $\|U_n - T_n\|_2^2 \le (1 + j_0)\varepsilon^2 + (4C/\beta_0^{2K_0 \log 1/\varepsilon}) e^{-\theta j_0}$. By choosing

$$j_0(\varepsilon) = 4K_0 \frac{1}{\theta} \log \frac{1}{\beta_0} \log \frac{1}{\varepsilon}$$

we obtain the estimate

$$\|U_n - T_n\|_2^2 \le r(\varepsilon) = \mathcal{O}\left(\varepsilon^2 \log \frac{1}{\varepsilon}\right).$$

□

Proof of Theorem 4. — We will prove the Central Limit Theorem for $\psi \in \mathcal{H}_\gamma$. More specifically, we will show that $F_n(t) \to \mathcal{N}(0, \sigma)$ in distribution, where

$$F_n(t) = \nu\left(\left\{ z : \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \psi \circ f^j \le t \right\}\right)$$

$$\sigma^2 = \sum_{j=0}^{\infty} \int \psi\, \psi \circ f^j \, d\nu \ge 0.$$

We will in the following assume $\sigma > 0$. It is in fact true that $\sigma = 0$ iff $\psi = \varphi - \varphi \circ f$ for some function $\varphi \in L^2$. (This fact was communicated to us by Bill Parry.) Hence the CLT is true for $\sigma = 0$ with the Normal Distribution Function $\Phi_\sigma(t)$ interpreted as the unit step function.

We shall first see how we can use Gordin's theorem to conclude that the CLT holds for test functions from the class X.

Let $\overline{\varphi} \in X$ with $\int \overline{\varphi}\,\overline{\rho}\, d\overline{m} = 0$. The spectral properties of the Perron-Frobenius operator $\mathcal{P}$ guaranties that

$$||\mathcal{P}^j(\overline{\varphi}\,\overline{\rho})|| \le C\tau^j \qquad \forall j \ge 0.$$

Then $|\mathcal{P}^j(\overline{\varphi}\,\overline{\rho})|_1 \le C\tau^j \ \forall j \ge 0$ and from (**) it follows that condition (*) in Gordin's theorem is valid. We conclude that

$$(***) \qquad \widetilde{\nu}\left(\left\{z : \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} \overline{\varphi} \circ F^j(z) \le t\right\}\right) \xrightarrow{\text{dist}} \mathcal{N}(0,\sigma).$$

Note that from the invariance of the measure $\widetilde{\nu}$ under F it follows immediately that

$$\sigma^2(\widetilde{\psi}^{(N)}) = \mathrm{Var}[T_n] = \mathrm{Var}[S_n(\widetilde{\psi}^{(N)})] = \mathrm{Var}[S_n(\widetilde{\psi})] = \sigma^2(\widetilde{\psi}).$$

We will also use the notation $\sigma_\varepsilon^2 = \sigma^2(\overline{\psi}_N) = \mathrm{Var}[U_n]$.

From Lemma 7 we conclude that

$$\sigma^2\left(\overline{\psi}_{N(\varepsilon)}\right) \to \sigma^2\left(\widetilde{\psi}\right) \qquad \text{as} \qquad \varepsilon \to 0$$

and hence we can assume that $\sigma_\varepsilon > 0$ by chosing ε sufficiently small.

By moving up to the tower Δ and using the invariance of $\widetilde{\nu}$ under F we can also write

$$F_n(t) = \widetilde{\nu}\left(\left\{z : \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} \widetilde{\psi}^{(N)} \circ F^j(z) \le t\right\}\right).$$

We wish to compare $F_n(t)$ with the distribution function

$$G_n(t) = \widetilde{\nu}\left(\left\{z : \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} \overline{\psi}_N \circ F^j(z) \le t\right\}\right).$$

It follow from (***) with $\overline{\varphi}$ replaced by $\overline{\psi}_N$ that $G_n(t) \to \mathcal{N}(0,\sigma_\varepsilon)$ in distribution. Now pick $\eta > 0$. By Tjebyshev's inequality

$$\begin{aligned} F_n(t) &\le G_n(t+\eta) + \widetilde{\nu}\{|T_n - U_n| \ge \eta\} \\ &\le G_n(t+\eta) + \frac{1}{\eta^2}\,||T_n - U_n||_2^2 \\ &\le G_n(t+\eta) + \frac{1}{\eta^2} r(\varepsilon). \end{aligned}$$

Letting $n \to \infty$ we conclude that $\overline{\lim_{n\to\infty}} F_n(t) \le \Phi_{\sigma_\varepsilon}(x+\eta) + \frac{1}{\eta^2} r(\varepsilon)$. Here

$$\Phi_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-u^2/2\sigma^2} du.$$

By letting $\varepsilon \to 0$ we have

$$\overline{\lim_{n\to\infty}} F_n(t) \le \Phi_\sigma(x+\eta).$$

Now let $\eta \to 0$. It follows that $\overline{\lim}_{n\to\infty} F_n(t) \le \Phi_\sigma(x)$. The proof that

$$\underline{\lim}_{n\to\infty} F_n(t) \ge \Phi_\sigma(x)$$

is completely analogous. □

References

[B] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphism*, volume 470. Springer Lecture Notes in Math., 1975.

[BC1] M. Benedicks and L. Carleson. On iterations of $1-ax^2$ on $(-1,1)$. *Ann. Math.*, 122:1–25, 1985.

[BC2] M. Benedicks and L. Carleson. The dynamics of the Hénon map. *Ann. Math.*, 133:73–169, 1991.

[BY] M. Benedicks and L.-S. Young. Sinai-Bowen-Ruelle measures for certain Hénon maps. *Invent. Math.*, 112:541–576, 1993.

[C] N. I. Chernov. Ergodic and statistical properties of piecewise linear hyperbolic automorphisms of the 2-torus. *J. of Stat. Phys.*, 69:111–134, 1992.

[G] M. I. Gordin. The central limit theorem for stationary processes. *Soviet Math. Dokl.*, 10(5):1174–1176, 1969.

[K] G. Keller. Un théoreme de la limite centrale pour une classe de transformations monotones par morceaux. *Comptes Rendus de l'Académie des Sciences, Série A*, 291:155–158, 1980.

[L] F. Ledrappier. Propriétés ergodiques des mesures de Sinai. *Publ. Math. IHES*, 59, 1984.

[PS] C. Pugh and M. Shub. Ergodic attractors. *Trans. Amer. Math. Soc.*, 312:1–54, 1989.

[R] D. Ruelle. *Thermodynamic formalism*. Addison-Wesley, New-York, 1978.

[Y] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. To appear in *Ann. Math.*

M. Benedicks, Department of Mathematics, Royal Institute of Technology, S-100 44 Stockholm, Sweden • *E-mail :* `michaelb@math.kth.se`

L.-S. Young, Department of Mathematics, University of California Los Angeles, Los Angeles, CA 90024, USA • *E-mail :* `lsy@math.ucla.edu`

Astérisque

CÉSAR CAMACHO
BRUNO AZEVEDO SCÁRDUA

Complex foliations with algebraic limit sets

Astérisque, tome 261 (2000), p. 57-88
<http://www.numdam.org/item?id=AST_2000__261__57_0>

Astérisque
261, 2000, p. 57–88

COMPLEX FOLIATIONS WITH ALGEBRAIC LIMIT SETS

by

César Camacho & Bruno Azevedo Scárdua

Dedicated to Adrien Douady on the occasion of his $60^{\underline{th}}$ *birthday.*

Abstract. — We regard the problem of classification for complex projective foliations with algebraic limit sets and prove the following:

Let $\mathcal{F}$ be a holomorphic foliation by curves in the complex projective plane $\mathbb{C}P(2)$ having as limit set some singularities and an algebraic curve $\Lambda \subset \mathbb{C}P(2)$. If the singularities $\operatorname{sing}\mathcal{F} \cap \Lambda$ are generic then either $\mathcal{F}$ is given by a closed rational 1-form or it is a rational pull-back of a Riccati foliation $\mathcal{R} : p(x)dy - (a(x)y^2 + b(x)y)dx = 0$, where Λ corresponds to $\overline{(y=0)} \cup \overline{(p(x)=0)}$, on $\overline{\mathbb{C}} \times \overline{\mathbb{C}}$.

The proof is based on the solvability of the generalized holonomy groups associated to a reduction process of the singularities $\operatorname{sing}\mathcal{F} \cap \Lambda$ and the construction of an affine transverse structure for $\mathcal{F}$ outside an algebraic curve containing Λ.

1. Introduction

Let $\mathcal{F}$ be a holomorphic codimension one foliation on the complex projective 2-space $\mathbb{C}P(2)$. Given any leaf L of $\mathcal{F}$ the *limit set* of L is defined as $\lim(L) = \bigcap_\nu \overline{L \backslash K_\nu}$ where $K_\nu \subset K_{\nu+1}$ is an exhaustion of L by compact subsets $K_\nu \subset L$. The *limit set* of the foliation $\mathcal{F}$ is defined as $\lim \mathcal{F} = \overline{\bigcup_L \lim(L)}$. We are interested in classifying those foliations whose limit set is a union of singularities of $\mathcal{F}$ and an algebraic curve $\Lambda \subset \mathbb{C}P(2)$. There are two reasons for this, first because these foliations exhibit the simplest dynamic behavior we can imagine and also because they must support an important class of first integrals. The parallel with the actions of Kleinian groups on the Riemann sphere comes naturally to mind. These foliations will correspond to actions with a finite set of limit points (one or two) while the first integrals of these foliations will correspond to the automorphic functions of such Kleinian group

1991 ***Mathematics Subject Classification.*** — 32L30, 58F18.
Key words and phrases. — Holomorphic foliation, limit set, holonomy.

B. Scárdua was supported by the CNPq-Brasil and the Ministère Affaires Étrangères (France).

actions. Here we will show that this similarity is not only apparent. Indeed the Kleinian groups will appear naturally as the holonomy groups of the Riccati foliation that, it will be shown here, is the ultimate model for these foliations.

The problem of classifying such foliations $\mathcal{F}$ was considered in **[1]** and **[17]**. In both cases it is proved that, under generic assumptions, there are a rational map $F\colon \mathbb{C}P(2) \to \mathbb{C}P(2)$ and a linear foliation $\mathcal{L} : \lambda_1 x dy - \lambda_2 y dx = 0$ on $\mathbb{C}P(2)$ such that $\mathcal{F} = F^*(\mathcal{L})$. In particular, it follows that no saddle-nodes appear in the resolution of $\operatorname{sing} \mathcal{F} \cap \Lambda$, and in fact all the singularities as well as all the holonomy groups appearing in this resolution are abelian and linearizable. Using **[9]** we can construct examples where $\mathcal{F}$ is a *Riccati foliation* with algebraic limit set on $\mathbb{C}P(2)$, containing the invariant line $\overline{(y=0)}$:

$$\mathcal{F} : p(x)dy - (y^2 a(x) + y b(x))dx = 0$$

where $a(x), b(x), p(x)$ are polynomials, and $(x, y) \in \mathbb{C}^2 \subset \mathbb{C}P(2)$ is an affine chart (see Example 1.3 below) and $\Lambda \cap \mathbb{C}^2 = (p(x) = 0) \cup (y = 0)$.

In the Riccati case, the holonomy groups are solvable and we have an additional compatibility condition as in **[2]**. However we may have saddle-nodes in the resolution of $\operatorname{sing} \mathcal{F} \cap \Lambda$. The aim of this paper is to solve the problem above in the case the foliation may have certain saddle-node singularities in its resolution along Λ.

Let therefore $\mathcal{F}$ be a foliation on $\mathbb{C}P(2)$ and let $\Lambda \subset \mathbb{C}P(2)$ be an algebraic invariant curve (perhaps reducible). We will say that $\operatorname{sing} \mathcal{F} \cap \Lambda$ *has* the pseudoconvexity property (*psdc*) if the *invariant (by $\mathcal{F}$) part* Γ of the resolution divisor D of $\operatorname{sing} \mathcal{F} \cap \Lambda$ is connected and its complement is a *Stein manifold* (alternatively, Γ is a very ample divisor on the ambient (algebraic) manifold of the resolution of $\operatorname{sing} \mathcal{F} \cap \Lambda$ denoted by $\widetilde{\mathbb{C}P(2)}$), so that we can apply Levi's extension theorem **[21]** which allows us to extend analytically to all $\mathbb{C}P(2)$, any analytic object defined on a neighborhood of Γ. This property is verified if $\mathcal{F}$ has no dicritical singularities over Λ **[1]**. There is another remarkable case where property (*psdc*) is verified, as we can find in **[17]**. A singularity $q_o \in \operatorname{sing} \widetilde{\mathcal{F}} \cap D$ is a *corner* if $q_o = D_i \cap D_j$, where $D_i \neq D_j$ are invariant components of D.

Also, we say that a saddle-node singularity $q_o \in D$ is in *good position* relatively to D, if its *strong separatrix* is contained in some component of Γ. A saddle-node $x^{k+1} dy - y(1 + \lambda x^k)dx + \text{h.o.t.} = 0$ is *analytically normalizable* if we may choose local coordinates (x, y) as above for which we have $\text{h.o.t.} = 0$. In this case it will be called *normally hyperbolic* if we have $\lambda \notin \mathbb{Q}$. In this case we call $(x = 0)$ the *strong separatrix* and $(y = 0)$ the *central manifold* of the saddle-node. We recall that according to **[12]** a saddle-node singularity is analytically classified by the local holonomy of this strong separatrix. In particular, the saddle-node is analytically normalizable if, and only if, its strong separatrix holonomy is an analytically normalizable flat diffeomorphism.

Finally, we introduce the following technical condition (see Example 1.4):

(C_1) *The saddle-nodes in the resolution of* $\operatorname{sing}\mathcal{F} \cap \Lambda$ *are analytically normalizable, and the ones in the corners are normally hyperbolic.*

Our main result is the following:

Theorem 1.1. — *Let* $\mathcal{F}$ *be a codimension one holomorphic foliation on* $\mathbb{C}P(2)$ *having as limit set some singularities and an algebraic curve* $\Lambda \subset \mathbb{C}P(2)$. *Assume that* $\operatorname{sing}\mathcal{F} \cap \Lambda$ *satisfies property (psdc) and condition* C_1. *Then, either* $\mathcal{F}$ *is given by a closed rational* 1*-form or it is a rational pull-back of a Riccati foliation* $\mathcal{R} : p(x)dy - (a(x)y^2 + b(x)y)dx = 0$, *where* Λ *corresponds to* $\overline{(y=0)} \cup \overline{(p(x)=0)}$, *on* $\overline{\mathbb{C}} \times \overline{\mathbb{C}}$.

The proof of this theorem relies on the study of the *singular* and *virtual* holonomy groups **[2]**, **[5]**, **[19]** and **[1]** respectively, of the irreducible components of the divisor given by the resolution of $\operatorname{sing}\mathcal{F}\cap\Lambda$. *The limit set of the leaves* $\widetilde{L}$ *of* $\widetilde{\mathcal{F}}$ *induces discrete pseudo-orbits in each of these groups, so that they are solvable* **[14]**. The solvability of these groups, allows (under our restrictions on $\operatorname{sing}\mathcal{F} \cap \Lambda$) the construction of a "transversely formal" meromorphic 1-form $\widetilde{\eta}$, defined over the invariant part Γ of the resolution divisor of $\operatorname{sing}\mathcal{F} \cap \Lambda$. This 1-form is closed and satisfies the relation $d\widetilde{\omega} = \widetilde{\eta} \wedge \widetilde{\omega}$, where $\widetilde{\omega}$ is a meromorphic 1-form with isolated singularities which defines the foliation $\widetilde{\mathcal{F}}$, obtained from the resolution of $\operatorname{sing}\mathcal{F} \cap \Lambda$. Moreover, $\widetilde{\eta}$ has (simple) poles over Γ which coincides with the limit set of $\widetilde{\mathcal{F}}$. Using a result of Hironaka-Matsumara (see **[5]**, **[8]**), we conclude that (since $\widetilde{\mathbb{C}P(2)}\backslash\Gamma$ is a Stein manifold) the 1-form $\widetilde{\eta}$ is in fact rational on $\widetilde{\mathbb{C}P(2)}$. This corresponds to the existence of a Liouvillian first integral for $\mathcal{F}$ on $\mathbb{C}P(2)$, and also to the existence of an affine transverse structure for $\widetilde{\mathcal{F}}$ in $\widetilde{\mathbb{C}P(2)}\backslash C$, where $C \subset \widetilde{\mathbb{C}P(2)}$ is an algebraic invariant curve containing $\widetilde{\Lambda}$, where $\widetilde{\Lambda}$ is the strict transform of Λ, **[18]**. This affine transverse structure can be extended as a projective transverse structure to $(\widetilde{\mathbb{C}P(2)}\backslash C) \cup \widetilde{\Lambda}$. In particular, all the singular holonomy groups associated to the components of Γ are solvable analytically normalizable. This implies by (a careful reading of the last part of) **[2]** that either $\mathcal{F}$ is given by a closed rational 1-form or by a rational pull-back of a Riccati foliation.

Example 1.2. — Let $\mathcal{F}$ be a rational pull-back of a hyperbolic linear foliation $\mathcal{L}$: $xdy - \lambda ydx = 0, \lambda \in \mathbb{C}\backslash\mathbb{R}$, on $\mathbb{C}P(2)$. Clearly $\mathcal{F}$ has an algebraic limit set consisting of some singularities and an algebraic curve Λ as in Theorem 1.1.

Example 1.3. — Let us take any finitely generated group of Moebius transformations $G \subset \mathrm{SL}(2,\mathbb{C})$. Assume that the limit set of G is a single point, which can be assumed to be the origin $0 \in \overline{\mathbb{C}}$. The limit point 0 is a fixed point of G. According to **[9]** we can find a Riccati foliation $\mathcal{F} : p(x)dy - (a(x)y^2 + b(x)y + c(x))dx = 0$ on $\overline{\mathbb{C}} \times \overline{\mathbb{C}}$, whose holonomy group of the line $\overline{(y=0)}$ is conjugated to the group G. Moreover we can assume that the singularities of $\mathcal{F}$ over this horizontal line are reduced and non degenerate. The line $\overline{(y=0)}$ is invariant by $\mathcal{F}$ so that $c(x) = 0$, and also it is contained in the limit set of $\mathcal{F}$ and satisfies condition C_1 in the statement above. This example can also be seen in $\mathbb{C}P(2)$ using a birational transformation. This will create a dicritical singularity. This example will satisfy the (*psdc*) property for a proper choice of Λ.

Example 1.4. — This is a counterexample to a more general statement. Let $\mathcal{F}$ be given by $\omega = dy - (a(x)y + b(x))dx = 0$ over $\mathbb{C}^2 \subset \mathbb{C}P(2)$. If we consider the vector field $X(x,y) = (1, a(x)y + b(x))$, then X is complete and tangent to $\mathcal{F}$ over $\mathbb{C}^2$. Moreover the orbits of X are diffeomorphic to $\mathbb{C}$. It is not difficult to see, using the flow of X, that the leaves of $\mathcal{F}$ accumulate the line at infinity $L_\infty = \mathbb{C}P(2)\backslash\mathbb{C}^2$, so that $\lim \mathcal{F} = L_\infty$. However, generically, the resolution of $\operatorname{sing}\mathcal{F} \cap L_\infty$ exhibits some non analytically normalizable saddle-node. Indeed, this resolution is quite simple and shows that there are saddle-nodes with non convergent central manifolds [5]. On the other hand, in general, $\mathcal{F}$ is not a rational pull-back of a Riccati foliation of the form stated in Theorem 1.1.

Acknowledgements. — Part of this work was conceived during a post-doctorade stage of the second author, at the Université de Rennes I. He wants to thank the IRMAR and specially D. Cerveau for the kind hospitality and for valuable conversations.

2. Formal normal forms and resolution of singularities

Let $\mathcal{F}$ and $\Lambda \subset \lim \mathcal{F}$ be as in Theorem 1.1. Let $\pi\colon (\widetilde{\mathbb{C}P(2)}, \widetilde{\mathcal{F}}, D) \to (\mathbb{C}P(2), \mathcal{F}, \Lambda)$ be the resolution morphism of Seidenberg, for $\operatorname{sing}\mathcal{F} \cap \Lambda$ [20]. Thus $\widetilde{\mathbb{C}P(2)}$ is a compact complex surface which is obtained from $\mathbb{C}P(2)$ by a finite sequence of blowing-up's, denoted π. The proper morphism π induces therefore a foliation by curves $\widetilde{\mathcal{F}} = \pi^*\mathcal{F}$ on $\widetilde{\mathbb{C}P(2)}$. The divisor $D = \pi^{-1}(\Lambda)$ of the resolution is a finite union $D = \bigcup_{j=0}^m D_j$, of projective lines $D_j \cong \mathbb{C}P(1)$, $j \neq 0$, and of the strict transform of Λ, $D_0 = \overline{\pi^{-1}(\Lambda\backslash \operatorname{sing}\mathcal{F})}$. The foliation $\widetilde{\mathcal{F}}$ has singularities of the following two types (called *irreducible* singularities):

(i) $xdy - \lambda ydx + \text{h.o.t.} = 0$ (non degenerate)
(ii) $y^{p+1}dx - [x(1 + \lambda y^p) + \text{h.o.t.}]dy = 0$ (saddle-node).

We consider the foliation $\widetilde{\mathcal{F}} = \pi^*\mathcal{F}$ and denote by Γ the invariant (by $\widetilde{\mathcal{F}}$) part of D, which consists of the invariant projective lines and of the strict transform of Λ. Let ω be a rational 1-form which defines $\mathcal{F}$ on $\mathbb{C}P(2)$ and denote by $\widetilde{\omega}$ the strict transform of $\pi^*\omega$. Therefore the 1-form $\widetilde{\omega}$ has isolated singularities and we can assume that its polar set intersects the divisor D transversely and at regular points of $\widetilde{\mathcal{F}}$. Clearly we have $\lim(\widetilde{\mathcal{F}}) \subset \Gamma$.

Lemma 2.1. — *We have* $\lim(\widetilde{\mathcal{F}}) = \Gamma$. *In particular all the saddle-nodes in* $\operatorname{sing}\widetilde{\mathcal{F}} \cap \Gamma$ *are in good position with respect to* Γ.

Proof. — Recall that by hypothesis Γ is connected. Let us fix a saddle-node $q_o \in D_j$, which is not a corner. Assume by contradiction that the strong separatrix S of q_o is not contained in D_j. We consider the local of S around q_o at a small transverse disk $\Sigma \cong \mathbb{D}$, with $\Sigma \cap S = q_1 \notin \operatorname{sing} \widetilde{\mathcal{F}}$. This holonomy map $h_o\colon (\Sigma, q_1) \to (\Sigma, q_1)$ is a flat local diffeomorphism, that is, a local diffeomorphism tangent to the identity. Thus, the orbits of h_o accumulate the origin q_1, so that the local leaves of $\widetilde{\mathcal{F}}$ around q_o and crossing Σ, must accumulate the strong separatrix S, and therefore we obtain $S \subset \lim \mathcal{F}$ (notice that S is transverse to D, so that it corresponds to a separatrix of $\mathcal{F}$ not contained in Λ), which contradicts the hypothesis that $\lim \mathcal{F} = \Lambda$. Now, we fix a saddle-node on a corner $q_o = D_i \cap D_j$. First we prove that if $\lim \widetilde{\mathcal{F}}$ contains the central manifold of q_o, say D_i then it contains the strong manifold, in this case D_j. In fact, by the hypothesis we may write $\widetilde{\mathcal{F}}$ as

$$y(1+\lambda x^k)dx - x^{k+1}dy = 0, \quad D_i = (y=0), \quad D_j = (x=0).$$

Now, in a sector $(x,y) \in U \times \mathbb{C}$ near $0 \in \mathbb{C}^2$, where $U = \{x \in \mathbb{C}^*;\ \operatorname{Re}(x^p) > 0\}$ the leaves of $\widetilde{\mathcal{F}}$ have a saddle-like behavior in the sense that there are sections $\Sigma_i = (x=1)$ and $\Sigma_j = (y=1)$, such that any leaf $\widetilde{L}$ of $\widetilde{\mathcal{F}}\big|_{U\times\mathbb{C}}$, not contained in $(y=0)$, is at a positive distance from $0 \in \mathbb{C}^2$ and, if we denote $r_\nu = \Sigma_\nu \cap \widetilde{L}$, we have: if $r_i \to (0,1)$ then $r_j \to (1,0)$.

Now we prove the converse: If $\lim \widetilde{\mathcal{F}}$ contains the central manifold of q_o, D_i, then it contains the strong manifold, D_j. In fact, using the normal form above we may conclude that the local holonomy of the central manifold around q_o, is linearizable of the form $h\colon (\Sigma_i, (1,0)) \to (\Sigma_i, (1,0)), h(y) = \exp(2\pi i\lambda)\cdot y$. If $\lambda \in \mathbb{R}\backslash\mathbb{Q}$, then it is a non rational rotation so that the accumulations of the leaves in the section Σ_i do not correspond to algebraic limit sets. Thus $\lambda \in \mathbb{C}\backslash\mathbb{R}$, and therefore, either h or h^{-1} is an attractor, so that any leaf which intersects Σ_i accumulates the origin $(0,1) \in \Sigma_i$.

We also remark that $\lim \widetilde{\mathcal{F}}$ contains all the strong separatrices of the saddle-nodes in Γ. In fact, from the analytic normal form above we have a multivalued first integral $f(x,y) = (y/x^\lambda)\exp(1/kx^k)$. This first integral shows that the leaves accumulate on the strong manifold $(x=0)$. Also from the same arguments of **[1]**, **[17]** we have (in the non degenerated corners) the passage of the limit set $\lim(\widetilde{L})$ from one to the other adjacent component of D: It is in fact, only necessary to use the fact that if q_o is a non degenerate corner say of the form, $xdy - \lambda ydx + \text{h.o.t.} = 0$ such that $\lambda \in \mathbb{R}_+\backslash\mathbb{Q}_+$, then by Poincaré Linearization Theorem this singularity is linearizable, and therefore it is not difficult to see that the local leaves around q_o are not proper. On the other hand, in the case $\lambda \in \mathbb{R}_-$, any leaf which accumulates q_o and which is not a separatrix, accumulates both separatrices. Finally, we remark that since by the hypothesis the limit set of $\mathcal{F}$ is algebraic, it follows that all the strong manifolds in $\operatorname{sing}\widetilde{\mathcal{F}} \cap \Gamma$ are contained in Γ, that is, the saddle-nodes are in good position relatively to Γ. □

Now we fix local transverse sections $\mathcal{S}_j$, $\mathcal{S}_j \cap D_j = p_j \notin \operatorname{sing} \widetilde{\mathcal{F}}$, $(\mathcal{S}_j, p_j) \cong (\mathbb{C}, 0)$. Let us write G_j for the holonomy group $\operatorname{Hol}(\widetilde{\mathcal{F}}, D_j, \mathcal{S}_j)$ of $D_j \subset \Gamma$ (see **[2]**). The following definition is found in **[1]**:

Definition 2.2. — The *virtual holonomy group* of $\widetilde{\mathcal{F}}$ relative to the component D_j at the section $\mathcal{S}_j$ is defined as

$$\operatorname{Hol}^v(\widetilde{\mathcal{F}}, D_j, \mathcal{S}_j) = \{f \in \operatorname{Diff}(\mathcal{S}_j, p_j) \mid \widetilde{L}_z = \widetilde{L}_{f(z)},\ \forall\, z \in (\mathcal{S}_j, p_j)\}$$

Clearly this group contains the *holonomy group* of $\widetilde{\mathcal{F}}$ relative to D_j at the section $\mathcal{S}_j$, denoted by $\operatorname{Hol}(\widetilde{\mathcal{F}}, D_j, \mathcal{S}_j)$ (see **[2]** for the definition of the holonomy group). Let us write G_j^v for the virtual holonomy group $\operatorname{Hol}^v(\widetilde{\mathcal{F}}, D_j, \mathcal{S}_j)$.

We will write *projective* holonomy group to denote the holonomy group of any component D_j of D.

We denote by $\operatorname{Diff}(\mathbb{C}, 0)$ respectively $\widehat{\operatorname{Diff}}(\mathbb{C}, 0)$ the group of germs of biholomorphisms respectively the group of formal biholomorphisms of $(\mathbb{C}, 0)$. We also denote by $\mathcal{X}(\mathbb{C}, 0)$, respectively $\widehat{\mathcal{X}}(\mathbb{C}, 0)$ the Lie algebra of the germs of singular holomorphic vector fields at $0 \in \mathbb{C}$, respectively the Lie algebra of singular formal vector fields in one complex variable.

According to Lemma 2.1 the limit set of any non algebraic leaf $\widetilde{L}$ induces discrete pseudo-orbits in each projective virtual holonomy group. These groups are solvable as a consequence of the following result due to I. Nakai:

Proposition 2.3 (**[14]**). — *Let $G \subset \operatorname{Diff}(\mathbb{C}, 0)$ be a subgroup which has some discrete pseudo-orbit. Then G is solvable.*

Corollary 2.4. — *Let $\mathcal{F}$ be as in Theorem 1.1. Then each projective or virtual holonomy group of* $\operatorname{sing} \mathcal{F} \cap \Lambda$ *is solvable.*

We also have the following result concerning subgroups with discrete pseudo-orbits:

Theorem 2.5 (**[10]**). — *Let $G \subset \operatorname{Diff}(\mathbb{C}, 0)$ be a nonabelian subgroup with discrete pseudo-orbits outside the origin. Then G is either formally conjugate to some group*

$$G_\nu^2 := \langle z \mapsto az, z \mapsto z/(1 + z^\nu)^{1/\nu} \rangle$$

where a^ν has order 2; or it is analytically conjugate to some group

$$G_{\nu,\tau}^2 := \langle z \mapsto az, z \mapsto z/(1 + z^\nu)^{1/\nu}, z \mapsto z/(1 + \tau z^\nu)^{1/\nu} \rangle$$

where a^ν has order 2 and $\tau \in \mathbb{C} \setminus \mathbb{R}$; or finally it is analytically conjugate to some group

$$G_\nu^n := \langle z \mapsto az, z \mapsto z/(1 + z^\nu)^{1/\nu} \rangle$$

where a^ν has order $n \in \{3, 4, 6\}$.

We shall consider the subgroups

$$\mathbb{H}_k = \left\{\varphi \in \text{Diff}(\mathbb{C},0) \mid \varphi(z)^k = \frac{\mu_\varphi z^k}{1+a_\varphi z^k}, \quad \mu_\varphi \in \mathbb{C}^*, \ a_\varphi \in \mathbb{C}\right\}$$

where $k \in \mathbb{N}^*$. According to **[4]** any solvable non abelian subgroup of $\text{Diff}(\mathbb{C},0)$ is formally conjugated to a subgroup of some $\mathbb{H}_k$, this conjugacy is analytic except for some special case. We also use the following result:

Lemma 2.6 (**[15]**). — *Let $G \subset \widehat{\text{Diff}}(\mathbb{C},0)$ be a subgroup. Then:*

(i) G is abelian if, and only if, there exists a formal vector field $\xi \in \widehat{\mathcal{X}}(\mathbb{C},0)$ such that $g * \xi = \xi$, $\forall\, g \in G$. In the case G is not linearizable the vector field ξ is unique.

(ii) G is solvable non abelian if, and only if, there exists a formal vector field $\xi \in \widehat{\mathcal{X}}(\mathbb{C},0)$ such that $g * \xi = c_g \cdot \xi$, $c_g \in \mathbb{C}^*$, $\forall\, g \in G$, where $c_g \neq 1$ for some $g \in G$. The vector field ξ is unique up to multiplicative constants.

As it is well-known **[11]**, given a formal vector field $\xi = a(z)d/dz$, with $\xi(0) = 0$, there exists a formal diffeomorphism $\Theta \in \widehat{\text{Diff}}(\mathbb{C},0)$, such that

$$\Theta_*\xi(z) = \xi_{\lambda,k}(z) := \frac{z^{k+1}}{1+\lambda z^k},$$

where $k \in \mathbb{N}$ and $\lambda \in \mathbb{C}$ are formal invariants associated to ξ. It is clear that if $\varphi \in \widehat{\text{Diff}}(\mathbb{C},0)$ satisfies $\varphi_*\xi_{0,k}(z) = \xi_{0,k}(z)$, then $\varphi(z) \in \mathbb{H}_k$. On the other hand it is not difficult to see that if $G \subset \widehat{\text{Diff}}(\mathbb{C},0)$ is solvable and non abelian, then the vector field ξ given by Lemma 2.6 above must exhibit $\lambda = 0$ **[15]**. Therefore we have the following definition:

Definition 2.7. — Let $G \subset \widehat{\text{Diff}}(\mathbb{C},0)$ be a solvable subgroup. A formal *normalizing* coordinate $w \in (\mathbb{C},0)$ for G is anyone for which the vector field ξ of Lemma 2.6 above writes as $\xi(w) = \xi_{\lambda,k}(w)$.

Clearly, if G is solvable and non abelian, then the formal normalizing coordinate is unique up to composition with elements $\varphi \in \mathbb{H}_k$, where k is given by G as above.

If G is abelian then given two normalizing coordinates u and w with $u'(0) = w'(0)$, we have $u = \varphi(w)$, where $\varphi(z) = \exp t\xi_{\lambda,k}(z)$ for some $t \in \mathbb{C}$.

The group G is called *analytically normalizable* if the associated vector field ξ is convergent, otherwise we will say that G is *non analytically normalizable.* In other words, a solvable (perhaps abelian) subgroup $G \subset \text{Diff}(\mathbb{C},0)$ is analytically normalizable if it is analytically conjugated to its formal model.

Proposition 2.8. — *Let $\mathcal{F}, \Lambda$ be as in Theorem 1.1. Denote by D the resolution divisor of* $\operatorname{sing}\mathcal{F} \cap \Lambda$. *Let D_j be a component of D. Assume that the holonomy of the component D_j is solvable non abelian $G_j \subset \mathbb{H}_k$ by a formal conjugation. Then, given a singularity $q_o \in \operatorname{sing}\widetilde{\mathcal{F}} \cap D_j$ there exists a formal diffeomorphism $\Phi \in \widehat{\text{Diff}}(\widetilde{\mathbb{C}P(2)}, q_o)$, such that $\Phi^*(\mathcal{F})$ has one of the following normal forms where $\Phi^*(D_j) = (y=0)$:*

(a) $\omega_\lambda = xdy - \lambda ydx$, $\lambda \in \mathbb{C}\backslash\mathbb{Q}$ *(q_o is formally linearizable non resonant);*
(b) $\omega_{n/m} = nxdy + mydx$, $n, m \in \mathbb{N}$, $\langle n, m\rangle = 1$ *(q_o has a formal first integral);*
(c) $\omega_{k,\ell} = kxdy + \ell y(1 + \frac{\sqrt{-1}}{2\pi}x^\ell y^k)dx$ *(q_o is resonant non formally linearizable);*
(d) $\omega_k = y^{k+1}dx - xdy$ *(q_o is a saddle-node with strong manifold tangent to D_j);*
(e) $\omega^{p,\lambda} = x^{p+1}dy - y(1 + \lambda x^p)dx$ *(q_o is a saddle-node with strong manifold transverse to D_j).*

Moreover, we may assume that Φ converges except for case (c), and that G_j is analytically normalizable except for cases (b) and (c).

Proof. — First we assume that q_o is non degenerate, say

$$\widetilde{\omega}(u, v) = udv - \lambda vdu + \text{h.o.t.}, \qquad \lambda \in \mathbb{C}^*$$

for some local *holomorphic* coordinates (u, v) centered at q_o. If $\lambda \notin \mathbb{Q}$ then it follows that $\widetilde{\omega}$ is formally linearizable at q_o, that is, we have (a). Assume now that $\lambda = -n/m \in \mathbb{Q}_-$ with $n, m \in \mathbb{N}$, $\langle n, m\rangle = 1$. Then we consider the local holonomy $\varphi(v)$ of D_j at q_o. According to the hypothesis on the holonomy group of D_j, there exists a formal change of coordinates $\widehat{\psi} \in \widehat{\text{Diff}}(\mathbb{C}, 0)$ such that

$$\widehat{\psi} \circ \varphi \circ \widehat{\psi}^{-1}(w) = \frac{cw}{(1 + aw^k)^{1/k}}$$

where the linear part is $c = \exp(\frac{2k\pi\sqrt{-1}n}{m})$. On the other hand it is well-known that an homography which is not tangent to the identity is linearizable by another homography. If $kn/m \notin \mathbb{N}$, then $c \neq 1$ and therefore the singularity is therefore formally linearizable as in (b). Assume that q_o is not formally linearizable and (therefore) that $k/n = \ell/m$ for some $\ell \in \mathbb{N}$. Then according to **[11]** there exists a formal conjugacy at q_o which takes $\widetilde{\omega}$ into the form (c).

Assume now that q_o is a saddle-node singularity. If the strong manifold of q_o is tangent to D_j then $c = 1$ in the expression of $\widehat{\psi} \circ \varphi \circ \widehat{\psi}^{-1}(w)$ above and therefore this local holonomy is formally conjugated to the local holonomy of the strong manifold $(v = 0)$ of the saddle-node $v^{k+1}du - udv = 0$. Therefore $\widetilde{\omega}$ must be of the form (d) above. If the strong manifold of q_o is transverse to D_j then its has a formal normal form as in (e) as a consequence of **[12]**.

Now we remark that in case (a) G_j must be analytically normalizable because it contains an element with nonperiodic linear part. In case (d) G_j is analytically normalizable because it contains the holonomy of the strong manifold of ω_k, which is assumed to be analytically normalizable. We remark that in case (e) G_j is again analytically normalizable because, as it follows from Lemma 2.1, q_o is a corner. Indeed, in this case G_j contains the holonomy of the central manifold of $\omega^{p,\lambda}$, which is a nonrational linearizable rotation, and therefore has nonperiodic linear part. Finally we remark that according to **[13]** a singularity q_o has a formal first integral if, and only if, q_o has a holomorphic first integral, so that we can assume that Φ is convergent in case (a). This finishes the proof of the proposition. □

Remark 2.9. — According to **[11]**, **[12]**, **[13]** any reduced singularity admits a formal integrating factor.

3. Virtual holonomy and Singular holonomy

In this section we follow **[5]**, **[19]** and **[2]**. Let us consider the following situation: $\mathcal{F}$ is a foliation on a compact complex surface M, $D \subset M$ is a compact (codimension one) invariant divisor with normal crossings, $D = \bigcup_j D_j$ where the D_j are irreducible smooth components. As in §2 we fix local transverse sections $\mathcal{S}_j$, $\mathcal{S}_j \cap D_j = p_j \notin \operatorname{sing}\mathcal{F}$, $(\mathcal{S}_j, p_j) \cong (\mathbb{C}, 0)$, and write G_j for the holonomy group $\operatorname{Hol}(\mathcal{F}, D_j, \mathcal{S}_j)$ of D_j (see **[2]**). Given any other transverse section Σ_j to $\mathcal{F}$ such that $\Sigma_j \cap D_j = q_j$, there is a conjugacy between $\operatorname{Hol}(\mathcal{F}, D_j, \Sigma_j)$ and $\operatorname{Hol}(\mathcal{F}, D_j, \mathcal{S}_j)$ induced by lifting to the leaves of $\mathcal{F}$ a simple path joining p_j to q_j, in the leaf $D_j \setminus \operatorname{sing}\mathcal{F}$. Thus, up to conjugacy, we can identify these groups and in particular $\operatorname{Hol}(\mathcal{F}, D_j, \Sigma_j)$ is solvable if, and only if, G_j is.

Now we fix a *corner* $q_o = D_i \cap D_j$. We assume that all the virtual the holonomy groups G_ν^v are solvable (perhaps abelian) for $\nu \in \{i, j\}$. In the non abelian case we denote by k_ν the *ramification order* of G_ν^v, so that $G_\nu \subset G_\nu^v \subset \mathbb{H}_{k_\nu}$ by a formal conjugacy.

The following lemma holds in general, *i.e.*, also for non normally hyperbolic saddle-nodes:

Lemma 3.1. — *Assume that the holonomy group G_j is analytically normalizable, and that if q_o is a saddle-node then its strong manifold is contained in D_j. Then $\mathcal{F}$ is analytically normalizable at the singular point q_o.*

Proof. — First we assume that G_j is non abelian, so that there exists a local holomorphic coordinate $z \in \Sigma_j$, $z(q_j) = 0$, where Σ_j is a local transverse section with $\Sigma_j \cap D_j = q_j$ close to q_o, such that the local holonomy of $\mathcal{F}$ due to q_o and relative to D_j writes

$$\varphi(z) = \frac{\lambda z}{(1 + az^k)^{1/k}}.$$

Now, if $\lambda^k \neq 1$ then we can linearize this local holonomy and therefore the singularity q_o, which is not a saddle-node (recall that the holonomy of the strong manifold of a saddle-node is never linearizable). Assume now that we have $\lambda^k = 1$. In this case we have $\varphi(z)^k = z/(1 + az^k)$. If q_o is not a saddle-node then as in the proof of Proposition 2.8 it follows from **[11]** that q_o must be analytically conjugated to a singularity of the form $\omega_{k,\ell}$ as in Proposition 2.8, because the holonomies are analytically conjugated. If q_o is a saddle-node then by the hypothesis the strong manifold is contained in D_j. Therefore, by **[12]**, the analytic normalization of the holonomy of the strong manifold implies the analytic normalization of the singularity q_o. Now we consider the case G_j is abelian and analytically normalizable. According to the techniques of construction of integrating factors (see Lemma 5.4 below or [16]), we can construct a closed meromorphic 1-form Ω_j in a neighborhood of q_o, and which satisfy $\widetilde{\omega} \wedge \Omega_j \equiv 0$. This implies that q_o is analytically normalizable [3], [11], [13]. $\square$

According to Lemma 3.1 above, except for the non analytically normalizable cases on the holonomy groups, there exists a neighborhood U of q_o where $\mathcal{F}$ can be written in an analytic normal form.

We shall describe the notion of singular holonomy introduced in [5] and used in [2], [19]:

3.1. The case of a formal holomorphic first integral. — Let us assume that q_o is a non degenerate resonant corner say, $\mathcal{F} : nxdy + mydx + \text{h.o.t.} = 0$ in a neighborhood of $q_o : x = y = 0$, where $n/m \in \mathbb{Q}_+$. We will assume that $\mathcal{F}$ has a formal first integral at q_o and therefore [13], a local holomorphic first integral on a neighborhood of q_o. Thus this singularity is linearizable [13] and we can define the *Dulac correspondence* in a neighborhood of the singularity q_o. This correspondence is defined as follows: By the hypothesis $q_o \in D_j$ is a linearizable singularity corresponding to a local holomorphic first integral of $\mathcal{F}$, therefore we can choose local coordinates $(x, y) \in U$, a neighborhood of q_o, centered at this point, such that $D_i \cap U = (x = 0)$, $D_j \cap U = (y = 0)$, and such that $\mathcal{F}\big|_U$ is given by $nxdy + mydx = 0$. We fix the local transverse sections as $\Sigma_j = (x = 1)$ and $\Sigma_i = (y = 1)$, such that $\Sigma_i \cap D_i = q_i \neq q_o$ and $\Sigma_j \cap D_j = q_j \neq q_o$. Let us denote by $G_j = \text{Hol}(\mathcal{F}, D_j, \Sigma_j)$ and by $G_i = \text{Hol}(\mathcal{F}, G_i, \Sigma_i)$. Also denote by $h_o \in \text{Hol}(\mathcal{F}, D_i, \Sigma_i)$ the element corresponding to the corner q_o. Then we have $h_o(x) = \exp(-2\pi \frac{n}{m}\sqrt{-1}) \cdot x$. The Dulac correspondence is therefore given by $\mathcal{D}$: $(\Sigma_i, q_i) \to (\Sigma_j, q_j)$, $\mathcal{D}(x_o) = x_o^{m/n}$. We use this correspondence in order to associate to G_i a subgroup $G_j * (\mathcal{D}_* G_i) \subset \text{Diff}(\Sigma_j, q_j)$. Given an element $h \in G_i$ we look for elements $h^{\mathcal{D}} \in \text{Diff}(\Sigma_j, q_j)$, which are solutions of the *adjunction equation* $h^{\mathcal{D}} \circ \mathcal{D} = \mathcal{D} \circ h$.

Case 1. — G_i is abelian: Take any element $h \in G_i$. Since G_i is abelian we have $h(x) = \mu x \widetilde{h}(x^m)$ for some $\widetilde{h} \in \mathcal{O}_1, \widetilde{h}(0) = 1$. We take $\mu_1 = \mu^{m/n}$ and $h_1 = \widetilde{h}^{m/n}$ be one of the n-roots of μ^m and $\widetilde{h}^m$ respectively. Then we define $h^{\mathcal{D}} : (\Sigma_j, q_j) \to (\Sigma_j, q_j)$, by

$$h^{\mathcal{D}}(y) = \mu_1 y h_1(y^n)$$

Clearly we have $h^{\mathcal{D}} \in G_j^v$.

Case 2. — $G_i \subset \mathbb{H}_{k_i}$ is non abelian, analytically normalizable and $nk_i/m = k_j \in \mathbb{N}$: In this case we have an analytic embedding $G_i \subset \mathbb{H}_{k_i}$. Take an element

$$h(x) = \frac{\lambda x}{(1 + ax^{k_i})^{1/k_i}} \in G_i.$$

We consider determinations of

$$y_o \longmapsto \frac{\lambda^{m/n} y_o}{\left(1 + a y_o^{nk_i/m}\right)^{m/nk_i}}$$

By definition the maps $h^{\mathcal{D}}$ are all these determinations. Clearly the maps $h^{\mathcal{D}}$ belong to the virtual holonomy group G_j^v.

Case 3. — G_i is solvable non abelian and not analytically normalizable: In this case it follows from **[4]** that the group of the commutators $[G_i, G_i]$ is cyclic, say, $[G_i, G_i] = \langle h_1 \rangle$ for some $h_1 \in G_i$ and G_i is generated by some power or root of h_o and some power h_1^ℓ, $\ell \in \mathbb{Z}$. Notice that if $n = m = 1$ then we have $\mathcal{D}(x) = x$, and given any $h \in G_i$, we may define $h^{\mathcal{D}} \in G_j^v$ as $h^{\mathcal{D}}(y) = h(y)$, in the coordinates above. Thus we may assume that $n \neq m$. We regard this case: First we consider the case the virtual holonomy group G_j^v is abelian. Then all its elements commute with the local holonomy g_o around q_o, associated to the separatrix contained in D_j. Therefore, using the same construction of Case 1 above we may consider the adjunction of G_j^v to the holonomy group G_i, as a subgroup of the virtual holonomy group G_i^v. If G_j^v contains some element g of infinite order then we have two possibilities to consider:

(a) g has non periodic linear part: In this case, g induces an element h in the adjunction holonomy and therefore in the virtual holonomy $h \in G_i^v$, which also has non periodic linear part. This implies that G_i^v (which is solvable by hypothesis), is analytically normalizable **[4]**. Therefore we may exclude this case.

(b) Every element g in G_j^v has periodic linear part: In this case we may find some non trivial element $g \in G_j^v$, which is tangent to the identity $g(y) = y + ay^{\ell+1} + \text{h.o.t.}$, $a \neq 0$. Then, g induces an element h in the virtual holonomy G_i^v, which has infinity order and some power tangent to the identity. Moreover, since G_j^v is analytically normalizable, it follows that g and h are analytically normalizable. This implies that the powers of h are analytically normalizable and therefore since the group of flat elements in G_i^v is cyclic, G_i^v is analytically normalizable. We exclude therefore this case, and conclude that all the elements in G_j^v have finite order. It follows that G_j^v is a group whose elements are rational rotations, and that each finitely generated subgroup is in fact a finite linearizable group. In particular G_j is a finite linearizable group.

Now we consider the case G_j^v is solvable non abelian, and analytically normalizable. In this case, once again we may use the same procedure of Case 2 above in order to induce non trivial analytically normalizable flat elements in the virtual holonomy G_i^v, and conclude that this is in fact analytically normalizable. Thus we exclude this case.

Summarizing, we conclude that *if G_i is exceptional* (that is, solvable non abelian and not analytically normalizable) *then G_j^v is either a group of rational rotations and therefore with finite finitely generated subgroups, or an exceptional group.* In this last case, we use **[4]** to conclude that G_j^v is generated by some root of the local holonomy

g_o associated to q_o (we may have $g_o = \mathrm{Id}$), and some flat element g_1. Moreover each flat element in G_j^v is some power of g_1. We may then use the ideas of **[14]** in order to conclude that some power of h_1 corresponds to some power of g_1 by means of the Dulac correspondence adjunction. In fact, we can give the ideas: It is possible to use the Dulac corresponce given by $\mathcal{D}(x) = x^{m/n}$, in order to consider the sets of "pseudo-orbits"

$$\{g_1^{k_r} \circ \mathcal{D} \circ h_1^{\ell_r} \circ \mathcal{D}^{-1} \circ \cdots \circ g_1^{k_2} \circ \mathcal{D} \circ h_1^{\ell_2} \circ \mathcal{D}^{-1} \circ g_1^{k_1} \circ \mathcal{D} \circ h_1^{\ell_1}(x)\} \subset \Sigma_i,$$

where $x \in \Sigma_i$, ℓ_t, k_t are integral numbers, and $\mathcal{D}^{-1}$ is the correspondence $\Sigma_j \to \Sigma_i$, $y \mapsto y^{n/m}$. These sets are contained in a same leaf of $\mathcal{F}$ for each fixed x, and as in **[14]**, if the powers g_1^ℓ, and h_1^k are never related by the congugation equation, then we will have accumulations for the leaves of $\mathcal{F}$, outside the origin in Σ_i. On the other hand, in the case we are interested in, we have discrete intersections of the leaves with the transverse sections, outside the origin, so that we will conclude that some power h_1^ℓ, passes to the virtual holonomy G_j^v, as some power g_1^k.

3.2. The case of a non degenerate non resonant corner. — Let us assume that q_o is a non degenerate non resonant corner say, $\mathcal{F} : xdy - \lambda ydx + \mathrm{h.o.t.} = 0$ in a neighborhood of $q_o : x = y = 0$, where $\lambda \in \mathbb{C} \backslash \mathbb{Q}$. We only need to consider the following case (see Definition 4.2): *G_i is abelian and G_j is solvable non abelian.*

In this case the singularity q_o is analytically linearizable. In fact, since the group G_j is solvable non abelian, and since the local holonomy φ associated to q_o has non periodic linear part it follows that the group G_j is analytically normalizable **[4]**, and therefore there exists an analytic coordinate $w \in \Sigma_j, w(q_j) = 0$, such that

$$\varphi(w) = \frac{aw}{(1 + bw^{k_j})^{1/k_j}},$$

where $a^n \neq 1, \forall\, n \in \mathbb{N}^*$. Thus, we can change coordinates analytically in order to have $\varphi(u) = au$ for some coordinate $u = \phi(w)$, $\phi \in \mathbb{H}_{k_j}$. This implies that the singularity q_o is analytically linearizable **[13]**.

The adjunction holonomy group $G_j * (\mathcal{D}_* G_i) \subset \mathrm{Diff}(\Sigma_j, q_j)$ is defined as follows: There exists an analytic embedding $G_j \subset \mathbb{H}_{k_j}$. We may also assume that we have $\Sigma_i, \Sigma_j, D_i \cap U, D_j \cap U, G_i, G_j, q_i, q_j$ given in terms of (x, y) as in 3.1 above.

Since G_i is abelian and contains a non resonant linearizable diffeomorphism, G_i is linearizable. In fact, we can assume that the holonomy group $\mathrm{Hol}(\mathcal{F}, D_i, \Sigma_i)$ is linear in the local coordinate $x\big|_{\Sigma_i}$. Therefore any element $h \in G_i$ corresponds to an element $h(x) = \mu_h \cdot x$ in $\mathrm{Diff}(\Sigma_i, q_i)$. In the present case the Dulac correspondence is given by $\mathcal{D}(x_o) = x_o^{-\lambda}$. We are looking for elements $h^{\mathcal{D}} \in \mathrm{Diff}(\Sigma_j, q_j)$, which satisfy $h^{\mathcal{D}} \circ \mathcal{D} = \mathcal{D} \circ h$. Therefore we choose $h^{\mathcal{D}}$ as

$$h^{\mathcal{D}}(y) = \mu_h^{-\lambda} \cdot y$$

where $\mu_h^{-\lambda}$ runs over the solutions of $z^{-1/\lambda} = \mu_h$. Clearly the diffeomorphims $h^{\mathcal{D}}$ belong to the virtual holonomy G_j^v.

3.3. The case of a saddle-node corner. — Assume now that q_o is a saddle-node singularity which will be assumed to be analytically normalizable. Thus, there exists a system of local coordinates $(x,y) \in U$ centered at q_o such that, $\mathcal{F}\big|_U$ is given by $x^{k+1}dy - y(1+\lambda x^k)dx = 0$. We assume that $D_i = (x=0)$ and $D_j = (y=0)$. We also introduce the transverse sections $\Sigma_i = (y = e^{1/k})$ and $\Sigma_j = (x=1)$. The leaves of $\mathcal{F}\big|_U$ are the level curves of the multiform first integral $f(x,y) = yx^{-\lambda}\exp\big(1/kx^k\big)$. Therefore, the *Dulac correspondence* is defined by

$$D\colon \Sigma_i \longrightarrow \Sigma_j, \quad x_o \longmapsto y_o = \mathcal{D}(x_o) = x_o^{-\lambda}\exp\big(1/kx_o^k\big).$$

Now we show how the Dulac correspondence can still be used to define an adjunction for the holonomy. This adjunction will be from the strong manifold to the central manifold, that is, from G_i to G_j above. Given an element $h \in G_i$ we look for elements $h^{\mathcal{D}} \in \mathrm{Diff}(\Sigma_j, q_j)$, which are solutions of the *adjunction equation* $h^{\mathcal{D}} \circ \mathcal{D} = \mathcal{D} \circ h$.

Case 1. — G_i is solvable non abelian. In this case we have $\lambda = 0$ and

$$h_o(x) = \frac{x}{(1+ax^k)^{1/k}}$$

and therefore $k_i = k$. We claim that G is analytically normalizable. In fact, since it is solvable it follows from [4] that it is analytically normalizable if, and only if, the (abelian) subgroup of flat elements in G is analytically normalizable. But this is clear because this group contains h_o. This also implies that we can assume that x is a normalizing coordinate for G_i. The Dulac correspondence $\mathcal{D}$ is given by $\mathcal{D}(x_o) = \exp\big(1/kx_o^k\big)$. Now we take any element $h \in G_i$ and write

$$h(x) = \frac{\mu x}{(1+ax^k)^{1/k}}.$$

Since h is of the form given above we have the adjunction equation as

$$h^{\mathcal{D}}(\exp\big(1/kx^k\big)) = \exp\big(a/k\mu^k\big)\cdot\big(\exp\big(1/kx^k\big)\big)^{\mu^{-k}}.$$

Thus we have linear solutions of the form $h^{\mathcal{D}}(y) = \exp\big(a/k\mu^k\big)\cdot y^{\mu^{-k}}$. However these are not uniform analytic functions. *Assume that h^k is tangent to* 1, that is, $\mu^k = 1$. In this case we can take solutions of the form $g(y) = \exp\big(a/k\big)\cdot y$. These are linear diffeomorphisms and we will denote any of them by $h^{\mathcal{D}}$ as before. Notice that *if q_o is a corner then by the normal hyperbolicity hypothesis G_i must be abelian.*

Case 2. — G_i is abelian. Let us consider once again the local holonomy $h_o \in G_i$ associated to the strong manifold of the saddle-node q_o. First notice that we have

$$h_o(x) = \exp\Big(\frac{x^{k+1}}{1+\lambda x^k}\frac{d}{dx}\Big).$$

Let us denote by $\xi(x) = \dfrac{x^{k+1}}{1+\lambda x^k}\dfrac{d}{dx}$. Once again we claim that G is analytically normalizable, and it follows from the fact that h_o is analytically normalizable and that G is abelian. It also is clear that h_o is a normal form diffeomorphism in the sense of [**11**], and therefore if $z \in (\Sigma_i, q_i)$ is a normalizing coordinate for G_i, then we must have $h_o(z) = \exp\left(\mu \dfrac{z^{k+1}}{1+\lambda z^k}\dfrac{d}{dz}\right)$, for some $\mu \in \mathbb{C}^*$. Thus it follows that

$$\xi(x) = \frac{x^{k+1}}{1+\lambda x^k}\frac{d}{dx} = \mu \frac{z^{k+1}}{1+\lambda z^k}\frac{d}{dz}.$$

On the other hand the local holonomy $g_o \in \mathrm{Diff}(\Sigma_j, q_j)$ associated to the local separatrix $(y=0)$; i.e, the holonomy of the *central manifold* of the saddle-node, is linear in $y\big|_{\Sigma_j}$ given by

$$g_o(y) = \exp\left(2\pi\lambda\sqrt{-1}y\frac{d}{dy}\right) = \exp(2\pi\lambda\sqrt{-1})y.$$

Thus it is natural to pass the elements $h \in G_i$, tangent to the identity, to G_j as linear maps in the coordinate $y\big|_{\Sigma_j}$. Let $h \in G_i$ with $h'(0) = 1$. We have $h = \exp(t_h\xi)$ for some $t_h \in \mathbb{C}$. Now we remark that if we consider the multiform function $f(x) = x^{-\lambda}\exp\left(1/kx^k\right)$ defined on the transversal $\Sigma_i = (y = e^{1/k})$, it extends to a local multiform first integral for $\mathcal{F}$. Therefore given any element $\gamma \in \pi_1(D_i \setminus \operatorname{sing}\mathcal{F})$, if we denote by $h_\gamma \in G_i$, the corresponding holonomy diffeomorphism associated to this homotopy class, then we obtain $f_o(x) \circ h_\gamma = \exp(2\pi c_\gamma\sqrt{-1}) \cdot f_o(x)$, for any fixed determination $f_o(x)$ of $f(x)$. Therefore we obtain

$$x^{-\lambda}\exp\left(1/kx^k\right) \circ \exp\left(t_{h_\gamma}\frac{x^{k+1}}{1+\lambda x^k}\frac{d}{dx}\right) = \exp(2\pi c_\gamma\sqrt{-1}) \cdot x^{-\lambda}\exp\left(1/kx^k\right).$$

Take now an element $h \in G_i$ tangent to the identity say

$$h(x) = \exp\left(t_h\frac{x^{k+1}}{1+\lambda x^k}\frac{d}{dx}\right).$$

Since we have $f_o(x) \circ h(x) = \exp(2\pi c_h\sqrt{-1}) \cdot f_o(x)$ for some $c_h \in \mathbb{C}$, we consider $h^{\mathcal{D}}(y) := \exp(2\pi c_h\sqrt{-1})y$. It is now clear that $h^{\mathcal{D}} \in \mathrm{Diff}(\mathbb{C}, 0)$ satisfies the adjunction equation. Finally we observe that by our construction we have $h^{\mathcal{D}} \in G_j^v$.

Remark 3.2. — As it is well-known a resonant nonlinearizable corner $q_o = D_i \cap D_j$, has a formal normal form,

$$\omega = [n(1 + (\lambda - 1)(x^m y^n)^k)xdy + m(1 + \lambda(x^m y^n)^k)y\,dx] = 0,$$

where $\lambda = n/m$ [**11**]. As we have seen in Proposition 2.8, when one of the virtual holonomy groups is nonabelian, then we have a formal normal form

$$\omega = kxdy + \ell y\Big(1 + \frac{\sqrt{-1}}{2\pi}x^\ell y^k\Big)dx = 0.$$

These formal normal forms admit *integrating factors* (that is, formal functions h such that $\Omega = \omega/h$ is closed), and give us closed meromorphic 1-forms Ω which have non simple poles along $(x \cdot y = 0)$. Moreover, these closed 1-forms are unique up to multiplication by constants (this is a consequence of the fact that the singularity admits no meromorphic first integral, except the constants). Later on, we will see that this last remark shows that it is not necessary for our purposes, to define any adjunction associated to such a corner (see Case C in the proof of Proposition 5.6).

Definition 3.3 **([5], [19])**. — Under the hypothesis of 3.1, 3.2 or 3.3 above for q_o, G_i^v and G_j^v, we define the group $G_j * (\mathcal{D}_* G_i)$ as the subgroup of $\mathrm{Diff}(\Sigma_i, q_i)$ generated by G_j and by the elements $h^{\mathcal{D}}$ where $h \in G_i$. This group will be called the *Dulac adjunction of G_i to G_j*[1].

Given a singularity $q_o = D_i \cap D_j$ as in cases 3.1, 3.2 or 3.3 above we assume that G_i is analytically normalizable, and choose analytic normalizing coordinates $(x, y) \in U$ for the singularity q_o (see Lemma 3.1 and Proposition 2.8), as in these cases, *i.e.*, coordinates that give the foliation $\mathcal{F}\big|_U$ in its local normal form. We choose $D_i \cap U = (x = 0)$ and $D_j \cap U = (y = 0)$. We may also choose x in such a way that the restriction $x\big|_{\Sigma_i}$ is an analytic normalizing coordinate for the holonomy group G_i (Lemma 3.1). However, it is not always true that $y\big|_{\Sigma_j}$ also normalizes G_j. This is the subject of the following lemma.

Lemma 3.4. — *The adjunction holonomy group $G_j * (\mathcal{D}_* G_i)$ is a solvable group if, and only if, $y\big|_{\Sigma_j}$ normalizes G_j.*

Lemma 3.4 is proved in **[2]**[2]. It also follows, under our hypothesis of analytic normalization and normal hyperbolicity on the saddle-nodes, from an equivalent result of **[16]**, which is stated in terms of the solvability and convergence of some formal Lie algebras of vector fields.

In other words, Lemma 3.4 says that if $q_o = D_i \cap D_j$ with G_i, G_j solvable as above, then we can normalize simultaneously G_i, G_j and the singularity q_o if, and only if, the adjunction holonomy group $G_j * (\mathcal{D}_* G_i)$ is a solvable group. This same lemma holds in the formal case, where G_i is not assumed to be analytically normalizable. We will use this lemma in order to "glue" certain integrating factors associated to adjacent components D_i and D_j (see Proposition 5.2 and Proposition 5.6).

In order to proceed to the definition of the singular holonomy groups of the components D_j we introduce an order in the resolution divisor (here we follow as in **[2]**). This order is defined as follows: Let $\mathcal{Z}$ be the (finite) family of connected components

---

(1) This adjunction process, defined preliminary for holonomy groups, must be iterated whenever it is possible.

(2) We will give a more general version of this lemma in §5 Lemma 5.7.

of

$$D^* = D \backslash \{\text{saddle-nodes}\} \cup \{\text{nonlinearizable resonant singularities}\}$$

Given two elements $A, B \in \mathcal{Z}$ we will say that $A > B$ if and only if there exist components D_i, D_j of D such that $D_i \cap D_j = \{q\}$ is either a nonlinearizable resonant singularity, or a saddle-node, $D_i^* = D_i - \{q\} \subset A$, $D_j^* = D_j - \{q\} \subset B$, and the strong manifold of q lies over D_i. Clearly $>$ defines a partial order on $\mathcal{Z}$. Each totally ordered subfamily $\mathcal{Z}_1 \subset \mathcal{Z}$ has a maximal element. Take a supremum say $A_o \in \mathcal{Z}$. It is clear that given a component D_j such that

$$D_j \backslash \{\text{saddle-nodes and resonant nonlinearizable singularities}\}$$

belongs to A_o then no adjunction can be defined from any other component B of $\mathcal{Z}$ to A_o. Thus we consider the *singular holonomy* of any fixed component D_j in A_o as the maximal subgroup $\mathrm{Hol}^{\mathrm{sing}}(\mathcal{F}, D_j, \mathcal{S}_j) = G_j^{\mathrm{sing}} \subset \mathrm{Diff}(\mathcal{S}_j, p_j)$ obtained by iterating the adjunction process at all the other corners of D_j, $D_j \cap D_k$, $k \neq j$ where D_k is in A_o. Now we consider the new family $\mathcal{Z}_1 = \mathcal{Z} - \{A_o\}$. Take a maximal element $A_1 \in \mathcal{Z}_1 = \mathcal{Z} - \{A_o\}$ with respect to the (induced partial order) $>$, and consider the same iteration of the adjunction process for the corners in A_1 adding to this the adjuncted singular holonomy of D_j. This defines the singular holonomy group of any component $D_k \in A_1$. Thus we may exhaust $\mathcal{Z}$ and define the adjunction of the holonomy for all the components of D.

Definition 3.5 **([5], [19])**. — The subgroup

$$\mathrm{Hol}^{\mathrm{sing}}(\mathcal{F}, D_j, \mathcal{S}_j) = G_j^{\mathrm{sing}} \subset \mathrm{Diff}(\mathcal{S}_j, p_j)$$

obtained by the algorithmic process above will be called the *singular holonomy group* associated to the component D_j of D. The singular holonomy group is defined up to conjugacies, depending on the choice of the transverse sections $\mathcal{S}_j$.

Using the fact that the singular holonomy is always contained in the virtual holonomy, we conclude from Corollary 2.4 that:

Corollary 3.6. — *Let $\mathcal{F}$ be a foliation on $\mathbb{C}P(2)$ having as limit set some singularities and an algebraic curve Λ as in Theorem 1.1. Then each projective singular holonomy group of* $\operatorname{sing} \mathcal{F} \cap \Lambda$ *is solvable with discrete pseudo-orbits outside the origin.*

4. Logarithmic derivatives of an integrable differential 1-form

In what follows $\mathcal{F}$ is a foliation on a complex surface M defined by a meromorphic integrable 1-form $\omega = 0$ outside its polar divisor $(\omega)_\infty$. We consider a bimeromorphic map $\pi \colon \widetilde{M} \to M$ (*e.g.* a resolution morphism as in §2) and denote by $\widetilde{\omega}$ the pull-back 1-form $\pi^*\omega$. The 1-form $\widetilde{\omega}$ is meromorphic in $\widetilde{M}$, but may have non-isolated singularities (in the case of the resolution morphism these singularities appear over the projective lines D_j introduced by π). On the other hand, for the cases we are interested in, we may assume that ($(\omega)_\infty$ and therefore) $(\widetilde{\omega})_\infty$, has non invariant codimension one irreducible components, and meets the divisor transversely at regular points of $\widetilde{\mathcal{F}}$. In this case, the singular set $\operatorname{sing}\widetilde{\mathcal{F}}$ is contained in the zero divisor $(\widetilde{\omega})_0$.

Definition 4.1 **([18])**. — Let $D \subset \widetilde{M}$ be an invariant divisor (not necessarily irreducible). A meromorphic 1-form $\widetilde{\eta}$ defined on $\widetilde{M}$ is called a *logarithmic derivative of $\widetilde{\omega}$ adapted to D* if

(1) $d\widetilde{\omega} = \widetilde{\eta} \wedge \widetilde{\omega}$, $d\widetilde{\eta} = 0$,

(2) the polar divisor $(\widetilde{\eta})_\infty$ of $\widetilde{\eta}$ has order one along (any component of) D, and consists of the union of $(\widetilde{\omega})_\infty \cup (\widetilde{\omega})_0$ and an invariant divisor of $\mathcal{F}$,

(3) the residue of $\widetilde{\eta}$ along any *noninvariant* irreducible component L of $(\widetilde{\omega})_\infty \cup (\widetilde{\omega})_0$ is equal to either $-$(the order of the poles of $\widetilde{\omega}$ along L), or (the order of $(\widetilde{\omega})_0$ along L), respectively. We remark that, *a priori*, $\widetilde{\eta}$ may have nonsimple poles on $\widetilde{M}$.

The existence of an adapted logarithmic derivative is related to some conditions in the resolution of $\operatorname{sing}\mathcal{F} \cap \Lambda$ **[2]**, **[19]** (see also Theorem 5.1 below).

Definition 4.2 **([2])**. — The foliation $\mathcal{F}$ has a *Liouvillian resolution* (*relative to* Λ) if the divisor $D = \bigcup_{j=0}^{m} D_j$ of the resolution of $\operatorname{sing}\mathcal{F} \cap \Lambda$ satisfies:

(i) The singular holonomy group of every D_i is solvable.

(ii) If $q = D_i \cap D_j$ is a non degenerate corner such that $G_i \subset \mathbb{H}_{k_i}$ and $G_j \subset \mathbb{H}_{k_j}$ are nonabelian, then q is a resonant singularity. Moreover, if q is linearizable (has a holomorphic first integral), then we can find local coordinates $(x, y) \in U$ centered at q such that $D_i \cap U = (x = 0)$, $D_j \cap U = (y = 0)$ and $\mathcal{F}\big|_U : k_j x dy + k_i y dx = 0$.

(iii) If $q = D_i \cap D_j$ is a saddle-node corner whose strong manifold is contained in D_i, then G_i is solvable, and G_j is abelian analytically linearizable.

The motivation for the definition above is given by **[2]**, **[18]** and the following result:

Proposition 4.3 **([2])**. — *Let $\mathcal{F}$ be a holomorphic foliation on $\mathbb{C}P(2)$ given by a rational 1-form ω, and suppose that there exists an algebraic $\mathcal{F}$-invariant curve Λ having only non dicritical singularities. Assume that each component of the resolution divisor of* $\operatorname{sing}\mathcal{F} \cap \Lambda$ *contains a linearizable non resonant singularity and that all the saddle-nodes in the resolution divisor are in good position relatively to this divisor. Then the following two assertions are equivalent:*

(i) The 1-form ω admits a rational logarithmic derivative η, adapted to the invariant curve $\Lambda \subset \mathbb{C}P(2)$.

(ii) $\mathcal{F}$ exhibits a Liouvillian resolution relative to Λ, and all the singular holonomy groups are analytically normalizable.

Theorem 4.4. — *Let $\mathcal{F}$ be a foliation on $\mathbb{C}P(2)$ having as limit set some singularities and an algebraic curve Λ. Assume that* $\operatorname{sing}\mathcal{F} \cap \Lambda$ *satisfies $(\mathcal{C}_1)$ and that the invariant part of the resolution divisor is connected. Then $\mathcal{F}$ has a Liouvillian resolution relative to Λ.*

Proof. — According to Corollary 3.6, it remains to prove that

(1) if $q = D_i \cap D_j \in \operatorname{sing} \widetilde{\mathcal{F}}$ is a non degenerate singularity such that the holonomy groups of D_i and of D_j are non abelian. Then q is resonant. Moreover, if q has a holomorphic first integral, then we can write it as $k_j x dy + k_i y dx = 0$ for some local coordinates with $D_j : \{y = 0\}$ and $D_i : \{x = 0\}$.

(2) If $q = D_i \cap D_j \in \operatorname{sing} \widetilde{\mathcal{F}}$ is a saddle-node whose strong manifold is tangent to D_i, then the holonomy group of D_j is abelian analytically linearizable.

Proof of (1). — In fact, if the holonomy group of D_j is not abelian then, since the virtual holonomy G_j^v exhibits discrete pseudo-orbits, it follows that G_j^v contains only elements with periodic linear part **[10]**. In particular the singularity q must be a resonant singularity. Assume now that q is linearizable say, as $mxdy + nydx = 0$ with $m, n \in \mathbb{N}$, $\langle m, n\rangle = 1$, and $(y = 0) \subset D_j$ and $(x = 0) \subset D_i$ as usual. We must prove that $m/n = k_j/k_i$. We consider first the case $m = n = 1$, and G_i is analytically normalizable. In this case as in item 3.1 of §3. we have that the Dulac correspondence is given by

$$\mathcal{D}\colon \Sigma_i \to \Sigma_j \quad \mathcal{D}(\{(x,1)\}) = \{(1,x)\} \in \Sigma_j.$$

We take any element $h \in G_i$, write

$$h(x) = \frac{\lambda x}{(1 + a x^{k_i})^{1/k_i}}$$

and define

$$G_j^v \ni h^{\mathcal{D}}(y) = y \mapsto \frac{\lambda y}{\left(1 + a y^{k_i}\right)^{1/k_i}}$$

using the fact that $m = n = 1$ (see 3.1 §3). Applying now Theorem 2.5 for G_i and G_j^v we conclude that $k_i \leq k_j$. The same way, we may conclude that $k_j \leq k_i$ and therefore $k_i = k_j$. Now we treat the case $m \neq n$ as in 3.1 §3. We repeat the main argument: Take elements $g_1 \in G_j^v$ and $h_1 \in G_i^v$ that are tangent to the identity of orders k_j and k_i respectively. Using the Dulac correspondence $\mathcal{D}(x) = x^{m/n}$, we introduce the sets of "pseudo-orbits"

$$\left\{ g_1^{k_r} \circ \mathcal{D} \circ h_1^{\ell_r} \circ \mathcal{D}^{-1} \circ \cdots \circ g_1^{k_2} \circ \mathcal{D} \circ h_1^{\ell_2} \circ \mathcal{D}^{-1} \circ g_1^{k_1} \circ \mathcal{D} \circ h_1^{\ell_1}(x) \right\} \subset \Sigma_i,$$

where $x \in \Sigma_i$, ℓ_t, k_t are integral numbers, and $\mathcal{D}^{-1}$ is the inverse correspondence $\Sigma_j \to \Sigma_i$, $y \mapsto y^{n/m}$. These sets are contained in a same leaf of $\mathcal{F}$ for each fixed x, and as in Theorem 2.5 and in Nakai's Theorem **[14]**, if the powers g_1^{ℓ}, and h_1^k are never related by the conjugation equation, then we will have accumulation points for the leaves of $\mathcal{F}$, outside the origin in Σ_i. Thus we conclude that some power h_1^{ℓ}, passes to the virtual holonomy G_j^v, as some power g_1^k. This implies that $m/n = k_j/k_i$.

Proof of (2). — In fact, choose analytic coordinates (x, y) at q, such that $\widetilde{\mathcal{F}}$ is given by $x^{p+1}dy - y(1 + \lambda x^p)dx = 0$. By hypothesis the saddle-node is normally hyperbolic so that the holonomy group G_j contains some element with non periodic linear part.

According to **[10]** since G has some discrete pseudo-orbit this implies that G_j cannot be solvable non abelian. Thus it follows that the holonomy group G_j is abelian and, since it contains some analytically linearizable non rational rotation, G_j is analytically linearizable. □

Remark 4.5. — Since $G_j^{\mathrm{sing}} \subset G_j^v$ also has discrete orbits, it follows (with the same proof) that in (2) above we have G_j^{sing} abelian analytically linearizable.

5. Formal logarithmic derivatives near the limit set

In this section we prove a generalization of Theorem 4.4 above:

Theorem 5.1. — *Let $\mathcal{F}$ be a foliation on M a complex projective surface, and let $\Lambda \subset M$ be an invariant irreducible analytic curve. Assume that* $\operatorname{sing}\mathcal{F} \cap \Lambda$ *satisfies property $(\mathcal{C}_1)$ and that all the saddle-nodes in the resolution divisor are in good position. If $\mathcal{F}$ is given by a meromorphic 1-form ω with isolated singularities, then the following assertions are equivalent:*

(i) The strict transform $\widetilde{\omega}$ of the 1-form ω, admits a transversely formal logarithmic derivative $\widetilde{\eta}$ over Γ, adapted to the invariant curve Γ.

(ii) $\mathcal{F}$ has a Liouvillian resolution relative to Λ.

We shall describe briefly the notion of transversely formal object. Suppose that $\Gamma \subset M$ is an algebraic codimension one divisor. We denote by $\mathcal{I}_\Gamma$ the sheaf of ideals defining $\Gamma \subset M$, and by $\mathcal{O}_M$ the sheaf of regular functions on M. For any $m \in \mathbb{N}$ we have $(\mathcal{I}_\Gamma)^{m+1} \subset (\mathcal{I}_\Gamma)^m \subset \mathcal{O}_M$.

The *infinitesimal tubular neighborhood of order m* of Γ in M is the locally ringed space

$$\Big(M, \frac{\mathcal{O}_M}{(\mathcal{I}_\Gamma)^m}\Big)$$

The *formal completion* of M *along* Γ is the locally ringed space $\widehat{\Gamma} = (M, \mathcal{O}_{\widehat{\Gamma}})$, whose structural sheaf is defined by the projective limit

$$\mathcal{O}_{\widehat{\Gamma}} = \varprojlim \frac{\mathcal{O}_M}{(\mathcal{I}_\Gamma)^m}$$

The *ring of (transversely) formal rational functions on M along* Γ, $K(\widehat{\Gamma})$, is defined as the ring of rational functions on the formal completion $\widehat{\Gamma}$ of M along Γ.

It is proved in **[7]** that:

(i) $K(\widehat{\Gamma}) = H^o(\widehat{\Gamma}, \mathcal{K}_{\widehat{\Gamma}})$ where $\mathcal{K}_{\widehat{\Gamma}}$ is the quotient field of the sheaf of total quotient rings of $\mathcal{O}_{\widehat{\Gamma}}$,

(ii) there exists a natural inclusion of sheaves $\mathcal{K}_\Gamma \to \mathcal{K}_{\widehat{\Gamma}}$, where $\mathcal{K}_\Gamma$ is the sheaf of germs of meromorphic functions defined on neighborhoods of points in Γ.

We refer to **[5]** for more details.

With the notions above we extend the notion of logarithmic derivative to the notion of *transversely formal logarithmic derivative* over a resolution divisor Γ.

The first step in the proof of Theorem 5.1 is the following (see also **[16]** for similar constructions):

Proposition 5.2. — *Let $\mathcal{F}$ be a foliation in the complex surface M, and let Λ be an invariant compact curve such that* $\operatorname{sing}\mathcal{F}\cap\Lambda$ *satisfies $(\mathcal{C}_1)$ and all the saddle-nodes in the resolution divisor are in good position. Assume that $\mathcal{F}$ has a Liouvillian resolution relative to Λ. Let $D_j \subset \Gamma$ be an irreducible component of the resolution divisor of* $\operatorname{sing}\mathcal{F}\cap\Lambda$. *Then there exists a closed transversely formal meromorphic* 1*-form $\widetilde{\eta}_j$ defined over D_j, such that $\widetilde{\eta}_j$ is a transversely formal logarithmic derivative of $\widetilde{\omega}$, adapted to D_j.*

Proof. — It follows from the definition of Liouvillian resolution that the groups G_j^{sing} are solvable. We distinguish the non abelian case and the abelian case considered in the following two lemmas. Clearly they complete the proof of Proposition 5.2. □

Lemma 5.3. — *Under the hypothesis of Proposition 5.2, assume that G_j^{sing} is non abelian and formally embedded in $\mathbb{H}_{k_j}$. Let $\ell_j = \operatorname{ord}((\widetilde{\omega})_0, D_j)$, then there exists a transversely formal closed logarithmic derivative $\widetilde{\eta}_j$ of $\widetilde{\omega}$, adapted to D_j, such that* $\operatorname{Res}_{D_j}\widetilde{\eta}_j = k_j + 1 + \ell_j$.

Proof. — In order to simplify the notation we write $k = k_j$. We follow **[16]**: Given any regular point $q \in D_j$ and a local transverse section Σ_q with $\Sigma_q \cap D_j = q$, we can choose a formal coordinate $y \in \Sigma_q$ centered at q, such that y defines the formal embedding $G_j \cong \operatorname{Hol}(\widetilde{\mathcal{F}}, D_j, \Sigma_q) \subset \mathbb{H}_{k_j}$. Given such a coordinate we consider the formal vector field $\widehat{\xi}_q = y^{k+1}d/dy$. We have $g * \widehat{\xi}_q = c_g \cdot \widehat{\xi}_q$, $\forall\, g \in \operatorname{Hol}(\widetilde{\mathcal{F}}, D_j, \Sigma_q)$. Thus $\widehat{\xi}_q$ can be extended into a global section $\widehat{\sigma}$ of the quotient sheaf $\widehat{\mathcal{S}}/\mathbb{C}^*$, where $\widehat{\mathcal{S}}$ is the sheaf of the transversely formal symmetries of $\widetilde{\mathcal{F}}$ over $D_j \setminus \operatorname{sing}\widetilde{\mathcal{F}}$ **[16]**. We define the 1-form $\widetilde{\eta}_j = d(\widetilde{\omega}(\widehat{\sigma}))/\widetilde{\omega}(\widehat{\sigma})$, which is a well defined transversely formal closed 1-form over $D_j \setminus \operatorname{sing}\widetilde{\mathcal{F}}$. The 1-form $\widetilde{\eta}_j$ is a logarithmic derivative of $\widetilde{\omega}$ adapted to $D_j \setminus \operatorname{sing}\widetilde{\mathcal{F}}$.

An alternative way of constructing $\widetilde{\eta}_j$ is the following (**[2]**, **[5]**, **[19]**): For each point $q \in D_j \setminus \operatorname{sing}\widetilde{\mathcal{F}}$ as above, we extend the coordinate y into a transversely formal coordinate defined over a certain neighborhood of q in $D_j \setminus \operatorname{sing}\widetilde{\mathcal{F}}$. This local extension is a consequence of the local trivialization for $\widetilde{\mathcal{F}}$. Thus we obtain a collection $\{(x_\alpha, y_\alpha), U_\alpha\}_{\alpha\in A}$ where the U_α are open sets which cover a neighborhood of $D_j \setminus \operatorname{sing}\widetilde{\mathcal{F}}$ in $\widetilde{\mathbb{C}P(2)}$, and also $(y_\alpha = 0) = U_\alpha \cap D_j$, $\widetilde{\omega}\big|_{U_\alpha} = g_\alpha dy_\alpha$ for some transversely formal function g_α over $U_\alpha \cap D_j$. The coordinates x_α are analytic in U_α and the coordinates y_α are transversely formal over $U_\alpha \cap D_j$, obtained as the extensions of the coordinates y above. Finally if $U_\alpha \cap U_\beta \neq \phi$ then we have

$$y_\alpha^k = \frac{a\overline{y}_\beta{}^k}{1 + b\overline{y}_\beta{}^k}$$

for some $a, b \in \mathbb{C}$. Define the 1-form

$$\widetilde{\eta}_j\big|_{U_\alpha \cap D_j} := (k+1)\frac{dy_\alpha}{y_\alpha} + \frac{dg_\alpha}{g_\alpha}.$$

It is not difficult to see that this is a well-defined closed transversely formal meromorphic 1-form over $D_j \backslash \operatorname{sing} \widetilde{\mathcal{F}}$, and satisfies the relation $d\widetilde{\omega} = \widetilde{\eta}_j \wedge \widetilde{\omega}$. Moreover $(\widetilde{\eta}_j)_\infty$ is invariant and contains $(y_\alpha = 0)$ as a simple pole of residue $k+1$.

The extension of the 1-form $\widetilde{\eta}_j$ to a singularity $q \in D_j$ is done as follows. According to Proposition 2.8 there exists a formal logarithmic derivative η_q defined at the singular point q. If q is a saddle-node then by the hypothesis we can take η_q meromorphic in a neighborhood of q. This same holds in the case q admits a formal first integral or is a nonresonant nondegenerate singularity (Proposition 2.8). In general, in the nondegenerate case the formal integrating factors η_q can be extended in a transversely formal way along the separatrices through q, as a consequence of the resommation properties of the integrating factors along the separatrices through q **[19]**(3). Thus, over a neighborhood of q in D_j, the difference $\widetilde{\eta}_j - \eta_q$ writes $\widetilde{\eta}_j - \eta_q = h_q \cdot \widetilde{\omega}$ for some transversely formal function h_q, defined over a punctured neighborhood of q in D_j, which is an integrating factor for $\widetilde{\omega}$, that is $d(h_q \cdot \widetilde{\omega}) = 0$. We will show that h_q extends in a transversely formal way to a neighborhood of q in D_j (see Remark 2.9 and the first part of the proof of Lemma 5.3 above). We consider the following cases:

(a) q is formally linearizable non resonant, of the form ω_λ in Proposition 2.8. In this case we know that there exist formal coordinates (x, y) such that $D_j = (y = 0)$, $D_i = (x = 0)$ and $\widetilde{\omega}(x, y) = g(xdy - \lambda ydx)$, for some transversely formal function g over a neighborhood of q in D_j. We also have

$$\eta_q = \frac{dg}{g} + a\frac{dx}{x} + b\frac{dy}{y}$$

where $a, b \in \mathbb{C}^*$ satisfy

$$(*) \qquad 1 + \lambda = a + b\lambda.$$

(3)We say that a 1-form $\widetilde{\omega}$ admits a *formal integrating factor* $\widehat{h}$ defined at q; if we have

$$(*) \qquad d\Big(\frac{\widetilde{\omega}}{\widehat{h}}\Big) = 0$$

where $\widehat{h}$ is a formal series at q. Equation $(*)$ exhibits *resommation properties for* $\widehat{h}$ along a certain separatrix S of $\widetilde{\omega} = 0$ at q, if $\widehat{h}$ can be written $\widehat{h}(x, y) = \sum_{j=0}^{+\infty} a_j(x)y^j$, where $(x, y) \in U$ is a local holomorphic coordinate centered at q, such that $S \cap U = \{y = 0\}$, $a_j(x)$ is a holomorphic function converging in a small disk $\mathbb{D}_q \subset S$ centered at q, not depending on $j \in \mathbb{N}$. This occurs, by a Briot-Bouquet type argument **[11]**, **[12]**, **[16]**, **[19]**, for any separatrix of a non-degenerate singularity, and for the strong separatrix of a saddle-node.

We have

$$\widetilde{\eta}_j - \eta_q = h_q.\widetilde{\omega}$$

as transversely formal objects over $(y = 0), (x \neq 0)$. The function $h_q(x,y)$ is a transversely formal function defined over $(y = 0), (x \neq 0)$, and since the 1-form $\frac{1}{gxy}\widetilde{\omega}(x,y)$ is closed it follows that we have $d(h_q xyg) \wedge \left(\frac{dy}{y} - \lambda \frac{dx}{x}\right) = 0$. Now, since $\lambda \in \mathbb{C}\backslash\mathbb{Q}$, it follows from **[18]** that $h_q xyg = \mu \in \mathbb{C}$ is constant (use Laurent series). Therefore, the formal expression $h_q(x,y)$ extends as a formal meromorphic extension to $(x = 0)$.

(b) q has a formal first integral. In this case, we can choose (x,y) as above such that $\widetilde{\omega}(x,y) = g(nxdy + mydx)$ as in Proposition 2.8 (b). We have

$$\eta_q = \frac{dg}{g} + a\frac{dx}{x} + b\frac{dy}{y}$$

where $a, b \in \mathbb{C}^*$ satisfy $n - m = na - mb$. The formal expression

$$d(h_q xyg) \wedge \left(\frac{dy}{y} + \frac{m}{n}\frac{dx}{x}\right) = 0$$

implies that $(h_q xyg)(x,y) = \varphi(x^m y^n)$ for some formal function $\varphi(z)$ in one variable. Therefore, the formal expression for $h_q(x,y)$ extends as $\varphi(x^m y^n)/xyg$ to $(x = 0)$.

(c) q has a non linear formal normalization as in Proposition 2.8 (c). In this case we may choose (x,y) such that $\widetilde{\omega} = g\big(kxdy + \ell y(1 + \frac{\sqrt{-1}}{2\pi}x^\ell y^k)dx\big)$. We choose

$$\eta_q = (k+1)\frac{dy}{y} + (\ell+1)\frac{dx}{x} + \frac{dg}{g}.$$

Then we have

$$d(h_q g x^{\ell+1} y^{k+1}) \wedge d\Big(\frac{1}{x^\ell y^k} - \frac{\sqrt{-1}\ell}{2\pi}\log x\Big) = 0.$$

On the other hand, since the singularity q has a non linear formal normalization, it follows that its local holonomy is not formally linearizable. This implies that $h_q g x^{\ell+1} y^{k+1} = \mu \in \mathbb{C}$ is a constant. Therefore $h_q(x,y)$ extends to $(x = 0)$.

(d) q is a saddle-node whose strong manifold is tangent to D_j. We may choose analytic coordinates (x,y) such that $(y = 0) \subset D_j$ and $\widetilde{\omega} = g(y^{k+1}dx - x(1+\lambda y^k)dy)$. We define

$$\eta_q = \frac{dx}{x} + (k+1)\frac{dy}{y} + \frac{dg}{g}.$$

Then the difference $\widetilde{\eta}_j - \eta_q$ has simple poles over $D_j\backslash\{q\} : (x \neq 0), (y = 0)$, therefore it follows that $\eta_q = \widetilde{\eta}_j$ in $(x \neq 0)$: In fact, since a saddle-node admits no formal first integral, it follows that up to multiplicative constants the 1-form

$$\frac{dx}{x} - \frac{dy}{y^{k+1}} - \lambda\frac{dy}{y}$$

is the only closed formal 1-form which defines $\mathcal{F}$ at q, and it has non simple poles over $(y = 0)$. Hence $\widetilde{\eta}_j$ extends as η_q to $(x = 0)$.

(e) q is a saddle-node whose strong manifold is transverse to D_j. In particular G_j is abelian analytically linearizable (see Proposition 4.3). This implies that $\operatorname{sing}\widetilde{\mathcal{F}}\cap D_j$ *contains no saddle-nodes in good position with respect to* D_j *and that the nondegenerate singularities are analytically linearizable.* Using the techniques of **[2]**, **[19]**, **[16]** one constructs a meromorphic closed 1-form $\widetilde{\Omega}_j$ in a neighborhood V_j of D_j minus the saddle-nodes in $\operatorname{sing}\widetilde{\mathcal{F}}\cap D_j$. This 1-form satisfies $\widetilde{\Omega}_j \wedge \widetilde{\omega} = 0$, and $\widetilde{\Omega}_j$ has simple poles, all of them contained in the divisor D. We have $\widetilde{\omega} = h_j \cdot \widetilde{\Omega}_j$ for some meromorphic h_j in V_j. Redefine now the 1-form $\widetilde{\eta}_j = dh_j/h_j$. Clearly $\widetilde{\eta}_j$ is a candidate for a logarithmic derivative of $\widetilde{\omega}$ adapted to D_j. Notice that $\widetilde{\eta}_j$ *has residue* 1 *over* D_j: In fact, the poles of $\widetilde{\Omega}_j$ are simple, and $\widetilde{\omega}$ has isolated singularities. We must show that $\widetilde{\eta}_j$ extends meromorphically to the saddle-node singularities in D_j. Fix such a singularity $q_o \in D_j$. We must have $q_o = D_l \cap D_j$ for some D_l (the saddle-nodes are in good position). We will show the extension of $\widetilde{\eta}_j$ to q_o using the fact that the adjunction holonomy group $G_j * \mathcal{D}_*(G_l)$ is solvable. Choose analytic coordinates $(x,y) \in U$ centered at q, such that $D_l \cap U = (x=0)$, and $D_j = (y=0)$, and such that $\widetilde{\omega}\big|_U(x,y) = g\cdot(x^{p+1}dy - ydx)$ for some meromorphic function g and some $p \in \mathbb{N}$. As we know the holonomy group G_l of the component D_l, must be solvable nonabelian $G_l \subset \mathbb{H}_p$. Moreover we can assume that the analytic coordinate $x\big|_{\Sigma_l}$, where $\Sigma_l = (y = e^{1/p})$, normalizes the group $\operatorname{Hol}(\widetilde{\mathcal{F}}, D_l, \Sigma_l) \cong G_l$. Take an element tangent to the identity $h \in G_l$, say

$$h(x) = \frac{x}{(1+ax^p)^{1/p}}.$$

Then the corresponding element $h^{\mathcal{D}} \in G_j * \mathcal{D}_*(G_l)$ is given by $h^{\mathcal{D}}(y) = e^{a/p}\cdot y$ Let now $z \in (\Sigma_j, q_j)$ be any analytic coordinate, where $\Sigma_j = (x=1)$ and $q_j = \Sigma_j \cap D_j$, which linearizes the holonomy group $G_j \cong \operatorname{Hol}(\widetilde{\mathcal{F}}, D_j, \Sigma_j)$. Then the fact that the singular holonomy group G_j^{sing} is solvable implies that either $y = \varphi(z)$ for some diffeomorphism $\varphi \in \mathbb{H}_k$ when $G_j^{\text{sing}} \subset \mathbb{H}_k$ by analytic conjugacy, or $y = \mu\cdot z$ for some $\mu \in \mathbb{C}^*$ when G_j^{sing} is abelian analytically linearizable. According to this we can assume that the analytic coordinate $y\big|_{\Sigma_j}$ linearizes G_j. This implies that the 1-forms $\widetilde{\eta}_j$ and

$$\eta_{q_o} := \frac{dy}{y} + (p+1)\frac{dx}{x} + \frac{dg}{g}$$

coincide over the transverse section Σ_j. This same argument shows that they coincide over an open neighborhood of q_j in D_j. Therefore $\widetilde{\eta}_j$ extends as η_{q_o} to the singularity q_o. □

Lemma 5.4. — *Under the hypothesis of Proposition 5.2, assume that* G_j^{sing} *is abelian. There exists a transversely formal logarithmic derivative* $\widetilde{\eta}_j$ *defined over* D_j, *this* 1*-form is obtained as follows: There exists a closed transversely formal* 1*-form* $\widetilde{\omega}_j$ *defined over* $D_j \backslash \operatorname{sing}\widetilde{\mathcal{F}}$, *which defines* $\widetilde{\mathcal{F}}$ *around* $D_j \backslash \operatorname{sing}\widetilde{\mathcal{F}}$ *in the sense that* $\widetilde{\omega}_j \wedge \widetilde{\omega} = 0$.

The 1*-form* $\widetilde{\eta}_j$ *is obtained by extending* $d\widetilde{h}_j/\widetilde{h}_j$*, where* $\widetilde{\omega} = \widetilde{h}_j \cdot \widetilde{\omega}_j$*, to the singularities* $q \in \operatorname{sing} \widetilde{\mathcal{F}} \cap D_j$.

Proof. — In this case the group G_j^{sing} may be non linearizable. Anyway, according to Lemma 2.6, there exists a formal singular holomorphic vector field in one complex variable $\xi \in \widehat{\mathcal{X}}(\mathbb{C}, 0)$ which is invariant by the natural action induced by G_j^{sing}, *i.e.*, $g * \xi = \xi$, $\forall\, g \in G_j^{\text{sing}}$. Now, using the techniques of **[16]** we can extend the vector $\widehat{\xi}$ into a *transversely formal symmetry* $\widehat{\sigma}$ for $\widetilde{\mathcal{F}}$, through the holonomy, (here one may use the invariance above), that is, a global section of the sheaf $\widehat{\mathcal{S}}$ of transversely formal symmetries of $\widetilde{\mathcal{F}}$ over $D_j \backslash \operatorname{sing} \widetilde{\mathcal{F}}$. This section induces a transversely formal integrating factor $\widehat{h}_j = \widetilde{\omega}(\widehat{\sigma})$ for $\widetilde{\omega}$ over $D_j \backslash \operatorname{sing} \widetilde{\mathcal{F}}$. We define the closed transversely formal 1-form $\widetilde{\eta}_j = d\widehat{h}_j/\widehat{h}_j$ over D_j. Since $\widetilde{\omega}/\widehat{h}_j$ is closed we have $d\widetilde{\omega} = \widetilde{\eta}_j \wedge \widetilde{\omega}$. Clearly $\widetilde{\eta}_j$ is a logarithmic derivative of $\widetilde{\omega}$ adapted to D_j and defined over $D_j \backslash \operatorname{sing} \widetilde{\mathcal{F}}$.

Now we show that this logarithmic derivative extends to the singularities in D_j. Fix a singularity $q \in \operatorname{sing} \widetilde{\mathcal{F}} \cap D_j$. As in the proof of Lemma 5.3, there exists a transversely formal logarithmic derivative η_q defined over a neighborhood of q in D_j, and we have $\widetilde{\eta}_j - \eta_q = h_q \cdot \widetilde{\omega}$ for some transversely formal function h_q which is an integrating factor for $\widetilde{\omega}$, defined over a punctured neighborhood of q in D_j. We consider the following cases:

(a) q is formally linearizable non resonant, of the form ω_λ in Proposition 2.8. In this case the same arguments of the proof of Lemma 5.3 above, apply to show that we have $hxyg =$ cte and therefore we have a natural extension of $\widetilde{\eta}_j$ to q.

(b) q has a formal first integral as $\omega_{k,\ell}$ in Proposition 2.8 (b). In this case, as above we can choose (x, y) above such that $\widetilde{\omega}(x, y) = g(nxdy + mydx)$, and

$$\eta_q = \frac{dg}{g} + a\frac{dx}{x} + b\frac{dy}{y}$$

where $a, b \in \mathbb{C}^*$ with $n - m = na - mb$. The formal expression $h_q(x, y)$ extends therefore as $\varphi(x^m y^n)/xyg$ to $(x = 0)$.

(c) q has a non linear formal normalization as in Proposition 2.8 (c). As above we conclude that $h_q(x, y)$ extends to $(x = 0)$.

(d) q is a saddle-node with the strong manifold tangent to D_j. As above the difference $\widetilde{\eta}_j - \eta_q$ has simple poles over $D_j \backslash \{q\} : (x \neq 0), (y = 0)$, and therefore it follows that $\eta_q = \widetilde{\eta}_j$ in $(x \neq 0)$. This implies that $\widetilde{\eta}_j$ extends as η_q to $(x = 0)$.

(e) q is a saddle-node whose strong manifold is transverse to D_j. In particular G_j^v is abelian analytically linearizable. This case is done as above, by using the fact that the adjunction holonomy groups associated to D_j are all abelian. This ends the proof of Lemma 5.4. □

Remark 5.5. — The following proposition is one of the main tools in this work, and we refer to **[19]**, **[2]** for a more general version and some details. This also may follow from the results in **[16]**, our hypothesis on the saddle-node singularities and the fact that the *virtual holonomy* groups are solvable (Corollary 2.4) and "contain" all the information concerning the foliation near the resolution divisor so that, in particular they must contain the "generalized holonomy" of **[16]**.

Proposition 5.6. — *Let $\mathcal{F}$ be a foliation on a complex surface M, and let Λ be an invariant compact curve such that the saddle-nodes in the resolution divisor of* $\operatorname{sing}\mathcal{F}\cap\Lambda$ *are in good position. Assume that $\mathcal{F}$ has a Liouvillian resolution relative to Λ, and that* $\operatorname{sing}\mathcal{F}\cap\Lambda$ *satisfies property C_1. There exists a closed transversely formal 1-form $\widetilde{\eta}$ defined over Γ, such that $d\widetilde{\omega}=\widetilde{\eta}\wedge\widetilde{\omega}$. This form $\widetilde{\eta}$ is a logarithmic derivative for $\widetilde{\omega}$ adapted to Γ.*

Proof. — We give the proof of the most relevant points in the construction. Let D_j be any component of Γ which we will write as $\Gamma=\bigcup_{i=0}^m D_i$. We consider a 1-form $\widetilde{\eta}_j$ over D_j given by Proposition 5.2 above. Fix a corner $q_o=D_i\cap D_j$ where D_i and D_j are components of Γ. We will prove that

$D_i\cap D_j\neq\varnothing\Rightarrow\widetilde{\eta}_i=\widetilde{\eta}_j$, *(for suitable and fixed choices of the 1-form $\widetilde{\eta}_j$), as formal expressions at q_o. In the case the 1-forms are meromorphic we also have the equality holding in a neighborhood of q_o.*

There are some cases to consider:

Case A. — q_o is a saddle-node singularity. In this case we have analytic coordinates (x,y) centered at q_o such that $\widetilde{\omega}=g\cdot\big(x(1+\lambda y^k)dy-y^{k+1}dx\big)$ and $D_i=(x=0)$, $D_j=(y=0)$. We have $\widetilde{\eta}_i-\widetilde{\eta}_j=f_{ij}\cdot\widetilde{\omega}$ for some formal meromorphic function f_{ij} such that $d(f_{ij}\cdot\widetilde{\omega})=0$ in a neighborhood of q_o. This implies that $f_{ij}gxy^{k+1}=a\in\mathbb{C}$. Therefore

$$\widetilde{\eta}_i-\widetilde{\eta}_j=a\cdot\Big(\lambda\frac{dy}{y}-\frac{dx}{x}+\frac{dy}{y^{k+1}}\Big),$$

and hence since the poles of $\widetilde{\eta}_i$ and $\widetilde{\eta}_j$ are simple we have $a=0$ and therefore $\widetilde{\eta}_i=\widetilde{\eta}_j$ in a neighborhood of q_o.

Now we assume that q_o is not a saddle-node. We have that q_o is a non degenerate singularity of the form $xdy-\lambda ydx+\text{h.o.t.}=0$ where (x,y) are local holomorphic coordinates centered at q_o and such that $D_i:(x=0)$ and $D_j:(y=0)$. As above $\widetilde{\eta}_i-\widetilde{\eta}_j=f_{ij}\cdot\widetilde{\omega}$, but now f_{ij} is a formal meromorphic function at q_o, which satisfies $d(f_{ij}\cdot\widetilde{\omega})=0$. We introduce the following types of singularity:

- ***type (1)*** **:** q_o admits a holomorphic first integral, G_i is abelian or nonabelian analytically normalizable, $G_i\subset\mathbb{H}_{k_i}$, with $nk_i/m\in\mathbb{N}$ as in Definition 4.2 and Theorem 4.4.
- ***type (2)*** **:** q_o is a non resonant analytically linearizable singularity and G_i is abelian.

We will use the following lemma whose proof we have extracted from **[19]** for the reader's convenience:

Lemma 5.7. — *Let $q_o = D_i \cap D_j$ be a corner of type* (1) *or* (2) *as above. Assume that the adjunction holonomy group $G_j * (\mathcal{D}_* G_i)$ is solvable. Then we may construct $\widetilde{\eta}_j$, $\widetilde{\eta}_i$ in such a way that we have $\widetilde{\eta}_j = \widetilde{\eta}_i$ as formal expressions at q_o. Moreover, for the cases G_i, G_j are analytically normalizable, let $(x,y) \in U$ be* analytic *coordinates such that $(x=0) = D_i \cap U$, $(y=0) = D_j \cap U$, $\widetilde{\mathcal{F}}\big|_U$ is in the normal form $xdy - \lambda ydx = 0$, $\lambda \in (\mathbb{C}\backslash\mathbb{Q}) \cup \mathbb{Q}_-$, and such that $x\big|_{\Sigma_i}$ is an analytic normalizing coordinate for* $\mathrm{Hol}(\widetilde{\mathcal{F}}, D_i, \Sigma_i)$, $\Sigma_i = (x=1)$. *Then the adjunction $G_j * (\mathcal{D}_* G_i)$ is solvable if, and only if, $y\big|_{\Sigma_j}$ $(\Sigma_j = (x=1))$ is an analytic normalizing coordinate for* $\mathrm{Hol}(\widetilde{\mathcal{F}}, D_j, \Sigma_j)$.

Proof. — We denote by $\ell_j = \mathrm{ord}((\widetilde{\omega})_o, D_j)$. First we assume that q_o is of type (2) We have G_i linearizable so that each $h \in \mathrm{Hol}(\widetilde{\mathcal{F}}, D_i, \Sigma_i)$ writes $h(x) = \mu_h \cdot x$ and induces an element $h^{\mathcal{D}}(y) = \nu_h \cdot y$ in $\mathrm{Diff}(\Sigma_j, q_j)$ where $\nu_h = \mu_h^{-\lambda}$. In particular the local holonomy h_0 around q_o induces $h_0^{\mathcal{D}}(y) = \left(e^{-2\pi i/\lambda}\right)^{-\lambda} y = y$. Thus we have two possibilities:

Case 1. — G_i is cyclic generated by h_0. In this case any element $h \in G_i$ writes $h = h_0^{m_h}$ for some $m_h \in \mathbb{Z}$. Now, if G_j is also abelian then we may assume that $\widetilde{\eta}_j = \widetilde{\eta}_i$ (are meromorphic and coincide) in a neighborhood of q_o. If G_j is nonabelian say, $G_j \subset \mathbb{H}_{k_j}$ by an analytic embedding, then we have necessarily $\mathrm{Res}_{D_j} \widetilde{\eta}_j = k_j + 1 + \ell_j$ (Lemma 5.3). This fixes $\mathrm{Res}_{D_i} \widetilde{\eta}_j$ as $\mathrm{Res}_{D_i} \widetilde{\eta}_j = 1 - k_j\lambda + \ell_i$ as it follows from equation $(*)$ in the proof of Lemma 5.3. On the other hand we may (since G_i is linearizable), choose $\widetilde{\eta}_i$ such that $\mathrm{Res}_{D_i} \widetilde{\eta}_i = 1 - k_j\lambda + \ell_i$ and therefore $\mathrm{Res}_{D_j} \widetilde{\eta}_i = k_j + \ell_i + 1$. Thus $\widetilde{\eta}_j - \widetilde{\eta}_i$ is holomorphic in a neighborhod of q_o and since

$$(\widetilde{\eta}_j - \widetilde{\eta}_i) \wedge \left(\frac{dy}{y} - \lambda \frac{dx}{x}\right) = 0$$

and $\lambda \notin \mathbb{Q}$, it follows that $\widetilde{\eta}_j = \widetilde{\eta}_i$ around q_o. Alternatively, we may use the fact that G_i is generated by h_0 in order to extend $\widetilde{\eta}_j\big|_{\Sigma_i}$ to a neighborhod of D_i and therefore obtain $\widetilde{\eta}_j = \widetilde{\eta}_i$ around q_o for $\widetilde{\eta}_i$ = extension of $\widetilde{\eta}_j\big|_{\Sigma_i}$ above obtained.

Case 2. — G_i is not cyclic. In this case there exists $h \in G_i$ such that $h^m \neq h_0^n$, $\forall\,(n,m) \ni \mathbb{Z} \times \mathbb{Z} - \{(0,0)\}$. Thus $\mu_h^m \neq e^{-2\pi i n/\lambda}$ and therefore $\mu_h^{-\lambda} \neq e^{2\pi i n/m}$ $\forall\,(n,m) \in \mathbb{Z} \times \mathbb{Z} - \{(0,0)\}$.

This shows that $h^{\mathcal{D}} \in G_j * (\mathcal{D}_* G_i) \subset \mathrm{Diff}(\Sigma_j, q_j)$ is of the form $h^{\mathcal{D}}(y) = \mu_h^{-\lambda} \cdot y$ and is not a rational rotation. Now, fixed a normalizing coordinate $z \in (\Sigma_j, q_j)$, $z(q_j) = 0$, for $G_j * (\mathcal{D}_* G_i)$ as a solvable group then (in the nonabelian case) we may write each element $g \in G_j * (\mathcal{D}_* G_i)$ as $g(z) = az/\sqrt[k_j]{1 + bz^{k_j}}$. In particular we may write $h^{\mathcal{D}}(z) = a \cdot z/\sqrt[k_j]{1 + bz^{k_j}}$ where $a = \mu_h^{-\lambda}$ is not a root of 1. This implies that $h^{\mathcal{D}}$ is linearizable by some coordinate $Z = T(z)$, where T is an homography. Therefore we may assume that $h^{\mathcal{D}}(z) = \mu_h^{-\lambda} \cdot z$. Since $\mu_h^{-\lambda}$ is not a root of 1 it follows that $z = \alpha \cdot y$ for some $\alpha \in \mathbb{C}^*$. Thus $y\big|_{\Sigma_j}$ also normalizes the group $G_j * (\mathcal{D}_* G_i) \supset G_j$. In particular we obtain $\widetilde{\eta}_j = \widetilde{\eta}_i$ around q_o.

– Now we consider the case q_o is of type (1) with G_i solvable nonabelian:

We write $\widetilde{\mathcal{F}}\big|_U : nxdy + mydx = 0$. Any element $h \in G_i$ writes

$$h(x) = \frac{ax}{\sqrt[k_i]{1 + bx^{k_i}}}$$

and induces an element $h^{\mathcal{D}} \in G_j * (\mathcal{D}_* G_i)$ of the form

$$h^{\mathcal{D}}(y) = \frac{a^{m/n} y}{\sqrt[k_j]{1 + by^{k_j}}} \in \mathrm{Diff}(\Sigma_j, q_j).$$

This implies that $G_j * (\mathcal{D}_* G_i)$ is nonabelian and we can choose a normalizing coordinate $z \in \Sigma_j$, $z(q_j) = 0$, such that $h^{\mathcal{D}}(z) = a^{m/n} z / \sqrt[k_j]{1 + cz^{k_j}}$. Now, if we choose $h \in G_i$ such that $b \cdot c \neq 0$ then necessarily we have $z^{k_j} = T(y^{k_j})$ for some homography $T(Z) \in \mathrm{SL}(2, \mathbb{C})$ and therefore $y\big|_{\Sigma_j}$ also normalizes the group $G_j * (\mathcal{D}_* G_i) \subset \mathrm{Diff}(\Sigma_j, q_j)$. In particular $\widetilde{\eta}_j = \widetilde{\eta}_i$ in a neighborhood of q_o.

– Finally, we consider the case where q_o is of type (1) with G_i abelian:

In this case again we write $\widetilde{\mathcal{F}}\big|_U : nxdy + mydx = 0$, and now any element $h \in G_i$, writes $h(x) = \mu x h_1(x^m)$ and induces $h^{\mathcal{D}}(y) = \mu^{m/n} y \cdot h_1^{m/n}(y^n)$ in the group $G_j * (\mathcal{D}_* G_i)$.

If G_i is not cyclic then there exists h such that $h^k \neq h_0^\ell$, $\forall\, (k, \ell) \in \mathbb{Z} \times \mathbb{Z} \backslash \{(0,0)\}$ and therefore we may proceed as above and conclude that $y\big|_{\Sigma_j}$ normalizes $G_j * (\mathcal{D}_* G_i)$ (notice that G_i not cyclic $\Rightarrow G_i$ analytically normalizable $\Rightarrow G_j$ is analytically normalizable). Thus we one reduced to the case G_i is cyclic. Thus G_i is in fact finite (because it is abelian and contains a rational rotation, h_0). We may consider two distinguished situations:

- If G_i is generated by h_0 then since $h_0^{\mathcal{D}} = \mathrm{Id}$ it follows that $G_j * (\mathcal{D}_* G_i) = G_j$ and we may extend $\widetilde{\eta}_j\big|_{\Sigma_i}$ to a neighborhood of D_i and therefore assume that $\widetilde{\eta}_j = \widetilde{\eta}_i$ at q_o.
- If G_i is not generated by h_0 then we have $G_i = \langle h_0^{1/\ell} \rangle$ for some $\ell \in \mathbb{N}$ and $h_0^{1/\ell} \in G_i$ induces an element $\big(h_0^{1/\ell}\big)^{\mathcal{D}} \in G_j * (\mathcal{D}_* G_i)$ that writes $\big(h_0^{1/\ell}\big)^{\mathcal{D}}(y) = e^{2\pi i/\ell} y$ which is not the identity. Since $G_j * (\mathcal{D}_* G_i) \supset G_j$ and is also solvable, we can construct $\widetilde{\eta}_j$ in such a way that $\big(h_0^{1/\ell}\big)^{\mathcal{D}}_* \widetilde{\eta}_j = \widetilde{\eta}_j$ in Σ_j, and this assures that $\big(h_0^{1/\ell}\big)_* \big(\widetilde{\eta}_j\big|_{\Sigma_i}\big) = \big(\widetilde{\eta}_j\big|_{\Sigma_i}\big)$ so that $\widetilde{\eta}_j\big|_{\Sigma_i}$ can be extended to a neighborhood of D_i and we may assume that $\widetilde{\eta}_j = \widetilde{\eta}_i$ at q_o. This ends the proof of the lemma. □

Case B. — $\lambda \notin \mathbb{Q}$. In this case, q_o is formally linearizable **[3]**, moreover, according to Definition 4.2 (ii) and Theorem 4.4, at least one of the holonomy groups of D_i and D_j is abelian, say G_i is abelian formally linearizable. However it is not clear that q_o is analytically linearizable, so that it is not immediate that we can introduce the adjunction holonomy group $G_j * \mathcal{D}_*(G_i)$. If both virtual holonomy groups are abelian then we can proceed as in the proof of Lemma 5.4 and construct the 1-forms $\widetilde{\eta}_i$ and $\widetilde{\eta}_j$ from integrating factors, say $\widetilde{\eta}_\nu = d \log h_\nu$, in a neighborhood of q_o, where $\widetilde{\omega}/h_\nu$

is closed, $\nu = i, j$. Clearly the quotient h_i/h_j is a first integral for $\widetilde{\omega}$ at q_o. Since the singularity is not formally linearizable it follows that $h_i = \text{cte} \cdot h_j$ and therefore we have $\widetilde{\eta}_i = \widetilde{\eta}_j$. Thus we may assume that G_j is non-abelian. But since it contains a diffeomorphism with non-periodic linear part (the local holonomy around q_o), it follows that G_j^v is analytically normalizable, and so it is the singularity q_o. This says that q_o is of type (2). The glueing of the forms $\widetilde{\eta}_i$ and $\widetilde{\eta}_j$ is therefore given by the fact that the adjunction holonomy group $G_j * \mathcal{D}_*(G_i)$ is solvable ([**2**], [**19**]) (see Lemma 5.7).

Case C. — q_o is not linearizable and $\lambda = -n/m \in \mathbb{Q}_-$ in the usual notation. If both singular holonomy groups are abelian, then according to the proof of Lemma 5.4 we construct the 1-forms $\widetilde{\eta}_i$ and $\widetilde{\eta}_j$ from integrating factors, say $\widetilde{\eta}_\nu = d\log h_\nu$, in a neighborhood of q_o, where $\widetilde{\omega}/h_\nu$ is closed, $\nu = i, j$. Clearly the quotient h_i/h_j is a first integral for $\widetilde{\omega}$ at q_o. Since the singularity is not formally linearizable it follows that $h_i = \text{cte} \cdot h_j$ and therefore we have $\widetilde{\eta}_i = \widetilde{\eta}_j$. Now we assume at least that only one of the holonomy groups say, G_j is nonabelian. In this case we have formal coordinates (x, y) at q_o such that

$$\widetilde{\omega} = g\Big(kxdy + \ell y\Big(1 + \frac{\sqrt{-1}}{2\pi}x^\ell y^k\Big)dx\Big) \qquad \text{(Proposition 2.8).}$$

Thus we have as usual $\widetilde{\eta}_i - \widetilde{\eta}_j = h_q \cdot \widetilde{\omega}$ where

$$d(h_q g x^{\ell+1} y^{k+1}) \wedge d\Big(\frac{1}{x^\ell y^k} - \frac{\sqrt{-1}\ell}{2\pi}\log x\Big) = 0.$$

On the other hand, since the singularity q_o has a non linear formal normalization, it follows that $h_q g x^{\ell+1} y^{k+1} = \mu \in \mathbb{C}$ is a constant. Finally we have $\mu = 0$ because both $\widetilde{\eta}_i$ and $\widetilde{\eta}_j$ have simple poles over Γ.

Case D. — q_o has a holomorphic first integral, $\lambda = -n/m$ as usual. If both virtual holonomy groups G_j^v and G_i^v are abelian, then we can consider the adjunction process from D_i to D_j and conversely. Thus it will follow that the 1-forms $\widetilde{\eta}_j$ and $\widetilde{\eta}_i$ may be constructed in compatible way, so that we have $\widetilde{\eta}_i = \widetilde{\eta}_j$ at q_o (indeed, q_o is of type (1)). Thus we may assume that some of the virtual holonomy groups say, G_j^v is non-abelian. If it is analytically normalizable then we have q_o of type (1) and we may apply Lemma 5.7. Let us however give a sole argument. Assume that G_i^v is also non-abelian. Then the 1-forms $\widetilde{\eta}_i$ and $\widetilde{\eta}_j$ have residues given by $\operatorname{Res}_{D_i}\widetilde{\eta}_i = \ell_i + k_i + 1$ and $\operatorname{Res}_{D_j}\widetilde{\eta}_j = \ell_j + k_j + 1$, in the usual notation. On the other hand we have that $m/n = k_j/k_i$ (Theorem 4.4). Using this and equation $(*)$ in the proof of Lemma 5.3, it follows that $\operatorname{Res}_{D_i}\widetilde{\eta}_j = \operatorname{Res}_{D_i}\widetilde{\eta}_i$ and also $\operatorname{Res}_{D_j}\widetilde{\eta}_j = \operatorname{Res}_{D_i}\widetilde{\eta}_j$. Thus, the difference $\widetilde{\eta}_j - \widetilde{\eta}_i$ is a closed (formal) holomorphic 1-form at q_o, so that we may write it as $\widetilde{\eta}_j - \widetilde{\eta}_i = d\varphi_{ij}(x^m y^n)$ for some (holomorphic) formal function $\varphi(z)$. Now, as in 3.1 of §3 *we can consider the adjunction holonomy groups* obtained from G_j and G_i. These groups are solvable and we have $G_i * \mathcal{D}^*(G_j) \subset G_i^v$ and $G_j * \mathcal{D}^*(G_i) \subset G_j^v$.

The solvability of these groups allows us to extend the 1-form $\widetilde{\eta}_i$ as a transversely formal 1-form over $D_j \setminus \operatorname{sing} \widetilde{\mathcal{F}}$, and therefore the integral $\varphi(x^m y^n)$ also extends as a transversely formal first integral over $D_j \setminus \operatorname{sing} \widetilde{\mathcal{F}}$. This implies, in the case $\varphi(z)$ is non-constant, that the virtual holonomy group G_j^v is finite **[13]**, **[16]**, and we would have a contradiction. Thus we may assume that G_j^v is non-abelian, but G_i^v is abelian. In particular *we may perform the adjunction from D_i to D_j and obtain a solvable subgroup $G_j * \mathcal{D}^*(G_i) \subset G_j^v$* (see 3.1 §3). We write the difference $\widetilde{\eta}_i - \widetilde{\eta}_j = \Omega_{ij}$, for some closed (formal) meromorphic 1-form at q_o. This 1-form can be extended by holonomy to a transversely formal closed meromorphic 1-form over $D_j \setminus \operatorname{sing} \widetilde{\mathcal{F}}$. This implies, in the $\Omega_{ij} \neq 0$, that the virtual holonomy G_j^v is abelian **[16]**, **[19]**, and we would have a contradiction. Thus $\Omega_{ij} = 0$ and therefore $\widetilde{\eta}_i = \widetilde{\eta}_j$ at q_o. Thus in any case either the glueing of $\widetilde{\eta}_i$ and $\widetilde{\eta}_j$ is immediate, or it is given by the fact that the adjunction holonomy is well-defined and solvable as it follows from Lemma 5.7.

Remark 5.8. — The solvability of the *virtual* holonomy groups is enough, under our hypothesis of normal hyperbolicity on the saddle-nodes, to conclude that we may choose all the 1-forms $\widetilde{\eta}_j$ in a simultaneously compatible way (see Remark 5.5). For instance we discuss the case the holonomy of D_j is finite, and there are D_i, D_k with non abelian holonomies such that $D_i \cap D_j \neq \varnothing \neq D_k \cap D_j$. In this case the 1-form $\widetilde{\eta}_j$ is not unique *over a neighborhood of* $D_j \setminus \operatorname{sing} \widetilde{\mathcal{F}}$. However, if $q_i = D_i \cap D_j$ and $q_k = D_k \cap D_j$ are not saddle-nodes then (they have finite local holonomies and therefore, by **[13]**, these singularities admit local holomorphic first integrals so that) we may perform the adjunction from the holonomy of D_j to D_k and D_i and conversely. Thus we consider the case where q_i and q_j are saddle-nodes. Since the holonomy of D_j is linearizable, it follows that these saddle-nodes are not in good position with respect to D_j, that is, their strong manifolds lie over D_i and D_k respectively. But in this case we can perform the adjunction of the holonomy of D_i to D_j and of the holonomy of D_k to D_j (see the construction of the singular holonomy which is below Lemma 3.4). Therefore, the fact that the singular holonomy G_j^{sing} is abelian linearizable (*cf.* Remark 4.5), is enough to assure the compatibility of the forms $\widetilde{\eta}_j$, $\widetilde{\eta}_i$ and $\widetilde{\eta}_k$ **[19]**. The collection of 1-forms $(\widetilde{\eta}_i)_{i=0}^m$ defines therefore a transversely formal 1-form $\widetilde{\eta}$ over Γ, which is a logarithmic derivative of $\widetilde{\omega}$, adapted to Γ.

Proof of Theorem 5.1. — According to Proposition 5.6, it remains to prove (i)$\Rightarrow$(ii) in Theorem 5.1. This is proved with a geometric interpretation of **[16]**, or following the steps of an analogous result stated in **[2]**, one may also find a proof in **[19]**, using our hypothesis of normal hyperbolicity on the saddle-nodes. □

6. Rationality of formal logarithmic derivatives

As in §2 we denote by Γ the invariant part of the divisor D, obtained in the resolution of $\operatorname{sing}\mathcal{F} \cap \Lambda$. The following proposition is a consequence of **[5]** which is based on a theorem of Hironaka-Matsumara **[8]**:

Proposition 6.1. — *Let $\widetilde{\alpha}$ be a transversely formal differential form defined over Γ in $\widetilde{\mathbb{C}P(2)}$, where $\Gamma \subset \widetilde{\mathbb{C}P(2)}$ is a normal crossing divisor, and $\widetilde{\mathbb{C}P(2)}$ is is obtained from $\mathbb{C}P(2)$ by a finite sequence of blowing-ups. Assume that $\Gamma \subset \widetilde{\mathbb{C}P(2)}$ satisfies property (psdc). Then $\widetilde{\alpha}$ extends rationally to $\widetilde{\mathbb{C}P(2)}$.*

Proposition 6.2. — *Let $\mathcal{F}$ be a foliation on $\mathbb{C}P(2)$, having as limit set some singularities and an algebraic curve Λ. Assume that* $\operatorname{sing}\mathcal{F} \cap \Lambda$ *satisfies property $(\mathcal{C}_1)$ and (psdc). Then there exists a closed rational 1-form $\widetilde{\eta}$ which is a rational logarithmic derivative for $\widetilde{\omega}$, adapted to the invariant part of the resolution divisor of* $\operatorname{sing}\mathcal{F} \cap \Lambda$. *In particular all the projective singular holonomy groups of* $\operatorname{sing}\mathcal{F} \cap \Lambda$ *are solvable analytically normalizable.*

Proof. — According to Theorem 4.4, $\mathcal{F}$ has a Liouvillian resolution relative to Λ. Using Propositions 5.6 and 6.1 we conclude that $\widetilde{\omega}$ admits a rational logarithmic derivative $\widetilde{\eta}$ adapted to Γ. We apply Proposition 6.1 to conclude the rationality of $\widetilde{\eta}$. The last part of the statement is a consequence of **[16]** or, also, of an improvement of **[2]** found in **[19]**. □

7. Proof of Theorem 1.1

According to Proposition 6.2 we know that $\widetilde{\mathcal{F}}$ (that is $\widetilde{\omega}$) admits a rational logarithmic derivative $\widetilde{\eta}$ on $\widetilde{\mathbb{C}P(2)}$, which is adapted to Γ. According to Proposition 6.2 above, all the singular holonomy groups appearing in the resolution of $\operatorname{sing}\mathcal{F} \cap \Lambda$ are analytically normalizable. Thus we may apply the last part of **[2]** which assures that either $\mathcal{F}$ is given by a closed rational 1-form or $\mathcal{F}$ is a rational pull-back of a Riccati foliation as in Theorem 1.1. □

We can relax the hypothesis on the limit set of $\mathcal{F}$ if we assume that $\mathcal{F}$ has a transcendent leaf whose limit set is an algebraic curve Λ plus some singularities but, in order to prove that $\Gamma = \lim(\widetilde{L})$ for the corresponding transcendent leaf $\widetilde{L}$ of $\widetilde{\mathcal{F}}$, we have to make an additional hypothesis on $\operatorname{sing}\mathcal{F} \cap \Lambda$.

$(\mathcal{C}_2)$ *Λ contains all its local separatrices, and all the saddle-nodes appearing in the resolution of* $\operatorname{sing}\mathcal{F} \cap \Lambda$ *have their strong manifolds contained in the limit set* $\lim \widetilde{L}$.

Using the same techniques as in the proof of Theorem 1.1 we can prove:

Theorem 7.1. — *Let $\mathcal{F}$ be a foliation on $\mathbb{C}P(2)$, having a transcendent leaf whose limit set is an algebraic curve Λ. Assume that* $\operatorname{sing}\mathcal{F} \cap \Lambda$ *satisfies property (psdc),* *conditions* $(C_1), (C_2)$. *Then, either* $\mathcal{F}$ *is given by a closed rational* 1*-form or it is a rational pull-back of a Riccati foliation* $\mathcal{R} : p(x)dy - (a(x)y^2 + b(x)y)dx = 0$, *where* Λ *corresponds to* $\overline{(y=0)} \cup \overline{(p(x)=0)}$, *on* $\overline{\mathbb{C}} \times \overline{\mathbb{C}}$.

A few words should be said, concerning the relations between our result and groups of linear rational transformations. Let G be a finitely generate *Fuchsian group*, i.e, a properly discontinuous group of diffeomorphisms of $\overline{\mathbb{C}}$, carrying a certain circle $C(G)$, (the *principal circle*), into itself. It is known that the limit set of G lies on $C(G)$. Moreover it is well-known that *if* $\lim(G)$ *has more than two points, either* $\lim(G) = C(G)$, *or* $\lim(G) \subset C(G)$ *is a nowhere dense perfect subset* **[6]**.

Let us assume that $\lim(G) = C(G)$. Using **[9]** we can realize G as the "suspension holonomy" of the line $\overline{(y=0)} \subset \overline{\mathbb{C}} \times \overline{\mathbb{C}}$, of a Riccati foliation

$$\mathcal{R}(G) : p(x)dy - (a(x)y^2 + b(x)y + c(x))dx = 0$$

on $\overline{\mathbb{C}} \times \overline{\mathbb{C}}$. The foliation $\mathcal{R}(G)$ satisfies $\lim(\mathcal{R}(G)) = M^3(G)$, for a real 3-dimensional singular subvariety $M^3(G)$, which is singular along the invariant vertical fibers $\overline{\mathbb{C}}_x$: $x \times \overline{\mathbb{C}}, p(x) = 0$, and such that the intersection $M^3(G) \cap \overline{\mathbb{C}}_x$ is a principal circle of a Fuchsian group conjugate to G, provided that the fiber $\overline{\mathbb{C}}_x$ is non invariant. Conversely, one can ask whether a foliation whose limit set is an invariant real singular hypersurface as $M^3(G)$ above, is in fact the pull-back of a Riccati foliation. This problem remains open.

References

[1] C. Camacho, A. Lins Neto, P. Sad: *Foliations with algebraic limit sets*, Ann. of Math. **136** (1992), 429–446.

[2] C. Camacho, B. Scárdua: *Liouvillian first integrals of holomorphic foliations*, to appear in Ergodic Theory and Dynamical Systems.

[3] D. Cerveau et J.-F. Mattei: *Formes intégrables holomorphes singulières*, Astérisque, vol. 97 (1982).

[4] D. Cerveau, R. Moussu: *Groupes d'automorphismes de* $(\mathbb{C}, 0)$ *et équations différentielles* $udy + \cdots = 0$, Bull. Soc. math. France **116** (1988), 495–488.

[5] D. Cerveau, B. Azevedo Scárdua: *On the integration of polynomial vector fields in dimension two*, Pre-print, IRMAR, April 1996.

[6] L.R. Ford: *Automorphic Functions*, Chelsea Pub. Company, N.Y. 1951.

[7] R. Hartshorne: *Ample subvarieties of algebraic varieties*, Lect. Notes in Math., vol. 156, Springer, 1970.

[8] H. Hironaka, H. Matsumara: *Formal functions and formal embeddings*, J. Math. Soc. Japan, Vol. 20, Nos.1-2, 1968.

[9] A. Lins Neto: *Construction of singular holomorphic vector fields and foliations in dimension two*, J. of Diff. Geometry **26** (1987),1–31.

[10] F. Loray: *Feuilletages holomorphes à holonomie résoluble*, Thèse Université de Rennes I, 1994.

[11] J. Martinet et J.-P. Ramis: *Classification analytique des équations différentielles non linéaires résonantes du premier ordre*, Ann. Sc. Éc. Norm. Sup. 4^e série **16** (1983), 571–621.

[12] J. Martinet et J.-P. Ramis: *Problème de modules pour des équations différentielles non linéaires du premier ordre*, Publ. Math. Inst. Hautes Études Sci. **55** (1982), 63–124.

[13] J.F. Mattei, R. Moussu: *Holonomie et intégrales premières*, Ann. Sc. Éc. Norm. Sup. 4^e série **10** (1980), 469–523.

[14] I. Nakai: *Separatrices for non solvable dynamics on* $\mathbb{C}, 0$, Ann. Inst. Fourier (Grenoble) **44** (1994), 569–599.

[15] E. Paul: *Groupes de difféomorphismes résolubles*, Prépublication Univ. Paul Sabatier, Toulouse (1995).

[16] E. Paul: *Formes singuliéres à holonomie résoluble*, Prépublication Univ. Paul Sabatier, Toulouse (1995).

[17] P. Sad: *Dicritical singularities in foliations with algebraic limit sets*, Bol. Soc. Bras. Mat. **26** (1995) 47–56.

[18] B. Azevedo Scárdua: *Transversely affine and transversely projective foliations*, Ann. Sc. Éc. Norm. Sup. 4^e série **30** (1997), 169–204.

[19] B. Azevedo Scárdua: *Integration of complex differential equations*, Journal of Dynamical and Control Systems **5** (1999), 1–50.

[20] A. Seidenberg: *Reduction of singularities of the differential equation $Ady = Bdx$*, Amer. J. Math. **90** (1968), 248–269.

[21] Y. Siu: *Techniques of extension of analytic objects*, Marcel Dekker, New York, 1974.

C. Camacho, IMPA, Est. D. Castorina 110, 22460-320 Rio de Janeiro, Brazil
E-mail : `camacho@impa.br` • *Url :* `http://www.impa.br/~camacho`

B. Azevedo Scárdua, IMPA, Est. D. Castorina 110, 22460-320 Rio de Janeiro, Brazil
E-mail : `scardua@impa.br` • *Url :* `http://www.impa.br/~scardua`

ALBERT FATHI

Une caractérisation des stades à virages circulaires

Astérisque, tome 261 (2000), p. 89-101
<http://www.numdam.org/item?id=AST_2000__261__89_0>

Astérisque
261, 2000, p. 89–101

UNE CARACTÉRISATION DES STADES À VIRAGES CIRCULAIRES

par

Albert Fathi

Résumé. — Nous donnons une minoration du volume d'un domaine compact convexe d'un espace euclidien dont le bord est de classe $C^{1,1}$. Nous caractérisons le cas d'égalité.

0. Introduction

On munit $\mathbb{R}^n$ du produit scalaire usuel. On notera $\|.\|$ la norme euclidienne usuelle sur $\mathbb{R}^n$. La sphère unité dans $\mathbb{R}^n$ est notée $\mathbb{S}^{n-1}$.

Afin de simplifier, dans cette introduction, nous allons considérer un convexe $K \subset \mathbb{R}^n$ d'intérieur non vide et de bord C^2. Pour $x \in \partial K$, notons $N(x)$ le vecteur unitaire normal à ∂K en x et pointant à l'extérieur de K. L'application $N : \partial K \to \mathbb{S}^{n-1}$ est C^1. Puisque les espaces tangents $T_x\partial K$ et $T_{N(x)}\mathbb{S}^{n-1}$ sont tous les deux égaux au sous-espace vectoriel $N(x)^\perp$ orthogonal à $N(x)$, la dérivée $DN(x)$ est donc une application linéaire de $N(x)^\perp$ dans lui-même, on sait qu'elle est symétrique et définie positive. Les valeurs propres de $DN(x)$, notées $\kappa_1(x), \ldots, \kappa_{n-1}(x)$, sont donc ≥ 0, on les appelle les courbures principales. On a $\|DN(x)\| = \max(\kappa_1(x), \ldots, \kappa_{n-1}(x))$, puisque $DN(x)$ est symétrique. Les rayons de courbure principaux $R_1(x), \ldots, R_{n-1}(x)$ sont les inverses $\kappa_1(x)^{-1}, \ldots, \kappa_{n-1}(x)^{-1}$. Soit $H_i(x)$ la $i^{\text{ème}}$ fonction symétrique de $\kappa_1(x), \ldots, \kappa_{n-1}(x)$. On note par σ la mesure d'aire induite par la métrique euclidienne sur ∂K. On définit $\mathcal{H}_i(K) = \int_{\partial K} H_i(x)\, d\sigma(x)$. Le théorème de Gauss-Bonnet donne $\mathcal{H}_{n-1}(K) = s_{n-1}$ l'aire de la sphère unité $\mathbb{S}^{n-1}$ dans $\mathbb{R}^n$.

On pose :

$$\begin{aligned} \kappa_{\text{sup}} &= \sup\{\kappa_i(x) \mid x \in \partial K,\ i = 1, \ldots, n-1\} \\ R_{\text{inf}} &= \inf\{R_i(x) \mid x \in \partial K,\ i = 1, \ldots, n-1\}. \end{aligned}$$

***Classification mathématique par sujets* (1991).** — 52A10, 52A20, 52A40, 53C99.
Mots clefs. — Convexe, rayon de courbure, Blaschke, volume.

En particulier, on a $\kappa_{\sup} = \sup\{\|DN(x)\| \mid x \in \partial K\}$.

Blaschke a montré qu'une boule de rayon $R_{\inf}$ roulait librement à l'intérieur de K, (*cf.* [**Bl**, pages 114–119] et [**Le**, Theorem 2.1, page 1055]). Si on définit l'application $\mathcal{N} : \partial K \times [0, \infty[\to \mathbb{R}^n$ par $\mathcal{N}(x,t) = x - tN(x)$, le théorème de Blaschke montre que $\mathcal{N}$ est un plongement de $\partial K \times [0, R_{\inf}[$ dans K. Il en résulte le théorème suivant :

Théorème 0.1. — *Si $K \subset \mathbb{R}^n$ est un compact convexe d'intérieur non vide et dont le bord est de classe C^2, alors, on a :*

$$V(K) \geq R_{\inf} S(K) + \sum_{i=1}^{n-1} (-1)^i R_{\inf}^{i+1} \frac{\mathcal{H}_i}{i+1},$$

où $V(K)$ est le volume de K et $S(K)$ est l'aire de son bord ∂K.

De plus, cette inégalité est une égalité si et seulement si K est une boule euclidienne.

Plus généralement, si K est un compact convexe d'intérieur non vide et dont le bord est C^1, on dit que K est $\mathrm{C}^{1,1}$ si le plan tangent $T_x\partial K$ est lipschitzien comme fonction de $x \in \partial K$. Il revient au même de dire que l'application normale $N : \partial K \to \mathbb{S}^{n-1}$ est lipshitzienne. Dans ce cas, les courbures sont définies presque partout. On peut voir que l'on peut faire rouler librement une boule de rayon $R > 0$ à l'intérieur d'un compact convexe K si et seulement si ∂K est $\mathrm{C}^{1,1}$ (*cf.* [**Ho**, Proposition 2.4.3, p. 97]). Il n'est donc pas étonnant que le théorème précédent s'étende au cas $\mathrm{C}^{1,1}$. Le cas d'égalité ayant lieu précisément quand $K = \{x \in \mathbb{R}^n \mid d(x, C) \leq R\}$ avec C compact convexe de dimension $\leq n-1$ et $R > 0$.

Dans le cas où $C = S$ est un segment de $\mathbb{R}^2$, un ensemble de la forme $\{x \in \mathbb{R}^2 \mid d(x, S) \leq R\}$ est ce que l'on appelle, pour des raisons visuelles évidentes, un *stade à virages circulaires*. Remarquons que le bord d'un tel stade est soit un cercle de rayon R si S est réduit à un point, soit la réunion de deux segments, de mêmes longueurs et parallèles à S et de deux demi-cercles de rayon R, par conséquent si S n'est pas réduit à un point, un tel stade n'est jamais C^2.

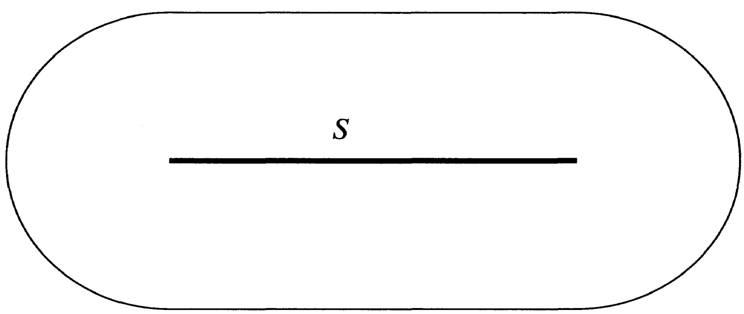


FIGURE 1. Un stade à virages circulaires.

Le théorème suivant donne une caractérisation des stades circulaires :

Théorème 0.2. — *Soit C une courbe convexe fermée plane de classe $C^{1,1}$. Notons par l sa longueur, par A l'aire entourée et par R_{inf} la borne inférieure des rayons de courbure. On a :*

$$A \geq lR_{\mathrm{inf}} - \pi R_{\mathrm{inf}}^2.$$

De plus, cette inégalité est une égalité si et seulement si C borde un stade à virages circulaires.

Remarque 0.3. — 1) Quand ce manuscrit était pratiquement terminé, Victor Bangert nous a communiqué les travaux d'Innami [**In**$_1$, **In**$_2$]. Certains des résultats d'Innami sont très proches des nôtres. En fait, bien que l'inégalité qu'Innami établit dans [**In**$_1$, Corollary 2], pour les hypersurfaces convexes lisses, soit moins bonne que la nôtre, les méthodes utilisées dans son travail sont les mêmes que les nôtres (en particulier, [**In**$_1$, Lemma 5] n'est rien d'autre que le théorème de Blaschke) et donc un examen de sa démonstration donne l'inégalité du théorème 0.1 ci-dessus. Bien sûr, sa minoration du volume étant moins bonne, le cas d'égalité est plus facile à analyser. Par ailleurs ses méthodes ne s'appliquent pas au cas des convexes de bord de classe $\mathrm{C}^{1,1}$.

2) Dans [**In**$_1$], Innami établit aussi une minoration de $V(K)$ pour le cas où K est de bord lisse mais n'est pas nécessairement convexe. Il faut, alors, remplacer R_{inf} par le rayon $r(K)$ de la plus grande boule euclidienne qui roule librement à l'intérieur de K. Un examen de sa démonstration, ou de la démonstration du théorème 1.1 ci-dessous, montre que l'on peut établir la même inégalité que dans le théorème 1.1, à condition de remplacer R_{inf} par $r(K)$, dès que ∂K est $\mathrm{C}^{1,1}$ (on peut voir que cette dernière condition implique $r(K) > 0$). Il serait intéressant d'analyser le cas d'égalité.

3) Dans [**Ga**], on établit l'inégalité $\pi l/A \leq \int_0^l \kappa(s)^2\, ds$, pour une courbe convexe C de classe C^2 dans le plan. Cette inégalité est optimale quand C s'approche d'un cercle. Si C s'approche d'un stade à virages circulaires quelconque, on voit que $\int_0^l \kappa(s)^2\, ds$ tend vers $2\pi/R_{\mathrm{inf}}$. L'inégalité de Gage devient donc dans ce cas limite $A \geq lR_{\mathrm{inf}}/2$, ce qui est moins bon que le théorème 1.2, dans le cas d'un stade qui n'est pas un disque.

4) Il est bon de signaler ici que si C est un convexe dans $\mathbb{R}^n$, les ensembles de la forme $C_R = \{x \in \mathbb{R}^n \mid d(x, C) = R\}$ sont tous, pour $R > 0$, des convexes à bord $\mathrm{C}^{1,1}$. Ceci résulte du fait, bien connu, que $x \mapsto d(x, C)$ est $\mathrm{C}^{1,1}$ sur $\mathbb{R}^n \setminus C$. Pour une démonstration, on peut consulter par exemple [**Ho**, theorem 2.1.30, p. 62], où il est montré que cette fonction est différentiable ; de plus, son gradient sur $\mathbb{R}^n \setminus C$ est donné par $x \mapsto (x - p_C(x))/\|x - p_C(x)\|$, où $p_C : \mathbb{R}^n \to C$ est la projection sur le point le plus proche de C. Or p_C est lipschitzienne. (On peut voir qu'une fonction distance à un fermé F dans une variété riemannienne M qui est différentiable en tous les points de $M \setminus F$ est, en fait, $\mathrm{C}^{1,1}$ sur $M \setminus F$).

1. Un théorème de Blaschke

Soit $K \subset \mathbb{R}^n$ un compact convexe. Les observations suivantes sont élémentaires et bien connues (*cf.* par exemple le paragraphe 3.1 sur la soustraction de Minkowski dans [**Sc**]). Si $r \geq 0$, on pose :

$$K_r^- = \{x \in K \mid d(x, \partial K) \geq r\}.$$

L'ensemble K_r^- est un compact convexe. La convexité résulte du fait que $x \in K_r^-$ si et seulement si $x + B(0,r) \subset K$. On peut voir que la frontière de K_r^- dans $\mathbb{R}^n$ est $\partial K_r^- = \{x \in K \mid d(x,K) = r\}$. On a aussi $\mathring{K}_r^- = \cup_{s>r} K_s^-$, où $\mathring{K}_r^-$ désigne l'intérieur du compact K_r^-. En particulier, si on pose $\rho_{\max} = \sup\{r \geq 0 \mid \exists x \in K,\ d(x, \partial K) = r\}$, on a $K_r^- = \varnothing$ pour $r > \rho_{\max}$, l'ensemble $K_{\rho_{\max}}^-$ est un convexe compact non vide de dimension $\leq n-1$ et, pour $r < \rho_{\max}$, la frontière ∂K_r^- est homéomorphe à $\mathbb{S}^{n-1}$ et $\mathring{K}_r^- \neq \varnothing$. (Le nombre $\rho_{\max}$ est le rayon de la plus grande boule euclidienne contenue dans K.)

Soit $K \subset \mathbb{R}^n$ un compact convexe d'intérieur non vide et dont le bord est de classe $\mathrm{C}^{1,1}$. Pour $x \in \partial K$, notons $N(x)$ le vecteur unitaire normal à ∂K en x orienté vers l'extérieur de K. L'application $N : \partial K \to \mathbb{S}^{n-1}$ est lipschitzienne. Posons :

$$\kappa_{\sup} = \lim_{\varepsilon \to 0} \max\left\{ \frac{\|N(x) - N(x')\|}{\|x - x'\|} \mid x, x' \in \partial K,\ x \neq x',\ \|x - x'\| \leq \varepsilon \right\},$$

et on définit $R_{\inf} = \kappa_{\sup}^{-1}$.

Si ∂K est C^2, alors cette définition coincide avec celle donnée plus haut, car les deux définitions donnent la constante Lipschitz de $N : \partial K \to \mathbb{R}^n$, où ∂K est muni de la métrique intrinsèque.

Si K a un bord de classe $\mathrm{C}^{1,1}$, on peut aussi définir l'application $\mathcal{N} : \partial K \times [0, \infty[\to \mathbb{R}^n$ par $\mathcal{N}(x,t) = x - tN(x)$.

Nous donnons une démonstration du théorème de Blaschke valable dans le cas $\mathrm{C}^{1,1}$.

Théorème 1.1 **(Blaschke).** — *Soit K un compact convexe d'intérieur non vide et dont le bord est de classe $C^{1,1}$. Pour tout $x \in \partial K$ et tout $r \in [0, R_{\inf}]$, on a*

$$d(x - rN(x), \partial K) = r.$$

Démonstration. — Pour $r \geq 0$, considérons l'application partielle

$$\mathcal{N}_r : \partial K \longrightarrow \mathbb{R}^n, \quad x \longmapsto \mathcal{N}(x,r).$$

On a pour tout $r \geq 0$ l'inclusion $\partial K_r^- \subset \mathcal{N}_r(\partial K)$. En effet, si $x \in K$ est tel que $d(x, \partial K) = r > 0$, soit $x_0 \in \partial K$ tel que $d(x, x_0) = r$, comme ∂K est C^1 et x_0 réalise le minimum de $d(y,x)$ pour $y \in \partial K$, on a que $x_0 - x$ est colinéaire à la normale $N(x_0)$, par conséquent $x = x_0 - rN(x_0) = \mathcal{N}_r(x_0)$, puisque $\|x - x_0\| = r$ et $N(x_0)$ pointe à l'extérieur de K.

Posons alors

$$R_0 = \sup\{r \in [0, R_{\inf}] \mid \forall\, x \in \partial K,\ \forall\, s \in [0,r],\ d(\mathcal{N}_s(x), \partial K) = s\}.$$

On a $\mathcal{N}(\partial K \times [0, R_0]) \subset K$. En effet, comme $d(\mathcal{N}_s(x), \partial K) = s$, pour $s \in [0, R_0]$, on a $\mathcal{N}_s(x) \notin \partial K$, pour $s \in\,]0, R_0]$. Or $N(x)$ pointe vers l'extéreiur de K, donc on a $\mathcal{N}_s(x) \in \mathring{K}$ pour $s > 0$ petit.

Supposons $R_0 < R_{\text{inf}}$. Fixons R tel que $R_0 < R < R_{\text{inf}} = \kappa_{\text{sup}}^{-1}$. Montrons qu'il existe $\varepsilon > 0$ tel que, pour tout $r \in [0, R]$, l'application partielle $\mathcal{N}_r : \partial K \to \mathbb{R}^n, x \mapsto \mathcal{N}(x, r)$ soit injective sur tout ensemble de ∂K de diamètre $\leq \varepsilon$. En effet, puisque $R\kappa_{\text{sup}} < 1$, il existe $k > \kappa_{\text{sup}}$ tel que $Rk < 1$. Par définition de κ_{sup}, il existe $\varepsilon > 0$ tel que, pour tout $x, x' \in K$, avec $\|x - x'\| \leq \varepsilon$, on ait $\|N(x) - N(x')\| \leq k\|x - x'\|$. Par conséquent

$$\|\mathcal{N}_r(x) - \mathcal{N}_r(x')\| \geq \|x - x'\| - r\|N(x) - N(x')\| \geq (1 - kr)\|x - x'\|.$$

L'injectivité résulte donc de $1 - kr \geq 1 - kR > 0$.

Toujours sous l'hypothèse $R_0 < R_{\text{inf}}$, montrons que l'on a $R_0 < \rho_{\max}$ et que l'application $\mathcal{N}_r$ est un homéomorphisme de ∂K sur ∂K_r^- pour tout $r \in [0, R_0]$.

En effet, pour $r \in [0, R_0]$, l'application $\mathcal{N}_r$ est localement injective, de plus, comme $d(\mathcal{N}_r(x), \partial K) = r$, pour tout $x \in \partial K$, l'application $\mathcal{N}_r$ envoie ∂K, qui est homéomorphe à $\mathbb{S}^{n-1}$ dans ∂K_r^-, la frontière, comme sous-ensemble de $\mathbb{R}^n$, de l'ensemble convexe compact K_r^-. Il résulte du théorème de Brouwer d'invariance du domaine que ∂K_r^- est homéomorphe à $\mathbb{S}^{n-1}$ (en particulier, on a $r < \rho_{\max}$ et donc $R_0 < \rho_{\max}$) et que $\mathcal{N}_r : \partial K \to \partial K_r^-$ est un revêtement. Comme les ∂K_r^- sont deux à deux disjoints, on voit que $\mathcal{N}$ induit un revêtement de $\partial K \times [0, R_0]$ (homéomorphe à $\mathbb{S}^{n-1} \times [0, 1]$) sur $\bigcup_{r \in [0, R_0]} \partial K_r^- = K \setminus \mathring{K}_{R_0}^-$ (lui aussi homéomorphe à $\mathbb{S}^{n-1} \times [0, 1]$). Or, sur $\partial K = \partial K_0^- = \mathcal{N}^{-1}(\partial K_0^-)$, l'application $\mathcal{N}$ est l'identité, donc le revêtement est de degré 1 et $\mathcal{N}$ est un homéomorphisme de $\partial K \times [0, R_0]$ sur $\{x \in K \mid 0 \leq d(x, \partial K) \leq R_0\}$.

Pour $r \in [0, R]$, l'application $\mathcal{N}_r$ est injective sur tout ensemble de diamètre $< \varepsilon$. Comme $\mathcal{N}_{R_0}$ est globalement injective, un raisonnement classique utilisant la compacité de ∂K montre qu'il existe $R' \in\,]R_0, R]$ tel que $\mathcal{N}_r$ reste globalement injective pour $r \in\,]R_0, R'[$. En effet, sinon on trouve $x_n, y_n \in \partial K$ et $r_n > R_0$ tels que $x_n \neq y_n, \mathcal{N}_{r_n}(x_n) = \mathcal{N}_{r_n}(y_n)$ et $\lim_{n\to\infty} r_n = R_0$, Par la compacité de ∂K, quitte à extraire des sous-suites, on peut supposer que $\lim_{n\to\infty} x_n = x_\infty$ et $\lim_{n\to\infty} y_n = y_\infty$. Par continuité, on obtient $\mathcal{N}_{R_0}(x_\infty) = \mathcal{N}_{R_0}(y_\infty)$ et donc $x_\infty = y_\infty$, par l'injectivité (globale) de $\mathcal{N}_{R_0}$. Pour n assez grand, on trouve donc $d(x_n, y_n) < \varepsilon$ et $r_n < R$. Ce qui contredit l'injectivité de $\mathcal{N}_{r_n}$ sur tout ensemble de diamètre $< \varepsilon$, pour $r \in [0, R]$, puisque $x_n \neq y_n$ et $\mathcal{N}_{r_n}(x_n) = \mathcal{N}_{r_n}(y_n)$.

Pour $R_0 \leq r < \min(R', \rho_{\max})$, on voit donc que $\mathcal{N}_r(\partial K)$ est homéomorphe à la sphère $\mathbb{S}^{n-1}$ et contient ∂K_r^- qui est homéomorphe à la même sphère. Le théorème de Brouwer d'invariance du domaine implique l'égalité $\mathcal{N}_r(\partial K) = \partial K_r^-$. On a donc $d(\mathcal{N}_r(x), \partial K) = r$, pour tout $x \in \partial K$ et tout $r \in\,]R_0, \min(R', \rho_{\max})[$, ce qui est contradictoire avec la définition de R_0. $\square$

Le corollaire suivant est en fait équivalent au théorème de Blaschke (*cf.* [**Th**, Corollaires p. 584] pour une autre démonstration) :

Corollaire 1.2. — *Sous les hypothèses de 1.1, la restriction de $\mathcal{N}$ à $\partial K \times [0, R_{\text{inf}}[$ est injective et à valeurs dans K.*

On a $K = \mathcal{N}(\partial K \times [0, R_{\text{inf}}[) \cup K^-_{R_{\text{inf}}}$, la réunion étant disjointe. De plus

$$K = \{x \in \mathbb{R}^n \mid d(x, K^-_{R_{\text{inf}}}) \leq R_{\text{inf}}\}.$$

Démonstration. — L'injectivité de $\mathcal{N}$ sur $\partial K \times [0, R_{\text{inf}}[$ est en fait déjà contenue dans la preuve de 1.1. Il n'est pas inintéressant de voir qu'elle résulte directement de l'énoncé de 1.1, par un argument géométrique bien connu.

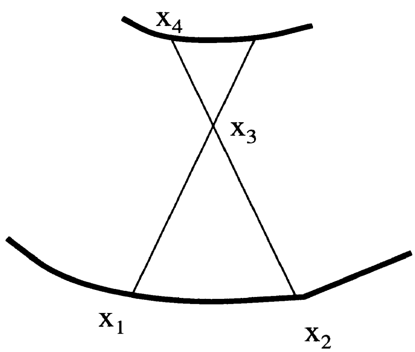


FIGURE 2

Supposons que x_1 et x_2 sont deux points distincts de ∂K tels qu'il existe $r_1, r_2 \in [0, R_{\text{inf}}[$ vérifiant $x_1 - r_1 N(x_1) = x_2 - r_2 N(x_2)$. Posons

$$x_3 = x_1 - r_1 N(x_1) = x_2 - r_2 N(x_2).$$

On trouve par le théorème de Blaschke $d(x_3, \partial K) = r_1 = r_2$. Posons par exemple $x_4 = x_1 - R_{\text{inf}} N(x_1)$ (*cf.* Figure 2). Les points x_4, x_3, x_2 ne sont pas colinéaires donc

$$\begin{aligned} d(x_4, x_2) &< d(x_4, x_3) + d(x_3, x_2) \\ &= \|x_1 - R_{\text{inf}} N(x_1) - (x_1 - r_1 N(x_1))\| + \|x_2 - r_2 N(x_2) - x_2\| \\ &= R_{\text{inf}} - r_1 + r_2 = R_{\text{inf}}. \end{aligned}$$

Par conséquent $R_{\text{inf}} = d(x_4, \partial K) \leq d(x_4, x_2) < R_{\text{inf}}$, ce qui est absurde.

Par la démonstration du théorème de Blaschke, l'application $\mathcal{N}_r$ est un homéomorphisme de ∂K sur $\partial K^-_r = \{x \in K \mid d(x, \partial K) = r\}$ pour $r < R_{\text{inf}}$. Le fait qu'alors $K = \mathcal{N}(\partial K \times [0, R_{\text{inf}}[) \cup K^-_{R_{\text{inf}}}$, la réunion étant disjointe, en résulte.

Du théorème de Blaschke, il résulte que $K \subset \{x \in \mathbb{R}^n \mid d(x, K^-_{R_{\text{inf}}}) \leq R_{\text{inf}}\}$. Montrons que cette inclusion est une égalité. En effet, si $x \notin K$ et $y \in K^-_{R_{\text{inf}}}$ est tel que $d(x, y) = d(x, K^-_{R_{\text{inf}}})$, le segment $[x, y]$, joignant x à y, coupe ∂K en un point que l'on notera y_0. Par définition de $K^-_{R_{\mathrm{inf}}}$, on a $d(y, y_0) \geq R_{\mathrm{inf}}$. Comme $d(x, y_0) > 0$, car $x \notin \partial K$, on obtient

$$d(x, K^-_{R_{\mathrm{inf}}}) = d(x, y) = d(x, y_0) + d(y, y_0) \geq d(x, y_0) + R_{\mathrm{inf}} > R_{\mathrm{inf}}.$$

Par conséquent $\{x \in \mathbb{R}^n \mid d(x, K^-_{R_{\mathrm{inf}}}) \leq R_{\mathrm{inf}}\} \subset K$. □

Nous aurons besoin de comprendre ce qui se passe dans le cas où $K^-_{R_{\mathrm{inf}}}$ est d'intérieur vide. Par exemple, si $K \subset \mathbb{R}^2$ est tel que $K^-_{R_{\mathrm{inf}}}$ soit d'intérieur vide, alors $K^-_{R_{\mathrm{inf}}}$ est un segment, éventuellement réduit à un point, et donc $K = \{x \in \mathbb{R}^n \mid d(x, K^-_{R_{\mathrm{inf}}}) \leq R_{\mathrm{inf}}\}$ est un stade à virages circulaires, qui n'est C^2 que si c'est un disque. Ce dernier résultat est vrai en toute dimension.

Complément 1.3. — *Si K est C^2, alors $K^-_{R_{\mathrm{inf}}}$ est d'intérieur vide si et seulement si K est une boule euclidienne (et dans, ce cas $K^-_{R_{\mathrm{inf}}}$ est réduit au centre de la boule).*

Posons pour simplifier $C = K^-_{R_{\mathrm{inf}}}$ et $R = R_{\mathrm{inf}} > 0$. Il suffit de démontrer la proposition suivante :

Proposition 1.4. — *Soit $C \subset \mathbb{R}^n$ est un compact convexe de dimension $\leq n-1$ non réduit à un point et $R > 0$. Considérons un segment S joignant deux points $c_1, c_2 \in C$ tels que $\|c_1 - c_2\| = \sup\{\|c - c'\| \mid c, c' \in C\}$, alors il existe un plan affine P contenant S et tel que $P \cap \{x \in \mathbb{R}^n \mid d(x, C) = R\}$ soit le stade à virages circulaires $\{x \in P \mid d(x, S) \leq R\}$. En particulier $\{x \in \mathbb{R}^n \mid d(x, C) = R\}$ n'est pas C^2.*

Démonstration. — Appelons $p_C : \mathbb{R}^n \to C$, la projection de $\mathbb{R}^n$ sur le compact convexe C, le point $p_C(x)$ est donc l'unique point de C dont la distance à x est $d(x, C)$.

Quitte à translater C, on peut supposer $c_1 = 0$. On posera $c = c_2$. Comme C n'est pas réduit à un point $c \neq 0$. On a :

$$p_C(tc) = \begin{cases} 0, & \text{pour } t \leq 0, \\ tc, & \text{pour } t \in [0,1], \\ c, & \text{pour } t \geq 1. \end{cases}$$

Montrons, par exemple, le dernier cas. Si $t \geq 1$ et $c' \in C$, on a

$$d(0, c) + d(c, tc) = d(0, tc) \leq d(0, c') + d(c', tc) \leq d(0, c) + d(c', tc),$$

d'où $d(c, tc) \leq d(c', tc)$ pour tout $c' \in C$, donc $p_C(tc) = c$.

Puisque C contient 0 et est de dimension $\leq n-1$, on peut trouver un hyperplan vectoriel H contenant C. Soit v un vecteur unitaire orthogonal à H. On a :

$$\forall x \in H,\ \forall \alpha \in \mathbb{R}, \quad p_C(x + \alpha v) = p_C(x).$$

En effet, comme $C \subset H$, on a, pour tout $c' \in C$, l'égalité

$$d(x + \alpha v, c') = (\|x - c'\|^2 + \alpha^2)^{1/2}.$$

Il suffit, alors, de minimiser sur $c' \in C$.

Dans le plan vectoriel P engendré par le segment $S = [0, c]$ et le vecteur v, on obtient alors :

$$\{x \in \mathbb{R}^n \mid d(x, C) \leq R\} \cap P = \{x \in P \mid d(x, S) \leq R\},$$

ce qui est le stade cherché.

Comme le segment S contient les deux points distincts 0 et c, le stade trouvé n'est pas C^2. Pour finir de montrer que $\{x \in \mathbb{R}^n \mid d(x, C) = R\}$ n'est pas C^2, il reste alors à remarquer que le plan P est transverse à $\{x \in \mathbb{R}^n \mid d(x, C) = R\}$, car il contient le point 0 qui est à l'intérieur du convexe $\{x \in \mathbb{R}^n \mid d(x, C) \leq R\}$, et par conséquent P ne peut être contenu dans aucun hyperplan de support de $\{x \in \mathbb{R}^n \mid d(x, C) \leq R\}$. $\square$

Remarque 1.5. — On peut montrer que si K est un compact convexe d'intérieur non vide avec ∂K de classe $C^{1,1}$, alors pour tout $r > K_{R_{\inf}}$, l'application $\mathcal{N} : \partial K \times [0, r] \to \mathbb{R}^n$ n'est pas injective.

Dans le cas où ∂K est suffisament lisse, c'est une conséquence du fait que l'on ne minimise plus les distances au-delà du premier point conjugué (ou focal). C'est aussi contenu dans la démonstration de Blaschke de son théorème.

Dans le cas où ∂K n'est que $\mathrm{C}^{1,1}$, cela vient d'un phénomène de régularité de la distance : on peut voir que si $\mathcal{N} : \partial K \times [0, r[\to \mathbb{R}^n$ est injective, alors $x \mapsto d(x, \partial K)$ est $\mathrm{C}^{1,1}$ sur l'ouvert $\mathcal{N}(\partial K \times]0, r[) \subset \mathbb{R}^n$.

2. Une minoration du volume d'un corps convexe de bord $\mathrm{C}^{1,1}$

Nous considérons toujours un compact convexe $K \subset \mathbb{R}^n$ d'intérieur non vide et de bord $\mathrm{C}^{1,1}$. L'application $N : \partial K \to \mathbb{S}^{n-1}$ est lipschitzienne. Par le théorème de Rademacher (*cf.* [**EG**, theorem 2, p. 81]), elle est dérivable pour σ-presque tout $x \in \partial K$, où σ est la mesure d'aire sur ∂K, sous-variété C^1 de $\mathbb{R}^n$ muni de sa métrique euclidienne usuelle. En tout point où $N(x)$ est dérivable, la dérivée $DN(x)$ est encore une application linéaire, symétrique et semi-définie positive de $N(x)^\perp$ dans lui-même (voir appendice). En un tel point, les valeurs propres, notées $\kappa_1(x), \ldots, \kappa_{n-1}(x)$, de $DN(x)$ sont donc ≥ 0, on les appelle les courbures principales. On a encore $\|DN(x)\| = \max(\kappa_1(x), \ldots, \kappa_{n-1}(x))$, puisque $DN(x)$ est symétrique. En un point où la fonction N est dérivable, on définit de même $H_i(x)$ comme la $i^{\text{ème}}$ fonction symétrique de $\kappa_1(x), \ldots, \kappa_{n-1}(x)$. Les fonctions $H_1, \ldots, H_{n-1}$ existent donc σ-presque partout. On définit $\mathcal{H}_i(K) = \int_{\partial K} H_i(x)\, d\sigma(x)$. On peut démontrer, dans ce cas aussi, le théorème de Gauss-Bonnet, c'est-à-dire que $\mathcal{H}_{n-1}(K) = s_{n-1}$ l'aire de la sphère unité $\mathbb{S}^{n-1}$ dans $\mathbb{R}^n$.

Théorème 2.1. — *Si $K \subset \mathbb{R}^n$ est un compact convexe d'intérieur non vide et dont le bord est de classe $C^{1,1}$, alors, on a :*

$$V(K) \geq R_{\inf} S(K) + \sum_{i=1}^{n-1} (-1)^i R_{\inf}^{i+1} \frac{\mathcal{H}_i(K)}{i+1},$$

où $V(K)$ est le volume de K et $S(K)$ est l'aire de son bord ∂K.

De plus, cette inégalité est une égalité si et seulement si

$$K = \{x \in \mathbb{R}^n \mid d(x, C) \leq R\},$$

où C est un compact convexe de $\mathbb{R}^n$ de dimension $n-1$.

En particulier, si K est de bord C^2, l'inégalité est une égalité si et seulement si K est une boule euclidienne.

Démonstration. — On sait que $\mathcal{N} : \partial K \times [0, R_{\inf}[\to K$ est injectif. Comme $\mathcal{N}(x, r) = x - rN(x)$, il est clair que $\mathcal{N}$ est lipschitzienne, donc la dérivée existe $\sigma \times dr$-presque partout et $D\mathcal{N}(x, r) : N(x)^{\perp} \times \mathbb{R} \to \mathbb{R}^n$ est donnée par

$$D\mathcal{N}(x, r)(u, t) = u - rDN(x)(u) + tN(x), \quad \sigma \times dr\text{-presque partout.}$$

Par conséquent, en $\sigma \times dr$-presque tout (x, r), le jacobien $\mathrm{Jac}_{x,r}(\mathcal{N})$ vaut

$$\det(\mathrm{Id}_{N(x)^{\perp}} - rDN(x)) = \prod_{j=1}^{n-1} (1 - r\kappa_j(x)) = 1 + \sum_{i=1}^{n-1} (-1)^i r^i H_i(x).$$

La formule du changement de variable, qui est valide car $\mathcal{N}$ est lipschitzien injectif (*cf.* [**EG**, theorem 1, p. 96]), donne :

$$\begin{aligned} V(\mathcal{N}(\partial K \times [0, R_{\inf}[)) &= \int_{\partial K \times [0, R_{\inf}[} \left[1 + \sum_{i=1}^{n-1} (-1)^i r^i H_i(x) \right] d\sigma(x) dr, \\ &= R_{\inf} S(K) + \sum_{i=1}^{n-1} (-1)^i R_{\inf}^{i+1} \frac{\mathcal{H}_i(K)}{i+1}. \end{aligned}$$

Or par 1.2, l'ensemble K contient $\mathcal{N}(\partial K \times [0, R_{\inf}[)$, d'où l'inégalité :

$$V(K) \geq R_{\inf} S(K) + \sum_{i=1}^{n-1} (-1)^i R_{\inf}^{i+1} \frac{\mathcal{H}_i}{i+1}.$$

Toujours par 1.2, cette inégalité est une égalité si et seulement si $V(K^-_{R_{\inf}}) = 0$. Comme $K^-_{R_{\inf}}$ est un convexe, le cas d'égalité est donc équivalent à $K^-_{R_{\inf}}$ de dimension $\leq n-1$. Il suffit alors d'appliquer 1.3 pour conclure. □

Dans le cas où $n = 2$, il n'y a qu'une intégrale de courbure $\mathcal{H}_1(K) = \mathcal{H}_{n-1}(K)$ qui vaut 2π par le théorème de Gauss–Bonnet. On trouve, alors que l'aire de K est supérieure à $R_{\inf} l - \pi R_{\inf}^2$, où l est la longueur de ∂K. C'est le théorème 0.2.

3. Complément

On considère, un compact convexe d'intérieur non vide et de bord de classe C^2. On peut se demander dans quel cas $\mathcal{N}$ est un plongement sur $\partial K \times [0, R_{\mathrm{inf}}]$. Bien sûr, le convexe K ne peut pas être une boule euclidienne, car alors $\mathcal{N}_{R_{\mathrm{inf}}}$ est constante. On peut obtenir d'autres convexes de bord C^∞ avec $\mathcal{N}_{R_{\mathrm{inf}}}$ non-injective, en prenant, par exemple dans le plan, un convexe avec un arc de cercle non-dégénéré dans son bord. Le résultat suivant est donc optimal :

Théorème 3.1. — *Soit $K \subset \mathbb{R}^n$ un convexe compact d'intérieur non vide et dont le bord est C^ω (*i.e.* analytique réel). Si K n'est pas une boule euclidienne alors $\mathcal{N}$ est un plongement sur $\partial K \times [0, R_{\mathrm{inf}}]$.*

En fait, en posant $C = K^-_{R_{\mathrm{inf}}}$ et $R = R_{\mathrm{inf}}$, ceci découle du théorème :

Théorème 3.2. — *Soit $C \subset \mathbb{R}^n$ un compact convexe non réduit à un point. Notons $p_C : \mathbb{R}^n \to C$ la projection, définie par $d(x, C) = d(x, p_C(x))$. Si $R > 0$ est tel que $\partial C_R = \{x \in \mathbb{R}^n \mid d(x, C) = R\}$ soit C^ω, alors p_C est injective sur ∂C_R.*

Démonstration. — Rappelons que $p_C(x)$ est aussi caractérisé par la propriété suivante :

Le point $p_C(x)$ est le seul point $c \in C$ tel que $\forall\, y \in C$, $\langle x - c, y - c\rangle \leq 0$.

Il en résulte que pour tout $c \in C$, l'ensemble $p_C^{-1}(c)$ est un cône convexe de sommet c. Notons F_c, le sous-espace affine engendré par $p_C^{-1}(c)$. La convexité de $p_C^{-1}(c)$ implique que l'intérieur de $p_C^{-1}(c)$ en tant que sous ensemble de F_c est non vide.

Montrons que F_c intersecte l'intérieur $\mathring{C}$ de C. On sait déjà par 1.4, que $\mathring{C}$ n'est pas vide. Si F_c n'intersectait pas $\mathring{C}$, il serait contenu dans un hyperplan affine H de support de C. Notons H_1 et H_2 les deux composantes connexes de $\mathbb{R}^n \setminus H$ avec $C \subset H_1 \cup H$. La caractérisation de p_C donnée plus haut montre que la demi droite D_2, issue de c, incluse dans H_2 et perpendiculaire à H en c, serait contenue dans $p_C^{-1}(c)$. Ce qui est absurde, car $p_C^{-1}(c) \subset F_c \subset H$.

Montrons que la dimension q de F_c est nécessairement 1. En effet, comme F_c intersecte l'intérieur du convexe $C_R = \{x \in \mathbb{R}^n \mid d(x, C)\}$, il est transverse au bord ∂C_R, par conséquent $F_c \cap \partial C_R$ est une hypersurface C^ω de F_c. Cette hypersurface est homéomorphe à $\mathbb{S}^{q-1}$, car c'est le bord du compact convexe $F_c \cap C_R$ d'intérieur non vide dans l'espace affine F_c de dimension q. De plus $F_c \cap \partial C_R$ contient l'intersection $p_C^{-1}(c) \cap \{x \in F_c \mid ||x - c|| = R\}$, or cette intersection est d'intérieur non vide dans la sphère $\{x \in F_c \mid ||x - c|| = R\}$ de dimension $q - 1$, car le cône $p_C^{-1}(c)$ est de sommet c et d'intérieur non vide comme sous-ensemble de F_c. La fonction analytique réelle $x \mapsto ||x - c||^2$ est donc nulle sur un ouvert non vide de $F_c \cap \partial C_R$ qui est une variété C^ω homéomorphe à $\mathbb{S}^{q-1}$. Si $q \geq 2$, on voit que $F_c \cap \partial C_R$ est connexe, donc $x \mapsto ||x - c||^2$ est constante sur $F_c \cap \partial C_R$. Comme $F_c \cap \partial C_R$ est une sous-variété compacte de codimension 1 dans F_c, nous obtenons

$$F_c \cap \partial C_R = \{x \in F_c \mid \|x - c\| = R\}.$$

Nous en concluons que $F_c \cap C = \{c\}$, ce qui est impossible, car $F_c \cap \mathring{C} \neq \varnothing$.

Par conséquent, le sous-espace affine F_c est une droite. Comme $F_c \cap \mathring{C} \neq \varnothing$, le compact convexe $C \cap F_c$ est un segment d'intérieur non vide, donc nous ne pouvons pas avoir $p_C^{-1}(c) = F_c$; ce qui force $p_C^{-1}(c) \cap \partial C_R$ à être réduit à un seul point. □

Remarque 3.3. — 1) Dans la preuve précédente, nous avons rappelé le fait bien connu suivant :

Si $C \subset \mathbb{R}^n$ est un compact convexe et $p_C : \mathbb{R}^n \to C$ est la projection, alors $p_C^{-1}(c)$ est un cône de sommet c pour tout $c \in C$. De plus, quand C est d'intérieur non vide, on peut voir que $p_C^{-1}(c)$ n'est pas un sous espace affine, car le sous espace affine F_c engendré par $p_C^{-1}(c)$ coupe l'intérieur de C.

Il s'ensuit que quand C est d'intérieur non vide les préimages de p_C restreint à un ensemble de la forme $\partial C_R = \{x \in \mathbb{R}^n \mid d(x, C) = R\}$ sont toutes homéomorphes à des disques de dimension (dépendant du point) $\leq n-1$; en particulier, ces préimages sont connexes.

2) On voit donc que si $K \subset \mathbb{R}^n$ est un compact, convexe, d'intérieur non vide, de bord $\mathrm{C}^{1,1}$ et tel que $K^-_{R_{\mathrm{inf}}}$ soit aussi d'intérieur non vide, alors, les préimages de $\mathcal{N}_{R_{\mathrm{inf}}}$ sont des disques topologiques. Ce qui n'est pas très étonnant car $\mathcal{N}_{R_{\mathrm{inf}}}$ est la limite d'homéomorphismes et par conséquent les préimages devraient être de «shape», au sens de Borsuk, trivial.

3) On aurait pu utiliser ce qui a été dit en 1), pour donner une démonstration légèrement différente du théorème de Blaschke.

Le théorème 3.1 a un corollaire intéressant.

Pour $k \geq 2$, introduisons l'espace $\mathcal{C}_k$ des hypersurfaces de classe C^k de $\mathbb{R}^n$ compactes et strictement convexes (c'est-à-dire ayant toutes leurs courbures principales > 0 en tout point). On munit $\mathcal{C}_k$ de la topologie C^k. Le sous-ensemble $\mathcal{C}_k$ est ouvert dans l'espace des hypersurfaces compactes C^k de $\mathbb{R}^n$, muni de la topologie C^k, c'est donc un espace de Baire. Si on enlève de $\mathcal{C}_k$ le sous-ensemble formé par les hypersurfaces qui sont des boules euclidiennes, on trouve encore un ouvert de l'espace des hypersurfaces compactes C^k de $\mathbb{R}^n$ muni de la topologie C^k. Cet ouvert est dense dans $\mathcal{C}_k$. On en déduit que les hypersurfaces C^ω de $\mathcal{C}_k$ qui ne sont pas des boules euclidiennes forment un ensemble dense dans $\mathcal{C}_k$. Si $S \in \mathcal{C}_k$, on définit l'application $N^S : S \to \mathbb{S}^{n-1}$, par $N^S(x)$ est la normale unitaire à S en x orientée vers l'extérieur. On définit aussi l'application $\mathcal{N}^S : S \times \mathbb{R} \to \mathbb{R}^n$, par $\mathcal{N}^S(x,r) = x - rN^S(x)$. Les applications N^S et $\mathcal{N}^S$ sont C^{k-1}. De plus, puisque $k \geq 2$, on peut définir R^S_{inf} comme dans l'introduction, il est clair que R^S_{inf} dépend continument de S dans la topologie C^2.

Corollaire 3.4. — *Si $k \geq 2$, pour un S générique (au sens de Baire) dans $\mathcal{C}_k$, muni de la topologie C^k, l'application $\mathcal{N}^S : S \times [0, R^S_{\text{inf}}] \to \mathbb{R}^n$ est injective.*

Démonstration. — Considérons $\mathcal{N}^S_{R^S_{\text{inf}}} : S \to \mathbb{R}^n$. Il n'est pas difficile de voir que, pour $\varepsilon > 0$:

$$\mathcal{U}_\varepsilon = \{S \in \mathcal{C}_k \mid \text{ les préimages de } \mathcal{N}^S_{R^S_{\text{inf}}} \text{ sont de diamètre } < \varepsilon\}$$

est un ouvert de $\mathcal{C}_k$, car $k \geq 2$. Cet ouvert est dense, puisqu'il contient toutes les hypersurfaces C^ω appartenant à $\mathcal{C}_k$ et qui ne sont pas des boules euclidiennes. Par conséquent, l'intersection $\bigcap_{n \geq 1} \mathcal{U}_{1/n}$ est un G_δ dense dans l'espace de Baire $\mathcal{C}_k$. Or cette intersection est précisément l'ensemble des hypersurfaces $S \in \mathcal{C}_k$ telles que $\mathcal{N}^S : S \times [0, R^S_{\text{inf}}] \to \mathbb{R}^n$ soit injective. □

4. Appendice : symétrie et positivité de DN

Dans cet appendice, nous montrons que DN est symétrique et semi-définie positive. Ce fait est bien connu. Ainsi que nous allons le voir, l'argument généralement donné dans le cas où ∂K est C^2 s'adapte aisément.

Supposons que N soit dérivable en $x_0 \in \partial K$. Quitte à changer de système de coordonnées, on peut supposer que $x_0 = 0$ et que $N(x_0) = N(0)$ est $(0, \dots, 0, 1)$, le dernier vecteur de la base canonique de $\mathbb{R}^n$. Par le théorème des fonctions implicites, au voisinage de 0, la sous-variété ∂K de classe C^1 est le graphe d'une fonction φ de classe C^1 définie sur un voisinage ouvert U de 0 dans $\mathbb{R}^{n-1}$. Notons $(x_1, \dots, x_{n-1})$ les coordonnées canoniques d'un point x de $\mathbb{R}^{n-1}$. Dans la carte

$$f : U \longrightarrow \mathbb{R}^n, \quad x = (x_1, \dots, x_{n-1}) \longmapsto (x_1, \dots, x_{n-1}, \varphi(x_1, \dots, x_{n-1})),$$

on a $N(x) = X(x)/\|X(x)\|$, avec $X(x) = (-\partial_1\varphi(x), \dots, -\partial_{n-1}\varphi(x), 1)$, où $\partial_j\varphi$ est la dérivée partielle de φ par rapport à x_j. En particulier, la dernière coordonnée de $N(x)$ est $\|X(x)\|^{-1}$. Elle est dérivable en 0, puisque N l'est. Il en résulte que X est dérivable en 0 et par conséquent φ a une dérivée seconde en 0. Cette dérivée seconde $D^2\varphi(0)$ est symétrique par le théorème de Schwarz. Montrons alors que la forme bilinéaire associée à $DN(0)$ n'est rien d'autre que $-D^2\varphi(0)$. Sur U, le produit scalaire $\langle \partial_i f(x), N(x) \rangle$ est identiquement nul. Puisque f a aussi une dérivée seconde en 0, par dérivation, on obtient

$$\langle \partial_i f(0), \partial_j N(0) \rangle = -\langle \partial^2_{j,i} f(0), N(0) \rangle = -\partial^2_{j,i}\varphi(0).$$

Pour montrer que $DN(0)$ est semi-définie positive, il suffit de voir que $D^2\varphi(0)$ est semi-définie négative. Puisque $N(0)$ est le dernier vecteur de la base canonique de $\mathbb{R}^n$, l'hypersurface convexe ∂K est située dans le demi-espace formé par les points $y \in \mathbb{R}^n$ tels que $\langle y, N(0) \rangle \leq 0$. Par conséquent, on a $\varphi \leq 0$ sur U, or $\varphi(0) = 0$. Il en résulte que φ a un maximum en 0 et donc $D^2\varphi(0)$ est semi-définie négative.

Références

[Bl] W. Blaschke. *Kreis und Kugel.* Chelsea, New York, 1949.

[EG] L.C. Evans & R.F. Gariepy. *Measure Theory & Fine properties of functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, Ann Arbor & London, 1992.

[Ga] M.G. Gage. *An isoperimetric inequality with applications to curve shortening.* Duke Math. J. **50** (1983), 1225–1229.

[Ho] L. Hörmander. *Notions of Convexity.* Progress in Mathematics vol. 127. Birkhaüser, Boston, Basel & Berlin, 1994.

[In_1] N. Innami. *Isoperimetric inequalities depending on injectivity radius from boundary,* in Complex Structures and Vector Fields, edited by S. Dimiev & K.Segikawa. World Scientific, Singapore, 1995.

[In_2] N. Innami. *Volume, surface area and inward injectivity radius.* Preprint, Niigata University.

[Le] K. Leichtweiss. *Convexity and Differential Geometry,* in Handbook of Convex Geometry **B**, edited by P.M. Gruber & J.M. Wills. North-Holland, Amsterdam, London, New York & Tokyo,1993.

[Sc] R. Schneider *Convex Bodies : The Brunn-Minkowski Theory.* Encyclopedia in Mathematics and its Applications **44**, Cambridge Univ. Press, Cambridge, 1993.

[Th] R. Thom. *Sur le cut-locus d'une variété plongée.* J. Differential Geometry, **6** (1972), 577–586.

A. FATHI, Unité de Mathématiques Pures et Appliquées, École Normale Supérieure de Lyon, 46, allée d'Italie, 69364 Lyon Cedex 07, France • *E-mail :* `afathi@umpa.ens-lyon.fr`

MICHAEL JAKOBSON
SHELDON NEWHOUSE

Asymptotic measures for hyperbolic piecewise smooth mappings of a rectangle

Astérisque, tome 261 (2000), p. 103-159
<http://www.numdam.org/item?id=AST_2000__261__103_0>

Astérisque
261, 2000, p. 103–159

ASYMPTOTIC MEASURES FOR HYPERBOLIC PIECEWISE SMOOTH MAPPINGS OF A RECTANGLE

by

Michael Jakobson & Sheldon Newhouse

To Adrien Douady on the occasion of his sixtieth birthday

Abstract. — We prove the existence of Sinai-Ruelle-Bowen measures for a class of C^2 self-mappings of a rectangle with unbounded derivatives. The results can be regarded as a generalization of a well-known one dimensional Folklore Theorem on the existence of absolutely continuous invariant measures. In an earlier paper [8] analogous results were stated and the proofs were sketched for the case of invertible systems. Here we give complete proofs in the more general case of noninvertible systems, and, in particular, develop the theory of stable and unstable manifolds for maps with unbounded derivatives.

1. Folklore Theorem and SRB Measures

A well-known Folklore Theorem in one-dimensional dynamics can be formulated as follows.

Folklore Theorem. — *Let $I = [0,1]$ be the unit interval, and suppose $\{I_1, I_2, \dots\}$ is a countable collection of disjoint open subintervals of I such that $\bigcup_i I_i$ has the full Lebesgue measure in I. Suppose there are constants $K_0 > 1$ and $K_1 > 0$ and mappings $f_i : I_i \to I$ satisfying the following conditions.*

(1) *f_i extends to a C^2 diffeomorphism from* Closure(I_i) *onto* $[0,1]$, *and*

$$\inf_{z \in I_i} |Df_i(z)| > K_0 \quad \textit{for all } i.$$

(2) $\sup_{z\in I_i} \dfrac{|D^2 f_i(z)|}{|Df_i(z)|} |I_i| < K_1$ *for all i.*

1991 ***Mathematics Subject Classification.*** — 37C40, 37A05, 28D05, 28D20, 37D25, 37D45.
Key words and phrases. — entropy, hyperbolic, ergodic, absolutely continuous measure, SRB-measure.

M. J. partially supported by NSF Grant 9303369.

where $|I_i|$ *denotes the length of* I_i. *Then, the mapping* $F(z)$ *defined by* $F(z) = f_i(z)$ *for* $z \in I_i$, *has a unique invariant ergodic probability measure* μ *equivalent to Lebesgue measure on* I.

For the proof of the Folklore theorem and the ergodic properties of μ see for example **[2]** and **[14]**.

In an earlier paper **[8]** we presented an analog of this theorem for piecewise C^2 diffeomorphisms with unbounded derivatives with proof sketched. We now wish to give a more general version of the results in **[8]**. We refer the reader to that paper for relevant remarks and references.

Let $\widetilde{Q}$ be a Borel subset of the unit square Q in the plane $\mathbf{R}^2$ with positive Lebesgue measure, and let $F : \widetilde{Q} \to \widetilde{Q}$ be a Borel measurable map. An F-invariant Borel probablility measure μ on Q is called a *Sinai – Ruelle – Bowen* measure (or SRB-measure) for F if μ is ergodic and there is a set $A \subset \widetilde{Q}$ of positive Lebesgue measure such that for $x \in A$ and any continuous real-valued function $\phi : Q \to \mathbf{R}$, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \phi(F^k x) = \int \phi d\mu. \tag{1}$$

The set of all points x for which (1) holds is called the *basin* of μ.

Note that if μ is an SRB measure, and m_1 is the normalized Lebesgue measure on its basin, then the bounded convergence theorem gives the weak convergence of the averages $\frac{1}{n}\sum_{k=0}^{n-1} F_\star^k m_1$ of the iterates of m_1 to μ. Hence, SRB measures occur as limiting mass distributions of sets of positive Lebesgue measure. This fact makes them natural objects to study.

We are interested in giving conditions under which certain two-dimensional maps F which piecewise coincide with hyperbolic diffeomorphisms f_i have SRB measures. As in the one-dimensional situation there is an essential difference between a finite and an infinite number of f_i. In the case of an infinite number of f_i, their derivatives grow with i and relations between first and second derivatives become crucial.

2. Hyperbolicity and geometric conditions

Consider a countable collection $\xi = \{E_1, E_2, \ldots,\}$ of full height closed curvilinear rectangles in Q. Assume that each E_i lies inside a domain of definition of a C^2 diffeomorphism f_i which maps E_i onto its image $S_i \subset Q$. We assume each E_i connects the top and the bottom of Q. Thus each E_i is bounded from above and from below by two subintervals of the line segments

$$\{(x,y) : y = 1,\ 0 \le x \le 1\} \quad \text{and} \quad \{(x,y) : y = 0,\ 0 \le x \le 1\}.$$

We assume that the left and right boundaries of E_i are graphs of smooth functions $x^{(i)}(y)$ with $|dx^{(i)}/dy| \le \alpha$ where α is a real number satisfying $0 < \alpha < 1$. We further assume that the images $f_i(E_i) = S_i$ are narrow strips connecting the left and right sides of Q and that they are bounded on the left and right by the two subintervals of the line segments

$$\{(x,y) : x = 0,\ 0 \leq y \leq 1\} \quad \text{and} \quad \{(x,y) : x = 1,\ 0 \leq y \leq 1\}$$

and above and below by the graphs of smooth functions $Y^i(X)$, $\left|dY^{(i)}/dX\right| \leq \alpha$. We will see later that the upper bounds on derivatives $\left|dx^{(i)}/dy\right| \leq \alpha$ and $\left|dY^{(i)}/dX\right| \leq \alpha$ follow from hyperbolicity conditions that we formulate below.

We call the $E_i^{\prime}s$ *posts*, the $S_i^{\prime}s$ *strips*, and we say the $E_i^{\prime}s$ are *full height* in Q while the $S_i^{\prime}s$ are *full width* in Q.

For $z \in Q$, let ℓ_z be the horizontal line through z. We define

$$\delta_z(E_i) = \operatorname{diam}(\ell_z \cap E_i), \quad \delta_{i,\max} = \max_{z\in Q} \delta_z(E_i), \quad \delta_{i,\min} = \min_{z\in Q} \delta_z(E_i).$$

We assume the following *geometric conditions*

G1. $\operatorname{int} E_i \cap \operatorname{int} E_j = \varnothing$ for $i \neq j$.
G2. $\operatorname{mes}(Q \setminus \cup_i \operatorname{int} E_i) = 0$ where mes stands for Lebesgue measure,
G3. $-\sum_i \delta_{i,\max} \log \delta_{i,\min} < \infty$.

We emphasize that the strips S_i can intersect in an arbitrary fashion, differently from condition G3 in ([**8**]).

In the standard coordinate system for a map $F : (x,y) \to (F_1(x,y), F_2(x,y))$ we use $DF(x,y)$ to denote the differential of F at some point (x,y) and F_{jx}, F_{jy}, F_{jxx}, F_{jxy}, etc., for partial derivatives of F_j, $j = 1,2$.

Let $J_F(z) = |F_{1x}(z)F_{2y}(z) - F_{1y}(z)F_{2x}(z)|$ be the absolute value of the Jacobian determinant of F at z.

Hyperbolicity conditions. — There exist constants $0 < \alpha < 1$ and $K_0 > 1$ such that for each i the map

$$F(z) = f_i(z) \quad \text{for } z \in E_i$$

satisfies

H1. $\left|F_{2x}(z)\right| + \alpha\left|F_{2y}(z)\right| + \alpha^2\left|F_{1y}(z)\right| \leq \alpha\left|F_{1x}(z)\right|$
H2. $\left|F_{1x}(z)\right| - \alpha\left|F_{1y}(z)\right| \geq K_0$.
H3. $\left|F_{1y}(z)\right| + \alpha\left|F_{2y}(z)\right| + \alpha^2\left|F_{2x}(z)\right| \leq \alpha\left|F_{1x}(z)\right|$
H4. $\left|F_{1x}(z)\right| - \alpha\left|F_{2x}(z)\right| \geq J_F(z)K_0$.

For a real number $0 < \alpha < 1$, we define the cones

$$\begin{aligned} K_\alpha^u &= \{(v_1,v_2) : \left|v_2\right| \leq \alpha\left|v_1\right|\} \\ K_\alpha^s &= \{(v_1,v_2) : \left|v_1\right| \leq \alpha\left|v_2\right|\} \end{aligned}$$

and the corresponding cone fields $K_\alpha^u(z), K_\alpha^s(z)$ in the tangent spaces at points $z \in \mathbf{R}^2$.

Unless otherwise stated, we use the max norm on $\mathbf{R}^2$, $\left|(v_1,v_2)\right| = \max(\left|v_1\right|, \left|v_2\right|)$.

The following simple proposition relates conditions H1-H4 above with the usual definition of hyperbolicity in terms of cone conditions. It shows that conditions H1 and H2 imply that the K_α^u cone is mapped into itself by DF and expanded by a factor no smaller than K_0 while H3 and H4 imply that the K_α^s cone is mapped into itself by DF^{-1} and expanded by a factor no smaller than K_0.

Proposition 2.1. — *Under conditions H1-H4 above, we have*

$$DF(K_\alpha^u) \subseteq K_\alpha^u \tag{2}$$

$$v \in K_\alpha^u \Rightarrow |DFv| \geq K_0|v| \tag{3}$$

$$DF^{-1}(K_\alpha^s) \subseteq K_\alpha^s \tag{4}$$

$$v \in K_\alpha^s \Rightarrow |DF^{-1}v| \geq K_0|v| \tag{5}$$

Proof. — H1 implies (2):

Let $v = (v_1, v_2) \in K_\alpha^u$. Then, $|v| = |v_1|$ since $\alpha < 1$ and $|v_2| \leq \alpha|v_1|$.

Write $DF(v_1, v_2) = (F_{1x}v_1 + F_{1y}v_2, F_{2x}v_1 + F_{2y}v_2) = (u_1, u_2)$.

Then, using H1, we have

$$\begin{aligned}
|u_2| &= |F_{2x}v_1 + F_{2y}v_2| \\
&\leq |F_{2x}||v_1| + |F_{2y}|\alpha|v_1| \\
&\leq |v_1|(|F_{2x}| + |F_{2y}|\alpha) \\
&\leq |v_1|(\alpha|F_{1x}| - |F_{1y}|\alpha^2) \\
&\leq \alpha|F_{1x}v_1 + F_{1y}v_2| \\
&= \alpha|u_1|
\end{aligned}$$

proving (2).

H2 implies (3):

Now, let $v = (v_1, v_2)$ be a unit vector in K_α^u, so that $|v| = |v_1| = 1$ and $|v_2| \leq \alpha$.

Using H2 and the fact that $DF(v) \in K_\alpha^u$, we have

$$\begin{aligned}
|DF(v)| &= |u_1| \\
&= |F_{1x}v_1 + F_{1y}v_2| \\
&\geq |F_{1x}| - \alpha|F_{1y}| \\
&\geq K_0
\end{aligned}$$

which is (3).

The proofs that H3 and H4 imply (4) and (5) are similar using the fact that

$$DF^{-1} = \frac{1}{J_z}\begin{pmatrix} F_{2y} & -F_{1y} \\ -F_{2x} & F_{1x} \end{pmatrix}$$

This completes our proof of Proposition 2.1. □

Remark. — In ([**8**]) different hyperolicity conditions were assumed which implied the invariance of cones and uniform expansion with respect to the sum norm $|v| = |v_1| + |v_2|$ (see [**3**] and [**7**] for related hyperbolicity conditions). The methods here can be adapted to work under the assumptions of ([**8**]).

The map

$$F(z) = f_i(z) \quad \text{for } z \in \operatorname{int} E_i$$

is defined almost everywhere on Q. Let $\widetilde{Q}_0 = \bigcup_i \operatorname{int} E_i$, and, define $\widetilde{Q}_n, n > 0$, inductively by $\widetilde{Q}_n = \widetilde{Q}_0 \cap F^{-1}\widetilde{Q}_{n-1}$. Let $\widetilde{Q} = \bigcap_{n\geq 0} \widetilde{Q}_n$ be the set of points whose forward orbits always stay in $\bigcup_i \operatorname{int} E_i$. Then, $\widetilde{Q}$ has full Lebesgue measure in Q, and F maps $\widetilde{Q}$ into itself.

The hyperbolicity conditions H1–H4 imply the estimates on the derivatives of the boundary curves of E_i and S_i which we described earlier. They also imply that any intersection $f_iE_i \cap E_j$ is full width in E_j. Further, $E_{ij} = E_i \cap f_i^{-1}E_j$ is a full height subpost of E_i and $S_{ij} = f_jf_iE_{ij}$ is a full width substrip in Q.

Given a finite string $i_0 \dots i_{n-1}$, indexed by non-negative integers, we define inductively

$$E_{i_0\dots i_{n-1}} = E_{i_0} \cap f_{i_0}^{-1} E_{i_1 i_2\dots i_{n-1}}.$$

Then, each set $E_{i_0\dots i_{n-1}}$ is a full height subpost of E_{i_0}.

Analogously, for a string $i_{-n+1} \dots i_0$ indexed by non-positive integers, we define.

$$S_{i_{-n+1}\dots i_0} = f_{i_0}(S_{i_{-n+1}\dots i_{-1}} \cap E_{i_0})$$

and get that $S_{i_{-n+1}\dots i_0}$ is a full width strip in Q. It is easy to see that $S_{i_{-n+1}\dots i_0} = (f_{i_0} \circ f_{i_{-1}} \circ \dots \circ f_{i_{-n+1}})(E_{i_{-n+1}\dots i_0})$ and that $f_{i_0}^{-1}(S_{i_{-n+1}\dots i_0})$ is a full-width strip in E_{i_0}.

For infinite strings, we have the following Proposition.

Proposition 2.2. — *Any C^1 map F satisfying the above geometric conditions G1-G3 and hyperbolicity conditions H1–H4 has a "topological attractor"*

$$\Lambda = \bigcup_{\dots i_{-n+1}\dots i_{-1} i_0} \bigcap_{k\geq 0} S_{i_{-k}\dots i_0}.$$

The infinite intersections $\bigcap_{k=0}^{\infty} S_{i_{-k}\dots i_0}$ define C^1 curves $\gamma = y(x)$, $|dy/dx| \leq \alpha$ which are the unstable manifolds for the points of the attractor. The infinite intersections $\bigcap_{k=0}^{\infty} E_{i_0\dots i_k}$ define C^1 curves $x(y)$, $|dx/dy| \leq \alpha$ which are the stable manifolds for the points of the attractor.

Proposition 2.2 is a well known fact in hyperbolic theory. For example it follows from Theorem 1 in **[3]**. See also **[10]**.

Remark 2.3. — The distortion condition D1 and *distortion estimates* below imply that if our maps f_i are C^2, then the unstable manifolds are actually C^2. Similar conditions on the inverses of f_i imply that the stable manifolds are C^2. There are analogous conditions (see section 6) to guarantee that the invariant manifolds are C^r for $r \geq 2$.

Remark 2.4. — The union of the stable manifolds contains the above set $\widetilde{Q}$ which has full measure in Q. The trajectories of all points in $\widetilde{Q}$ converge to Λ. That is the reason to call Λ a topological attractor, although F is not typically a well-defined mapping on all of Λ. However the convergence of Birkhoff averages to the unique SRB measure is a much stronger property. Condition D1 is natural in this context and may be necessary for the existence of the SRB measure. At present, we need to assume condition G3. This is used to prove absolute continuity of the stable foliation as in Section 10. It also implies that our SRB measure has finite entropy. We do not know if condition G3 is actually necessary for our results.

3. Distortion conditions and the main theorem

As we have a countable number of domains the derivatives of f_i grow. We will need to formulate certain assumptions on the second derivatives. Unless otherwise stated, we will use the norm $|v| = \max(|v_1|, |v_2|)$ on vectors $v = (v_1, v_2)$, and the associated distance function $d((x,y),(x_1,y_1)) = \max(|x - x_1|, |y - y_1|)$.

As above, for a point $z \in Q$, let l_z denote the horizontal line through z, and if $E \subseteq Q$, let $\delta_z(E)$ denote the diameter of the horizontal section $l_z \cap E$. We call $\delta_z(E)$ the $z - width$ of E.

In given coordinate systems we write $f_i(x,y) = (f_{i1}(x,y), f_{i2}(x,y))$. We use $f_{ijx}, f_{ijy}, f_{ijxx}, f_{ijxy}$, etc. for partial derivatives of $f_{ij}, j = 1,2$.

We define

$$|D^2 f_i(z)| = \max_{j=1,2,(k,l)=(x,x),(x,y),(y,y)} |f_{ijkl}(z)|.$$

Next we formulate distortion conditions. These will be used to control the fluctuation of the derivatives of iterates of F along vectors in K^u_α as in Lemma 7.1 and Proposition 8.1 below.

Suppose there is a constant $C_0 > 0$ such that the following *distortion condition* holds

D1
$$\sup_{z\in E_i, i\geq 1} \frac{|D^2 f_i(z)|}{|f_{i1x}(z)|}\delta_z(E_i) < C_0.$$

Theorem 3.1. — *Let F be a piecewise smooth mapping as above satisfying the geometric conditions G1–G3, the hyperbolicity conditions H1–H4, and the distortion condition D1.*

Then, F has an SRB measure μ whose basin has full Lebesgue measure in Q. Moreover, the natural extension of the system (F,μ) is measure-theoretically isomorphic to a Bernoulli shift, F has finite entropy with respect to the measure μ, and we have the formula

$$h_\mu(F) = \lim_{n\to\infty} \frac{1}{n} \log |DF^n(z)| \tag{6}$$

where the latter limit exists for Lebesgue almost all z and is independent of such z.

Remark 3.2. — Formula (6) says that the entropy can actually be computed by taking the logarithmic growth rate of the norms of $DF^n(z)$ for almost all z. It is actually true that if v is any unit vector in the K^u_α cone in the tangent space to such a z, then

$$h_\mu(F) = \lim_{n\to\infty} \frac{1}{n} \log \left|DF^n(z)(v)\right| \tag{7}$$

This last expression can easily be implemented numerically.

Remark 3.3. — If we assume that the interiors of the strips S_i are disjoint, then (F, μ) itself is isomorphic to a Bernoulli shift, and the entropy formula

$$h_\mu(F) = \int log|D^u F| d\mu$$

holds where $D^u F(z)$ is the norm of the derivative of F in the unstable direction at z.

Acknowledgement. — We wish to thank Francois Ledrappier and Dan Rudolph for useful conversations during the preparation of this paper.

4. Some estimates of partial derivatives

We will need to use the Mean Value Theorem for various partial derivatives of the mappings f_i at points near the domain E_i. Since the E_i are not necessarily convex subsets of $\mathbf{R}^2$, it will be useful to have our maps f_i extended to neighborhoods $\mathcal{E}_i$ of E_i which contain $\bigcup_{z\in E_i} B_{C\delta(z)}(z)$ where C is a fixed positive constant and $B_{C\delta(z)}(z)$ denotes the ball about z of radius $C\delta(z)$. Using the proof of the Whitney extension theorem in **[1]** it is possible to show that there is an extension $\widetilde{f_i}$ of f_i to such a neighborhood which satisfies the same properties H1-H4, D1, with possibly different constants. We will assume henceforth that our maps f_i have such extensions.

We collect here some estimates which follow from our assumptions.

Let $f(x, y) = (f_1(x.y), f_2(x, y))$ be one of our maps f_i on E_i.

Lemma 4.1. — *For $z \in E_i$, we have the estimates*

$$\frac{|f_{1y}(z)|}{|f_{1x}(z)|} \le \alpha \tag{8}$$

$$\frac{|f_{2x}(z)|}{|f_{1x}(z)|} \le \alpha \tag{9}$$

$$\frac{|f_{2y}(z)|}{|f_{1x}(z)|} \le \frac{1}{K_0^2} + \alpha^2 \tag{10}$$

Proof. — We have

$$Df_z = \begin{pmatrix} f_{1x} & f_{1y} \\ f_{2x} & f_{2y} \end{pmatrix}$$

and

$$Df_{fz}^{-1} = \frac{1}{J_z}\begin{pmatrix} f_{2y} & -f_{1y} \\ -f_{2x} & f_{1x} \end{pmatrix}$$

where $J_z = f_{1x}f_{2y} - f_{2x}f_{1y}$.

Using $Df_z\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) \in K_\alpha^u$ and $Df_{fz}^{-1}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right) \in K_\alpha^s$ immediately gives

$$\frac{|f_{2x}|}{|f_{1x}|} \leq \alpha, \quad \frac{|f_{1y}|}{|f_{1x}|} \leq \alpha.$$

Now, we know that $|Df_{fz}^{-1}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)| \geq K_0$ in the max norm, so $\frac{1}{|J_z|}\max(|f_{1y}|, |f_{1x}|) \geq K_0$.

Hence, either $|J_z|K_0 \leq |f_{1y}|$ or $|J_z|K_0 \leq |f_{1x}|$.

The first case gives

$$(|f_{1x}f_{2y}| - |f_{1y}f_{2x}|)K_0 \leq |f_{1y}|$$

or

$$\begin{aligned}\frac{|f_{2y}|}{|f_{1x}|} &\leq \frac{|f_{1y}|}{K_0|f_{1x}|^2} + \frac{|f_{1y}f_{2x}|}{|f_{1x}|^2} \\ &\leq \frac{\alpha}{K_0|f_{1x}|} + \alpha^2 \\ &\leq \frac{1}{K_0^2} + \alpha^2.\end{aligned}$$

Analogously, in the second case,

$$(|f_{1x}f_{2y}| - |f_{1y}f_{2x}|)K_0 \leq |f_{1x}|$$

or

$$\begin{aligned}\frac{|f_{2y}|}{|f_{1x}|} &\leq \frac{1}{K_0|f_{1x}|} + \alpha^2 \\ &\leq \frac{1}{K_0^2} + \alpha^2\end{aligned}$$

Thus, in any case, we have

$$\frac{|f_{2y}|}{|f_{1x}|} \leq \frac{1}{K_0^2} + \alpha^2.$$

□

We have assumed that our maps f_i have extensions to neighborhoods $\mathcal{E}_i$ of E_i with the following properties.

The map f_i takes $\mathcal{E}_i$ onto a set $\widetilde{S}_i \subset \mathbf{R}^2$ such that

(11) $$B_{C\delta_z(E_i)}(z) \subset \mathcal{E}_i \quad \text{for } z \in E_i$$

and

$$f_i \text{ satisfies H1–H4, D1 on } \mathcal{E}_i \tag{12}$$

Any C^1 curve $\gamma(t)$ such that $\gamma'(t) \in K_\alpha^u$ for all t will be called a K_α^u curve. Similarly, a K_α^s curve is a C^1 curve $\gamma(t)$ for which $\gamma'(t) \in K_\alpha^s$ for all t. In this paper, all of our K_α^u curves will actually be of class C^2, and this will be assumed without further mention.

Lemma 4.2. — *Let f_i have an extension to the neighborhood $\mathcal{E}_i$ as above. Then, there is a constant $C_1 > 0$ independent of i such that if z and w lie on a K_α^u curve in $\mathcal{E}_i$, then*

$$\frac{|f_{i1x}(z)|}{|f_{i1x}(w)|} \leq \exp\left(C_1 \frac{|z-w|}{\delta_z(E_i)}\right).$$

Proof. — Write $f = f_i$.

Since $|Df_z\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)| = \max(|f_{1x}(z)|, |f_{2x}(z)|) \geq K_0$ and $|f_{2x}(z)| \leq \alpha|f_{1x}(z)|$, we know that

$$|f_{1x}(z)| = \left|Df_z\begin{pmatrix}1\\0\end{pmatrix}\right| \geq K_0 > 1$$

so, for w near z, both $f_{1x}(z)$ and $f_{1x}(w)$ have the same sign. We assume this sign is positive (replace f by $-f$ otherwise).

Since f extends to the neighborhood $\mathcal{E}_i$, and, for some constant $C > 0$, this last set contains the balls of radius $C\delta_z(E_i) > 0$ about points z in E_i, the mean value theorem gives us that if $|z-w| \leq C\delta_z(E_i)$, then there is a τ on the line segment joining z and w such that

$$|\log f_{1x}(z) - \log f_{1x}(w)| \leq \left|\frac{f_{1xx}(\tau)}{f_{1x}(\tau)}\right||z-w| + \left|\frac{f_{1xy}(\tau)}{f_{1x}(\tau)}\right||z-w|\alpha$$

or

$$\frac{|f_{1x}(z)|}{|f_{1x}(w)|} \leq \exp\left((1+\alpha)C_0 \frac{|z-w|}{\delta_\tau(E_i)}\right)$$

using the distortion estimate D1.

Let $z = (x_0, y_0)$, let $z_r = (x_r, y_r)$ be the point of intersection of the horizontal line ℓ_z with the right boundary curve of E_i, and let $z_\ell = (x_\ell, y_\ell)$ be the point of intersection of the horizontal line ℓ_z with the left boundary curve of E_i. Since w lies on a K_α^u curve containing z, the line ℓ^0 through z and τ has equation $y - y_0 = \beta(x - x_0)$ for some β with $|\beta| \leq \alpha$. Also, since the right boundary curve of E_i through z_r is a K_α^s – *curve*, it is contained between the lines $\ell_r^- : x - x_r = -\alpha(y - y_r)$ and $\ell_r^+ : x - x_r = \alpha(y - y_r)$. Similar statements hold for the left boundary curve of E_i and the lines $\ell_\ell^- : x - x_\ell = -\alpha(y - y_\ell)$ and $\ell_\ell^+ : x - x_\ell = \alpha(y - y_\ell)$. Using the intersections of the lines $\ell^0, \ell_r^\pm, \ell_\ell^\pm$, an elementary argument gives that

$$\frac{1}{1+\alpha^2} \leq \frac{1}{1+|\beta|\alpha} \leq \frac{\delta_\tau(E_i)}{\delta_z(E_i)} \leq \frac{1}{1-|\beta|\alpha} \leq \frac{1}{1-\alpha^2}.$$

This gives the desired estimate for z, w with $|z - w| \leq C\delta_z(E_i)$.

To get the general estimate of the Lemma, we simply find a sequence $z_0 = z$, $z_1, \ldots z_j = w$ with $z_k \in E_i$, $|z_k - z_{k+1}| < C\delta_z(E_i)$, each z_k on the same K_α^u curve, and j dependent only on α, C, and $\delta_z(E_i)$. Using the estimate for each pair z_i, z_{i+1} then easily gives us the general estimate to complete the proof of the Lemma. □

In some of our arguments below, it will simplify matters if we can take the constant K_0 in (3) and (5) to be large. The next lemma shows that this can be arranged by replacing F by a fixed finite power F^t with $t > 0$.

Lemma 4.3. — *Suppose the maps f_i satisfy (2), (3), (4), (5), and D1 on the neighborhoods*

$$\bigcup_{z \in E_i} B_{C\delta_z(E_i)}(z),$$

and let $t > 0$ be a positive integer.

Then there are positive constants $C_0 = C_0(t)$, $C_2 = C_2(t)$ such that the maps $f_{i_{t-1}} \circ \cdots \circ f_{i_0}$ satisfy (2), (3), (4), and (5) with K_0 replaced by K_0^t and D1 with C_0 replaced by $C_0(t)$ on the neighborhoods

$$\bigcup_{z \in E_{i_0 \ldots i_{t-1}}} B_{C_2(t)\delta_z(E_{i_0 \ldots i_{t-1}})}(z)$$

Proof. — The proof is by induction on the number of elements in the composition. We assume that it holds for compositions of length t and prove it for those of length $t+1$.

Let $B_{C(i_0 \ldots i_t)}$ denote the set

$$\bigcup_{z \in E_{i_0 \ldots i_t}} B_{C_2(t+1)\delta_z(E_{i_0 \ldots i_t})}(z).$$

From Lemma 4.2, we can choose a constant $C_2(t+1) \in (0, C_2(t)) \subset (0,1)$ so that if $w \in B_{C(i_0 \ldots i_t)}$, then $f_{i_0}(w) \in B_{C(i_1 \ldots i_t)}$.

It is clear that the maps $f_{i_t} \circ \cdots \circ f_{i_0}$ satisfy (2), (3), (4), and (5) with K_0 replaced by K_0^{t+1}, so we only need to be concerned with the statement regarding D1.

If E is a subset of Q, $z \in E$, and $f(x,y) = (f_1(x,y), f_2(x,y))$, we set

$$\Theta_z(f, E) = \max_{i=1,2} \frac{|D^2 f_i(z)|}{|f_{1x}(z)|} \delta_z(E)$$

Let $f = f_{i_t} \circ \cdots \circ f_{i_1}$, $g = f_{i_0}$, $h = f \circ g$, $E_f = E_{i_1 \ldots i_t}$, $E_g = E_{i_0}$, $E_h = E_{i_0 \ldots i_t}$, and, for $z \in E_h$, write $\Delta f = \delta_{gz}(E_f)$, $\Delta g = \delta_z(E_g)$, and $\Delta h = \delta_z(E_h)$. Also, write $\Theta(f) = \Theta_{gz}(f, E_f)$, $\Theta(g) = \Theta_z(g, E_g)$, $\Theta(h) = \Theta_z(h, E_h)$.

Let us first estimate the quotient

$$\frac{|g_{1x}(w)|}{|g_{1x}(z)|}$$

for any $w \in \ell_z \cap \mathcal{E}_{i_0}$.

Note that $g_{1x}(w)$ and $g_{1x}(z)$ have the same sign. We assume it is positive. The argument when it is negative is similar.

Letting C_1 be the constant in Lemma 4.2, if $w, \overline{w} \in \ell_z \cap \mathcal{E}_{i_0}$, we have

$$\frac{|g_{1x}(w)|}{|g_{1x}(\overline{w})|} \leq \exp(2C_1) \tag{13}$$

We can connect w to z in $\ell_z \cap \mathcal{E}_{i_0}$ by a chain of points $w = w_0, w_1, \ldots w_k = z$ where $|w_i - w_{i+1}| < C_2(1)\delta_z(E_g)$, and $k \leq 3/C_2(1)$.

Hence, putting $\zeta = \exp(6C_1/C_2(1))$, we have

$$\frac{|g_{1x}(w)|}{|g_{1x}(z)|} \leq \prod_{0 \leq i < k} \frac{|g_{1x}(w_i)|}{|g_{1x}(w_{i+1})|} \leq \zeta \tag{14}$$

Interchanging z, w in the above argument gives $|g_{1x}(w)|/|g_{1x}(z)| \geq \zeta^{-1}$. From these two inequalities we get, for any $w, \tau \in \ell_z \cap \mathcal{E}_{i_0}$,

$$\begin{aligned} \zeta^{-2} &\leq \frac{|g_{1x}(\tau)|}{|g_{1x}(z)|} \frac{|g_{1x}(z)|}{|g_{1x}(w)|} \\ &= \frac{|g_{1x}(\tau)|}{|g_{1x}(w)|} \\ &\leq \zeta^2 \end{aligned}$$

By the Chain Rule for partial derivatives we have the following formulas for $i = 1, 2$

$$h_{ix} = f_{ix} g_{1x} + f_{iy} g_{2x}; \quad h_{iy} = f_{ix} g_{1y} + f_{iy} g_{2y} \tag{15}$$

$$h_{ixx} = f_{ixx} g_{1x}^2 + f_{ixy} g_{2x} g_{1x} + f_{iyx} g_{1x} g_{2x} + f_{iyy} g_{2x}^2 + f_{ix} g_{1xx} + f_{iy} g_{2xx} \tag{16}$$

$$h_{ixy} = f_{ixx} g_{1y} g_{1x} + f_{ixy} g_{2y} g_{1x} + f_{iyx} g_{1y} g_{2x} + f_{iyy} g_{2y} g_{2x} + f_{ix} g_{1xy} + f_{iy} g_{2xy} \tag{17}$$

$$h_{iyy} = f_{ixx} g_{1y}^2 + f_{ixy} g_{2y} g_{1y} + + f_{iyx} g_{1y} g_{2y} + f_{iyy} g_{2y}^2 + f_{ix} g_{1yy} + f_{iy} g_{2yy} \tag{18}$$

Let $w \in B_{C(i_0 \ldots i_t)}$. Except where otherwise mentioned, we compute the partial derivatives below at w.

From (15) and Lemma 4.1, we get

$$|h_{1x}| \geq |f_{1x} g_{1x}|(1 - \alpha^2) \tag{19}$$

From (16), we have, for $i = 1,2$,

$$\left|\frac{h_{i_{xx}}(w)}{h_{1_x}(w)}\Delta h\right| \leq (1-\alpha^2)^{-1}\left[\Theta(f)|g_{1_x}(w)|\frac{\Delta h}{\Delta f} + 2\Theta(f)|g_{2_x}(w)|\frac{\Delta h}{\Delta f}\right.$$
$$\left. + \Theta(f)|g_{2_x}(w)|\frac{\Delta h}{\Delta f} + 3\Theta(g)\frac{\Delta h}{\Delta g}\right]$$

Since the g-image of a horizontal line is a K^u_α curve and the boundaries of E_f are K^s_α curves, we can use the mean value theorem and a simple geometric estimate to get a constant $C_3(\alpha) > 0$ such that

$$|g_{1x}(\tau)|\Delta h \leq C_3(\alpha)\Delta f$$

for some point τ in $\ell_z \cap E_h$.

Putting all these estimates together gives

$$\left|\frac{h_{i_{xx}}(w)}{h_{1_x}(w)}\Delta h\right| \leq C_4(\alpha)(\Theta(f)\zeta^2 + \Theta(g))$$

Similar estimates can be given for the quantities

$$\left|\frac{h_{i_{xy}}(w)}{h_{1_x}(w)}\Delta h\right|, \quad \left|\frac{h_{i_{yy}}(w)}{h_{1_x}(w)}\Delta h\right|.$$

Thus, we simply define $C_0(t+1)$ so that it is larger than $C_4(\alpha)(C_0(t)\zeta^2 + C_0)$ and we have proved Lemma 4.3. □

5. Families of Fiber contractions

Fiber contraction maps were defined in [7] to provide a tool in the analysis of smoothness of stable and unstable manifolds. We collect here certain facts about parametrized families of fiber contraction maps and related concepts.

Let $(X, d_1), (Y, d_2)$ be complete metric spaces and give $X \times Y$ the metric

$$d((x,y),(x',y')) = \max(d_1(x,x'), d_2(y,y')).$$

Let $\pi_1 : X \times Y \to X, \pi_2 : X \times Y \to Y$ be the natural projections.

A pair of maps (F, f) is called a fiber contraction on $X \times Y$ if the following properties hold.

(1) $f : X \to X$ and $F : X \times Y \to X \times Y$ are continuous maps.
(2) $\pi_1 F = f\pi_1$.
(3) There is a constant $0 < K < 1$ such that for $x \in X, y, y' \in Y$, we have

$$d(F(x,y), F(x,y')) \leq K d_2(y,y').$$

We call f the *base map* and F the *total map* of the fiber contraction (F, f).

Let f be a continuous self-map of the complete metric space X. We say the a point $x_0 \in X$ is an attracting fixed point of f if for every $x \in X$, the sequence of iterates

$x, f(x), f^2(x), \ldots$ converges to x_0 as $n \to \infty$. Clearly if such an x_0 exists, it must be the unique fixed point of f.

Let A be a topological space and consider a family $\{f_\lambda\}_{\lambda \in A}$ of self-maps of the complete metric space X. We say that the family is continuous if the map $(\lambda, x) \to f_\lambda(x)$ from $A \times X$ to X is continuous.

A family $\{f_\lambda\}$ of self-maps of X is called a *uniform family of contractions* if

(1) there is a constant $0 < K < 1$ such that, for all λ, x, x',

$$d(f_\lambda x, f_\lambda x') \leq K d(x, x').$$

(2) the family $\{f_\lambda\}$ is continuous.

We say that a family $\{(F_\lambda, f_\lambda)\}$ of fiber contractions is a *uniform family of fiber contractions* if

(1) the fiber Lipschitz constants are uniformly less than 1. That is, there is a constant $0 < K < 1$ such that for any λ, x, y, y'

$$d(F_\lambda(x, y), F_\lambda(x, y')) \leq K d_2(y, y')$$

(2) the families $\{F_\lambda\}$ and $\{f_\lambda\}$ are continuous.

The following Proposition is standard (see *e.g.* **[6]**) and its proof will be omitted.

Proposition 5.1. — *If $\{f_\lambda\}$ is a uniform family of contractions of the complete metric space X, and x_λ is the fixed point of f_λ, then the family $\{x_\lambda\}$ depends continuously on λ.*

Proposition 5.2. — *Suppose $\{(F_\lambda, f_\lambda)\}$ is a uniform family of fiber contractions whose base maps $\{f_\lambda\}$ have attracting fixed points $\{x_\lambda\}$ depending continuously on λ. Then, each of the maps F_λ has an attracting fixed point of the form $(x_\lambda, y_\lambda) \in X \times Y$ and the family $\{(x_\lambda, y_\lambda)\}$ depends continuously on λ.*

Proof. — Letting x_λ be the fixed point of the base map f_λ, Hirsch and Pugh prove in **[7]** that F_λ has an attracting fixed point of the form (x_λ, y_λ) where y_λ is the fixed point of the map $F(x_\lambda, \cdot)$ on Y. Since x_λ depends continuously on λ, the family $\{F(x_\lambda, \cdot)\}$ is family of uniform contractions on Y. Therefore, by Proposition 5.1, the fixed points $\{y_\lambda\}$ depend continuously on λ. □

The following corollary is proved by induction using Propositions 5.1 and 5.2.

Corollary 5.3. — *Suppose $X_1 \times X_2 \times \cdots \times X_N$ is a sequence of complete metric spaces and $\{F_{\lambda,i}\}, 1 \leq i \leq N$ is a sequence of maps with the following properties.*

(1) *$\{F_{\lambda,1}\}$ is a uniform family of contractions on X_1.*

(2) *For $2 \leq i \leq N$, $\{F_{\lambda,i}, F_{\lambda,i-1}\}$ is a uniform family of fiber contractions on $\prod_{1 \leq j \leq i} X_j$.*

Then, each of the families $\{F_{\lambda,i}\}$ has an attracting family of fixed points $\{x_{\lambda,i}\}$ which depends continuously on λ.

6. Invariant Manifolds

We consider the collection $\xi = \{E_1, E_2, \dots\}$ of rectangles as above and the sequence $(f_1, f_2, \dots)$ of C^2 diffeomorphisms with $f_i(E_i) = S_i$ satisfying G1–G3, H1–H4, and D1. From Proposition 2.1, using the max norm on $\mathbf{R}^2$, we have, for each i,

$$Df_i(K^u_\alpha) \subseteq K^u_\alpha \tag{20}$$

$$v \in K^u_\alpha \Rightarrow \left|Df_i v\right| \geq K_0|v| \tag{21}$$

$$Df_i^{-1}(K^s_\alpha) \subseteq K^s_\alpha \tag{22}$$

$$v \in K^s_\alpha \Rightarrow \left|Df_i^{-1} v\right| \geq K_0|v| \tag{23}$$

For each finite sequence $i_{-n+1} \dots i_0 \dots i_{n-1}$ we have defined, in Section 2, the sets $E_{i_0 \dots i_{n-1}}, S_{i_{-n+1} \dots i_0}$.

Given a non-positive itinerary $\boldsymbol{i} = (\dots i_{-n} i_{-n+1} \dots i_0)$, we consider the set $W^u_{\boldsymbol{i}} = E_{i_0} \cap \bigcap_{n \geq 0} S_{i_{-n} \dots i_{-1}}$. Clearly, $W^u_{\boldsymbol{i}}$ is a closed, connected full-width subset of E_{i_0}. Its image $FW^u_{\boldsymbol{i}} = f_{i_0} W^u_{\boldsymbol{i}}$ is the set $\bigcap_{n \geq 0} S_{i_{-n} \dots i_0}$, a full-width connected subset of Q. The next result shows that $FW^u_{\boldsymbol{i}}$ is a C^2 curve which depends continuously on $\boldsymbol{i}$.

For convenience, we let $D^0 \psi = \psi$ for a function ψ.

Theorem 6.1. — *There is a constant $K > 0$ such that for each non-positive itinerary $\boldsymbol{i} = (\dots i_{-n} \dots i_0)$, the set $FW^u_{\boldsymbol{i}}$ is the graph of a C^2 function $g_{\boldsymbol{i}} : I \to I$ such that, for $z \in I$,*

$$\left|Dg_{\boldsymbol{i}}(z)\right| \leq \alpha \tag{24}$$

and

$$\left|D^2 g_{\boldsymbol{i}}(z)\right| \leq K. \tag{25}$$

Further, given $\varepsilon > 0$, there is a positive integer $N > 0$ such that if $\boldsymbol{i} = (\dots i_{-n} \dots i_0)$ and $\boldsymbol{j} = (\dots j_{-n} \dots j_0)$ are non-positive itineraries with $i_{-\ell} = j_{-\ell}$ for $0 \leq \ell \leq N$, then

$$\left|D^k g_{\boldsymbol{i}}(z) - D^k g_{\boldsymbol{j}}(z)\right| < \varepsilon \tag{26}$$

for $z \in I$ and $0 \leq k \leq 2$.

Remark 6.2. — The proof of Theorem 6.1 uses graph transform techniques as in **[7]**, **[12]**. However, since our maps have unbounded derivatives, and the off-diagonal terms of our derivatives are not small, certain modifications of the techniques in **[7]**, **[12]** are necessary.

It can be shown that if f_i is C^r for $r \geq 2$, then the curves $W^u_{\boldsymbol{i}}$ are C^r and depend continuously on $\boldsymbol{i}$ in the C^r sense provided the f_i satisfy the r-th order distortion condition

$$\sup_{\substack{z\in\mathcal{E}_i,\, i\geq 1\\ 2\leq k\leq r}} \frac{|D^k f_i(z)|}{|f_{i1x}(z)|^{k-1}}\delta_z(\mathcal{E}_i) < C_0$$

where $D^k f_i(z)$ is the supremum of the ℓ-th order partial derivatives of f_i at z for $\ell \leq k$.

We proceed toward the proof of Theorem 6.1.

Notice that if we replace $F = \{f_i\}$ by a positive power $F^t, t > 0$, and ξ by the collection $\{E_{i_0\dots i_{t-1}}\}$, we may assume that K_0 is as large as we wish in (21), (23). In the present section we will take $K_0 > 4$. Of course this changes the distortion constant C_0 in D1 to some $C_1 = C_1(f,t)$ but this will not cause us difficulties.

Let $\mathbf{N}$ be the set of positive integers, and let $\Sigma = \mathbf{N}^{\mathbf{Z}}$ be the space of doubly infinite sequences $\boldsymbol{i} = (\dots i_{-1} i_0 i_1 \dots)$ of elements of $\mathbf{N}$ with the product topology. Let $\sigma : \Sigma \to \Sigma$ be the usual left shift automorphism.

For an element $\boldsymbol{i} \in \Sigma$, let $\boldsymbol{i}^+ = (i_0 i_1 \dots)$ be its non-negative part, and let $\boldsymbol{i}^- = (\dots i_{-1} i_0)$ be its non-positive part. Set $W^s_{\boldsymbol{i}^+} = \bigcap_{n\geq 0} E_{i_0\dots i_n}$ and $W^u_{\boldsymbol{i}^-} = E_{i_0} \cap \bigcap_{n\geq 0} S_{i_{-n}\dots i_{-1}}$.

It follows from (20)–(23) that the sets $W^s_{\boldsymbol{i}^+}, W^u_{\boldsymbol{i}^-}$ intersect in a unique point and there is a continuous map $\pi : \Sigma \to Q$ defined by

$$\{\pi((\dots i_{-1} i_0 i_1 \dots))\} = W^s_{\boldsymbol{i}^+} \cap W^u_{\boldsymbol{i}^-}$$

Moreover, for each $\boldsymbol{i} \in \Sigma$ there is a splitting $T_{\pi(\boldsymbol{i})}\mathbf{R}^2 = E^u_{\pi(\boldsymbol{i})} \oplus E^s_{\pi(\boldsymbol{i})}$ which depends continuously on $\boldsymbol{i}$ and is such that Df_{i_0} maps $E^u_{\pi(\boldsymbol{i})}$ to $E^u_{\pi(\sigma\boldsymbol{i})}$ and $E^s_{\pi(\boldsymbol{i})}$ to $E^s_{\pi(\sigma\boldsymbol{i})}$. The arguments for these facts are analogous to standard arguments in hyperbolic theory (*e.g.*, to prove that C^1 perturbations of the Smale horseshoe diffeomorphism have a hyperbolic non-wandering set) and will not be given here.

Thus, the matrix of DF is diagonal with respect to the splitting $E^u \oplus E^s$ on the image of π.

For $z = \pi(\boldsymbol{i})$ and $v \in T_z\mathbf{R}^2$, we write $v = (v_1, v_2) \in E^u_z \oplus E^s_z$ and define $|v| = |v|_z = \max(|v_1|, |v_2|)$. This norm depends continuously on $\boldsymbol{i} \in \Sigma$.

We will identify all tangent spaces with the space $\mathbf{R}^2$ itself by standard translations.

It will be convenient to use the subundles E^u, E^s to define affine local coordinates near points $z, f_i z$ in which Df_{iz} becomes diagonal and in which Df_{iw} is nearly diagonal for $|w - z|$ no larger than a fixed multiple of $\delta_z(E_i)$. Here $i = i_0$ with $z = \pi(\boldsymbol{i})$.

Toward this end, let A_z be the affine automorphism of $\mathbf{R}^2$ such that

(1) $A_z(z) = z$.
(2) $DA_z\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) = \left(\begin{smallmatrix}1\\a_z\end{smallmatrix}\right) \in E^u_z$.
(3) $DA_z\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right) = \left(\begin{smallmatrix}b_z\\1\end{smallmatrix}\right) \in E^s_z$.

Since $E^u_z \subseteq K^u_\alpha$ and $E^s_z \subseteq K^s_\alpha$, we have $|a_z| \leq \alpha$, $|b_z| \leq \alpha$.

Let $\widetilde{f} = \widetilde{f}_i = A_{\widetilde{f}_i z}^{-1} f_i A_z$ be the local representative of f_i using the $A_{f_i z}, A_z$ coordinates. Note that $\widetilde{f}$ is defined on the affine image $A_z^{-1}(\mathcal{E}_i)$ of $\mathcal{E}_i$.

Then the matrix $D\widetilde{f}_z$ is diagonal. For w near z in $A_z(\mathcal{E}_i)$, let

$$D\widetilde{f}_w = \begin{pmatrix} \widetilde{f}_{1x}(w) & \widetilde{f}_{1y}(w) \\ \widetilde{f}_{2x}(w) & \widetilde{f}_{2y}(w) \end{pmatrix}$$

and set

$$\varepsilon_{12}(w) = \frac{|\widetilde{f}_{1y}(w)|}{|\widetilde{f}_{1x}(w)|}, \quad \varepsilon_{21}(w) = \frac{|\widetilde{f}_{2x}(w)|}{|\widetilde{f}_{1x}(w)|}, \quad \varepsilon_{22}(w) = \frac{|\widetilde{f}_{2y}(w)|}{|\widetilde{f}_{1x}(w)|}.$$

We wish to estimate $\varepsilon_{ij}(w)$ for w near z in $\mathcal{E}_i$. It follows from the definitions that $\varepsilon_{ij}(z) = 0, i \neq j$. Also, (21) and (23) imply $\varepsilon_{22}(z) \leq 1/K_0^2 < 1/16$ since $K_0 > 4$.

Lemma 6.3. — *There are constants $C_2 \in (0,1), C_3 > 0, C_4 > 0$, such that for $z \in E_i, w \in A_z^{-1}(\mathcal{E}_i)$, if $|w - z| < C_2\delta_z(E_i)$, then*

$$|\varepsilon_{ij}(w) - \varepsilon_{ij}(z)| \leq C_3 \frac{|z-w|}{\delta_z(E_i)}, \tag{27}$$

and

$$C_4 \frac{1}{\delta_z(E_i)} \leq |\widetilde{f}_{1x}(w)| \leq C_4^{-1} \frac{1}{\delta_z(E_i)} \tag{28}$$

Proof. — To begin with, let us choose $C_2 \in (0,1)$ so that if $|w - z| < C_2\delta_z(E_i)$, then $w \in \mathcal{E}_i \cap A_z^{-1}(\mathcal{E}_i)$ and $\widetilde{f}$ satisfies D1 for some (possibly different) constant C_0. Since $A_{\widetilde{f}_i z}^{-1}$ and A_z are uniformly bounded, it is possible to choose C_0 and C_2 independent of $z \in E_i$ and $i \geq 1$.

We next show that there are constants $C_5 > 0, C_6 > 0$ such that for $z \in E_i$ and $|z - w| < C_5\delta_z(E_i)$,

$$C_6^{-1} \leq \frac{|f_{1x}(w)|}{|\widetilde{f}_{1x}(w)|} \leq C_6. \tag{29}$$

Since

$$\begin{aligned} D\widetilde{f}_w \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= \begin{pmatrix} \widetilde{f}_{1x}(w) \\ \widetilde{f}_{2x}(w) \end{pmatrix} \\ &= DA_{fz}^{-1} Df_{A_z w} DA_z \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{J_{fz}} \begin{pmatrix} 1 & -b_{fz} \\ -a_{fz} & 1 \end{pmatrix} \begin{pmatrix} f_{1x} & f_{1y} \\ f_{2x} & f_{2y} \end{pmatrix} \begin{pmatrix} 1 & b_z \\ a_z & 1 \end{pmatrix} \end{aligned}$$

where $J_{fz} = 1 - a_{fz}b_{fz}$ and the partial derivatives of f_1, f_2 are evaluated at $A_z(w)$, we have

$$\begin{aligned}
J_{fz}\widetilde{f}_{1x}(w) &= f_{1x} + f_{1y}a_z - b_{fz}f_{2x} - b_{fz}f_{2y}a_z \\
J_{fz}\widetilde{f}_{1y}(w) &= f_{1x}b_z + f_{1y} - b_{fz}f_{2x}b_z - b_{fz}f_{2y} \\
J_{fz}\widetilde{f}_{2x}(w) &= -a_{fz}f_{1x} - a_{fz}f_{1y}a_z + f_{2x} + f_{2y}a_z \\
J_{fz}\widetilde{f}_{2y}(w) &= -a_{fz}f_{1x}b_z - a_{fz}f_{1y} + f_{2x}b_z + f_{2y}
\end{aligned}$$

Using the first equation above, the fact that $|J_{fz}| \geq 1 - \alpha^2$, and the estimates (8), (9), (10) at $A_z(w)$ we get

$$|\widetilde{f}_{1x}(w)| \leq C|f_{1x}(A_z(w))|$$

for some constant C. But, from Lemma 4.2 we have $|f_{1x}(A_z(w))|$ is bounded above by const $|f_{1x}(z)|$, so this gives the lower bound in (29).

For the upper bound, we will obtain the two estimates

$$|\widetilde{f}_{1x}(w)| \geq C\left|D\widetilde{f}_w \begin{pmatrix}1\\0\end{pmatrix}\right| \tag{30}$$

and

$$\left|D\widetilde{f}_w \begin{pmatrix}1\\0\end{pmatrix}\right| \geq C|f_{1x}(w)| \tag{31}$$

for some constant $C > 0$.

To prove (30), we note first note that the vector $v = \begin{pmatrix}1 & b_z\\ a_z & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}$ is in the cone K^u_α. Since Df_{A_z} preserves this cone, we have that $Df_{A_z}(v)$ is a constant multiple of the vector $\begin{pmatrix}1\\ \overline{a}\end{pmatrix}$ for some $\overline{a}$ with $|\overline{a}| \leq \alpha$.

Thus,

$$D\widetilde{f}_w \begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\widetilde{f}_{1x}(w)\\ \widetilde{f}_{2x}(w)\end{pmatrix}$$

is a constant multiple of the vector

$$\begin{pmatrix}1 & -b_{fz}\\ -a_{fz} & 1\end{pmatrix}\begin{pmatrix}1\\ \overline{a}\end{pmatrix} = \begin{pmatrix}1 - b_{fz}\overline{a}\\ -a_{fz} + \overline{a}\end{pmatrix}$$

This gives

$$\frac{|\widetilde{f}_{2x}(w)|}{|\widetilde{f}_{1x}(w)|} \leq \frac{2\alpha}{1-\alpha^2}$$

and, hence,

$$\left|\begin{pmatrix}\widetilde{f}_{1x}(w)\\ \widetilde{f}_{2x}(w)\end{pmatrix}\right| = \max\left(|\widetilde{f}_{1x}(w)|, |\widetilde{f}_{2x}(w)|\right)$$
$$\leq \max\left(1, \frac{2\alpha}{1-\alpha^2}\right)|\widetilde{f}_{1x}(w)|$$

and (30) follows.

Next we go to the proof of (31).

Since the matrix $\begin{pmatrix}1 & b_{fz}\\ a_{fz} & 1\end{pmatrix}$ and its inverse are uniformly bounded as are J_{fz} and its inverse, we have

$$\left|D\widetilde{f}_w\begin{pmatrix}1\\0\end{pmatrix}\right| \geq C\left|Df_{A_z(w)}\begin{pmatrix}1\\a_z\end{pmatrix}\right|$$

But,

$$Df_{A_z(w)}\begin{pmatrix}1\\a_z\end{pmatrix} = \begin{pmatrix}A_1\\A_2\end{pmatrix}$$

where $A_1 = f_{1x} + a_z f_{1y}$. So,

$$\left|Df_{A_z(w)}\begin{pmatrix}1\\a_z\end{pmatrix}\right| \geq |f_{1x}| - \alpha^2|f_{1x}|$$
$$\geq (1-\alpha^2)|f_{1x}|$$
$$\geq (1-\alpha^2)|f_{1x}(A_z^{-1}w)|$$
$$\geq C|f_{1x}(w)|$$

This completes the proof of (29).

Next, we give the proof of the estimate

$$|\varepsilon_{12}(w) - \varepsilon_{12}(z)| \leq C_3\frac{|z-w|}{\delta_z(E_i)} \tag{32}$$

The other estimates for (27) are similar.

Since $\widetilde{f}_{1y}(z) = 0$, we need to estimate $|\widetilde{f}_{1y}(w)/\widetilde{f}_{1x}(w)|$.

But,

$$J_{fz}\widetilde{f}_{1y}(w) = f_{1x}(A_z w)b_z + f_{1y}(A_z w) - b_{fz}f_{2x}(A_z w)b_z - b_{fz}f_{2y}(A_z w),$$

so,

$$|\widetilde{f}_{1y}(w)| \leq C\max_{i,j,k,\tau}|f_{ijk}(\tau)||A_z(w) - z|.$$

Now, we know that the quantities $\delta_z(E_i)/\delta_\tau(E_i)$, $|f_{1x}(w)|/|\widetilde{f}_{1x}(w)|$ are bounded above and below, and, by Lemma (4.2), the same holds for $|f_{1x}(\tau)|/|\widetilde{f}_{1x}(w)|$. This gives (32) and (27).

For (28), notice that $f(\ell_z \cap E_i)$ is a full-width K^u_α curve in Q. In the max metric, it has unit length. By the Mean Value Theorem there is a $\tau \in \ell_z \cap E_i$ such that

$$\left| Df(\tau) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right| \delta_z(E_i) = 1$$

But,

$$\begin{aligned} \left| Df(\tau) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right| &= \max\left(\left| f_{1x}(\tau) \right|, \left| f_{2x}(\tau) \right| \right) \\ &= \left| f_{1x}(\tau) \right| \end{aligned}$$

so,

$$\left| f_{1x}(\tau) \right| = \frac{1}{\delta_z(E_i)}.$$

Since, $\left| f_{1x}(\tau) \right| / \left| \widetilde{f}_{1x}(w) \right|$ is bounded above and below, (28) follows.

This completes the proof of Lemma 6.3. □

For $\varepsilon > 0$, let $B_\varepsilon(z) = \{w \in \mathbf{R}^2 : \left| w - z \right|_z \leq \varepsilon\}$. Here $\left| w - z \right|_z$ refers to the max norm in the image of the affine coordinate map A_z. The set $B_\varepsilon(z)$ is then a parallelogram centered at z with sides parallel to E^u_z, E^s_z. Write $B_\varepsilon(z) = B^u_\varepsilon(z) \times B^s_\varepsilon(z)$ where $B^u_\varepsilon(z)$ is a line segment centered at z parallel to E^u_z, and $B^s_\varepsilon(z)$ is a line segment centered at z parallel to E^s_z. A *full-width curve of slope less than 1* in $B_\varepsilon(z)$ is the graph of a function $\phi : B^u_\varepsilon(z) \to B^s_\varepsilon(z)$ in which ϕ is Lipschitz with Lipschitz constant less than 1.

With $z = \pi(\boldsymbol{i})$, let $z_0 = z, z_j = \pi(\sigma^j \boldsymbol{i})$ for $j \leq 0$.

Our next goal, as is usual in invariant manifold theory, is to find a sequence of numbers $\varepsilon_j > 0$ such that the neighborhoods $B_j = B_{\varepsilon_j}(z_j)$ have the following properties.

B1 If $z_j \in E_{i_j}$, then $B_j \subseteq \mathcal{E}_{i_j}$.

B2 $\widetilde{f}_{i_j}(B_j)$ overflows B_{j+1} in the sense that if γ is a full-width curve of slope less than 1 in B_j passing through z_j, then $\widetilde{f}_{i_j}(\gamma) \cap B_{j+1}$ is a full-width curve of slope less than 1 in B_{j+1} passing through z_{j+1}.

Let $\overline{\varepsilon}_j = C_2 \delta_{z_j}(E_{i_j})$ where C_2 is the constant of Lemma 6.3. By Lemma 6.3, for $\left| w - z_j \right| < \overline{\varepsilon}_j$, the matrix of $D\widetilde{f}_{i_j}(w)$ is hyperbolic with off-diagonal terms small compared to $\left| \widetilde{f}_{1x}(w) \right|$. This implies that the image $\widetilde{f}_{i_j}\gamma$ of a curve γ as above will have slope less than 1 in $B_{\overline{\varepsilon}_{j+1}}(z_{j+1})$. Letting $f = \widetilde{f}_{i_j}$, and using $a \sim b$ to mean a/b is bounded above and below, we have $length(f\gamma) \sim \left| \widetilde{f}_{1x}(\tau) \right| C_2 \delta_{z_j}(E_{i_j})$ for some $\tau \in \gamma$. In the proof of Lemma 6.3 we saw that $\delta_{z_j}(E_{i_j}) \sim 1/\left| f_{1x}(\tau_1) \right| \sim 1/\left| \widetilde{f}_{1x}(\tau_1) \right|$ and $\left| \widetilde{f}_{1x}(\tau) \right| \sim \left| \widetilde{f}_{1x}(z_j) \right| \sim \left| \widetilde{f}_{1x}(\tau_1) \right|$. It follows that $\widetilde{f}_{i_j}\gamma$ contains a neighborhood of fixed size C_7 about $\widetilde{f}_{i_j}(z_j)$ in $\widetilde{f}_{i_j}\gamma$.

Let

$$\varepsilon_j = \begin{cases} \overline{\varepsilon}_j & \text{if } \overline{\varepsilon}_j < C_7 \\ C_7 & \text{if } \overline{\varepsilon}_j > C_7. \end{cases}$$

Then, the overflowing property above is satisfied.

Now fix a non-positive itinerary $\boldsymbol{i} = (\ldots i_{-n} \ldots i_0)$. We first show that $W_{\boldsymbol{i}}^u$ contains the graph of a C^2 function $g_{\boldsymbol{i}} : B_0^u \to B_0^s$ such that, for all $w \in B_0^u$

$$|Dg_{\boldsymbol{i}}(w)| \leq 1 \tag{33}$$

and

$$|D^2 g_{\boldsymbol{i}}(w)| \leq K_2 \tag{34}$$

where K_2 is independent of $\boldsymbol{i}$.

We also will show that the functions $g_{\boldsymbol{i}}$ depend continuously on $\boldsymbol{i}$.

Once these things are done, the proof of Theorem 6.1 is completed as follows.

Let $\boldsymbol{j} = (\ldots i_{-1} i_0 j_1 j_2 \ldots)$ be a doubly infinite itinerary which agrees with $\boldsymbol{i}$ for non-positive indices. Let $z_0 = \pi(\boldsymbol{j})$. Then there is a $k > 0$ independent of $\boldsymbol{i}, \boldsymbol{j}$ such that $f_{i_{-k}}^{-1} \circ \cdots \circ f_{i_{-1}}^{-1}(W_{\boldsymbol{i}}^u) \subseteq B_{-k}$. Note that here we use the original maps f_{i_j}, not the affine representatives $\widetilde{f}_{i_j}$.

Thus $W_{\boldsymbol{i}}^u$ is the $f_{i_{-1}} \circ \cdots \circ f_{i_{-k}}$-image of a curve of bounded slope and bounded C^2 size. Letting $F^k = f_{i_{-1}} \circ \cdots \circ f_{i_{-k}}$ we have that $W_{\boldsymbol{i}}^u$ is the graph of a function $\Gamma(F^k, g)$ where F^k has bounded distortion and g has bounded C^1, C^2 sizes. Using the formulas (36), (37), and (38) which appear in the second derivative of the graph transform function then gives that $\Gamma(F^k, g)$ also has bounded C^2 size. The same argument then works for $\Gamma(F^{k+1}, g)$ and this gives (25). A similar argument gives the continuity statement in Theorem 6.1.

To get estimate (24) first note that hyperbolicity conditions imply that any vector v in the tangent space to a point in $W_{\boldsymbol{i}}^u$ which is not in K_α^u has its backwards iterates eventually in K_α^s and, hence, eventually expanded. Since the tangent vectors to $W_{\boldsymbol{i}}^u$ are eventually contracted in the past, they must be in K_α^u.

We now return to the affine representatives $\widetilde{f}_{i_j}$ of the maps f_{i_j}.

To obtain $g_{\boldsymbol{i}}$ satisfying (33), (34), it is convenient to use graph transform techniques as in **[7]**, **[12]**.

In view of Lemmas 4.3 and 6.3, we may assume that

$$K_0 > c, \quad \varepsilon_{ij}(w) < \frac{1}{4}, \quad \varepsilon_{22}(w) < \frac{1}{8} \tag{35}$$

for $w \in B_j$ where $c > 0$ is arbitrary. In the present section, it suffices to take $c > 4$. In section 8 below, we will take $c > 117$.

We define some function spaces.

Recall that $z_0 = \pi(\boldsymbol{i})$, $z_i = \pi(\sigma^i \boldsymbol{i})$ for $i \leq 0$. Let $\varepsilon_i = \varepsilon(\pi\sigma^i \boldsymbol{i})$, $B_i^u = B_{\varepsilon_i}^u(z_i)$, $B_i^s = B_{\varepsilon_i}^s(z_i)$.

Let $\mathcal{G}_{0i}$ be the space of Lipschitz functions g from B_i^u to B_i^s with Lipschitz constant less than or equal to 1. For such a g, let $graph(g) = \{(x,y) : y = g(x) \text{ for } x \in B_i^u\}$.

For $g_1, g_2 \in \mathcal{G}_{0i}$, set

$$d_{0i}(g_1, g_2) = \sup_{x \in B_i^u} \left|g_1 x - g_2 x\right|.$$

Let $\mathcal{G}_{1i}$ be the set of continuous functions $H : B_i^u \times \mathbf{R} \to \mathbf{R}$ such that for each $x \in B_i^u$, the map

$$v \to H(x, v)$$

is linear of norm no larger than 1.

Define the metric d_{1i} on $\mathcal{G}_{1i}$ by

$$d_{1i}(H_1, H_2) = \sup_{x \in B_i^u, |v| \leq 1} \left|H_1(x,v) - H_2(x,v)\right|$$

Let $\mathcal{G}_{2i}$ be the set of continuous functions $J : B_i^u \times \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ such that for each $x \in B_i^u$, the map

$$(v, w) \to J(x, v, w)$$

is symmetric and bilinear of norm no larger than K_2 for some constant K_2 to be specified later.

Set

$$d_{2i}(J_1, J_2) = \sup_{x \in B_i^u, |v| \leq 1} \left|J_1(x,v,v) - J_2(x,v,v)\right|$$

The spaces $(\mathcal{G}_{0i}, d_{0i}), (\mathcal{G}_{1i}, d_{1i}), (\mathcal{G}_{2i}, d_{2i})$ are bounded complete metric spaces.

Let $\mathbf{Z}^- = \{k \leq 0\}$ be the non-positive integers and consider the spaces

$$\mathcal{L}_0 = \{\phi : \mathbf{Z}^- \to \bigcup_i \mathcal{G}_{0i} : \phi_i \in \mathcal{G}_{0i} \ \forall\, i\}$$

$$\mathcal{L}_1 = \{\phi : \mathbf{Z}^- \to \bigcup_i \mathcal{G}_{1i} : \phi_i \in \mathcal{G}_{1i} \ \forall\, i\}$$

$$\mathcal{L}_2 = \{\phi : \mathbf{Z}^- \to \bigcup_i \mathcal{G}_{2i} : \phi_i \in \mathcal{G}_{2i} \ \forall\, i\}$$

with the metrics

$$\overline{d}_i(\phi, \psi) = \sum_{k \geq 0} \frac{1}{2^k} d_{ik}(\phi_k, \psi_k)$$

where $\phi, \psi \in \mathcal{L}_i, i = 0, 1, 2$.

The spaces $\mathcal{L}_i$ are also bounded complete metric spaces.

Let us recall the graph transform operator [**7**]. Let $f = \widetilde{f}_{i_j}$ for some i_j, and let $g \in \mathcal{G}_{0j}$. Write $f(x,y) = (f_1(x,y), f_2(x,y))$, and let $(1,g) : B_j^u \to B_j^u \times B_j^s$ be the graph map defined by $(1,g)x = (x, gx)$.

We define

$$\Gamma(f, g) = f_2 \circ (1, g) \circ [f_1 \circ (1, g)]^{-1}.$$

It follows from our hyperbolicity assumptions the $\Gamma(f, g)$ is a well-defined mapping from $\mathcal{G}_{0j}$ to $\mathcal{G}_{0,j+1}$ for $j \leq -1$.

Returning now to the spaces $\mathcal{L}_i$ of sequences of functions, let us use the notation $\boldsymbol{g} = (g_k)_{k\leq 0}$, for elements of $\mathcal{L}_0, H = (H_k)_{k\leq 0}$, for elements of $\mathcal{L}_1$, and $J = (J_k)_{k\leq 0}$, for elements of $\mathcal{L}_2$. If $\boldsymbol{g} = (g_k)_{k\leq 0}$ is a sequence of C^2 functions, we write $D\boldsymbol{g} = (Dg_k)_{k\leq 0}, D^2\boldsymbol{g} = (D^2 g_k)_{k\leq 0}$.

We will define continuous maps

$$\begin{aligned}
&\Phi_0 : \mathcal{L}_0 \longrightarrow \mathcal{L}_0,\\
&\Phi_1 : \mathcal{L}_0 \times \mathcal{L}_1 \longrightarrow \mathcal{L}_1,\\
&\Phi_2 : \mathcal{L}_0 \times \mathcal{L}_1 \times \mathcal{L}_2 \longrightarrow \mathcal{L}_2,\\
&\Xi_1 : \mathcal{L}_0 \times \mathcal{L}_1 \longrightarrow \mathcal{L}_0 \times \mathcal{L}_1,\\
&\Xi_2 : \mathcal{L}_0 \times \mathcal{L}_1 \times \mathcal{L}_2 \longrightarrow \mathcal{L}_0 \times \mathcal{L}_1 \times \mathcal{L}_2
\end{aligned}$$

with the following properties.

FB1. $\Xi_1(g, H) = (\Phi_0(g), \Phi_1(g, H))$ and $\Xi_2(g, H, J) = (\Phi_0(g), \Phi_1(g, H), \Phi_2(g, H, J))$ for each $(g, H, J) \in \mathcal{L}_0 \times \mathcal{L}_1 \times \mathcal{L}_2$.

FB2. If $(g_k)_{k\leq 0}$ is a sequence of C^2 maps with $g_k \in \mathcal{G}_{0k}, Dg_k \in \mathcal{G}_{1k}, D^2 g_k \in \mathcal{G}_{2k}$ for all k, then $\Xi_2(\boldsymbol{g}, D\boldsymbol{g}, D^2\boldsymbol{g})_k = (\Gamma(\widetilde{f}_{i_{k-1}}, g_{k-1}), D\Gamma(\widetilde{f}_{i_{k-1}}, g_{k-1}), D^2\Gamma(\widetilde{f}_{i_{k-1}}, g_{k-1}))$.

FB3. Φ_0 is a contraction mapping; *i.e.*, it is Lipschitz with Lipschitz constant less than 1.

FB4. The map Ξ_1 is a fiber contraction map over Φ_0 in the sense of **[7]**.

FB5. The map Ξ_2 is a fiber contraction map over Ξ_1.

Once these properties are established, we proceed as follows.

Let $z_0 = (x_0, y_0) \in \pi(\boldsymbol{i})$, let $\pi_2(x, y) = y$, and let $\boldsymbol{g} = (g_k)_{k\leq 0}$ be the sequence of constant maps

$$\begin{aligned}
g_0(x) &= y_0\\
g_{k-1}(x) &= \pi_2(\widetilde{f}_{i_{-1}} \circ \cdots \circ \widetilde{f}_{i_{k-1}})^{-1}(z_0).
\end{aligned}$$

for $x \in B^u_{k-1}$.

Using the fiber contraction theorem of **[7]** we have that the sequence $\Xi_2^n(\boldsymbol{g}, D\boldsymbol{g}, D^2\boldsymbol{g})$, $n \geq 1$, converges to a fixed point $(\widetilde{\boldsymbol{g}}, \widetilde{H}, \widetilde{J})$ of Ξ_2. Letting $\pi_0 : \mathcal{L}_0 \times \mathcal{L}_1 \times \mathcal{L}_2 \to \mathcal{L}_0$, $\pi_1 : \mathcal{L}_0 \times \mathcal{L}_1 \times \mathcal{L}_2 \to \mathcal{L}_1$, $\pi_2 : \mathcal{L}_0 \times \mathcal{L}_1 \times \mathcal{L}_2 \to \mathcal{L}_2$ be the natural projections, the definitions give

$$\begin{aligned}
\pi_0\Xi_2^n(\boldsymbol{g})_0 &= \Gamma(\widetilde{f}_{i_{-1}} \circ \cdots \circ \widetilde{f}_{i_{-n}}, g_{-n})\\
\pi_1\Xi_2^n(\boldsymbol{g})_0 &= D\Gamma(\widetilde{f}_{i_{-1}} \circ \cdots \circ \widetilde{f}_{i_{-n}}, g_{-n})\\
\pi_2\Xi_2^n(\boldsymbol{g})_0 &= D^2\Gamma(\widetilde{f}_{i_{-1}} \circ \cdots \circ \widetilde{f}_{i_{-n}}, g_{-n})
\end{aligned}$$

Since all three of these sequences converge, it follows that

$$\lim_{n\to\infty} \Gamma(\widetilde{f}_{i_{-1}} \circ \cdots \circ \widetilde{f}_{i_{-n}}, g_{-n}) = g_{\boldsymbol{i}}$$

is C^2 with $Dg_{\boldsymbol{i}} = \lim \pi_1 \Xi_2^n(\boldsymbol{g})_0$ and $D^2 g_{\boldsymbol{i}} = \lim \pi_2 \Xi_2^n(\boldsymbol{g})_0$. The function $g_{\boldsymbol{i}}$ will be the C^2 function whose graph is contained in (and hence equals) $W_{\boldsymbol{i}}^u$.

Let us now define the maps Φ_i and establish their properties.

Let $f = \widetilde{f}_{i_j}$ for some i_j and let g be a C^2 function such that $g \in \mathcal{G}_{0j}, Dg \in \mathcal{G}_{1j}, D^2 g \in \mathcal{G}_{2j}$.

Write $u(x) = [f_1 \circ (1, g)]^{-1}(x)$.

Then, differentiating $\Gamma(f, g) = f_2 \circ (1, g) \circ ([f_1 \circ (1, g)]^{-1}$ we get

$$\begin{aligned} D\Gamma(f, g) &= f_{2x}(ux, gux)Du(x) + f_{2y}(ux, gux)Dg(ux)Du(x) \\ &= f_{2x}Du(x) + f_{2y}Dg(ux)Du(x) \end{aligned}$$

$$\begin{aligned} D^2\Gamma(f, g) = {} & f_{2xx}Du(x)Du(x) + f_{2xy}Du(x)Dg(ux)Du(x) + f_{2yx}Du(x)Dg(ux)Du(x) \\ & + f_{2yy}Dg(ux)Dg(ux)Du(x)Du(x) + f_{2x}D^2u(x) \\ & + f_{2y}D^2g(ux)Du(x)Du(x) + f_{2y}Dg(ux)D^2u(x) \end{aligned} \tag{36}$$

We can compute formulas for Du, D^2u in terms of f, g by differentiating the formula $f_1(ux, gux) = x$ twice and solving for Du, D^2u.

We get

$$Du(x) = \left[f_{1x}(ux, gux) + f_{1y}(ux, gux)Dg(ux)\right]^{-1} \tag{37}$$

and

$$\begin{aligned} D^2u(x) = -Du(x) \big[& f_{1xx}(Du(x))^2 + 2f_{1xy}Dg(ux)Du(x)Du(x) \\ & + f_{1yy}(Dg(ux))^2 Du(x)^2 + f_{1y}(Du)^2 D^2g(ux) \big] . \end{aligned} \tag{38}$$

For $H \in \mathcal{G}_{1j}, J \in \mathcal{G}_{2j}$, let us write H_x for the map $H(x, \cdot)$, J_x for the map $J(x, \cdot, \cdot)$. Define

$$D_1 = D_1(u, H)_x = \left[f_{1x}(ux, gux) + f_{1y}(ux, gux)H_{ux}\right]^{-1}$$

$$\begin{aligned} D_2(u, H, J)_x = -D_1 \big[& f_{1xx}D_1D_1 + 2f_{1xy}H_{ux}D_1D_1 \\ & + f_{1yy}H_{ux}H_{ux}D_1D_1 + f_{1y}D_1D_1J_{ux} \big] \end{aligned}$$

$$R_1(f, g, h)_x = \left[f_{2x}(ux, gux) + f_{2y}(ux, gux)H_{ux}\right] D_1$$

$$\begin{aligned} R_2(f, g, H, J)_x = {} & f_{2xx}D_1D_1 + f_{2xy}D_1H_{ux}D_1 + f_{2yx}D_1H_{ux}D_1 \\ & + f_{2yy}H_{ux}H_{ux}D_1D_1 + f_{2x}D_2 + f_{2y}J_{ux}D_1D_1 + f_{2y}H_{ux}D_2 \end{aligned}$$

Finally, if $\boldsymbol{g} = (g_k)_{k \le 0} \in \mathcal{L}_0, H = (H_k)_{k \le 0} \in \mathcal{L}_1, J = (J_k)_{k \le 0} \in \mathcal{L}_2$, set

$$\Phi_0(\boldsymbol{g})_k = \Gamma(\widetilde{f}_{i_{k-1}}, g_{k-1})$$

$$\Phi_1(\boldsymbol{g}, H)_k = (\Phi_0(\boldsymbol{g})_k, R_1(f_{i_{k-1}}, g_{k-1}, H_{k-1}))$$

$$\Phi_2(\boldsymbol{g}, H, J)_k = (\Phi_0(\boldsymbol{g})_k, R_1(\widetilde{f}_{i_{k-1}}, g_{k-1}, H_{k-1}), R_2(\widetilde{f}_{i_{k-1}}, g_{k-1}, H_{k-1}, J_{k-1})).$$

and define Ξ_1, Ξ_2 as in FB1.

Then, Ξ_i and Φ_i satisfy properties FB1 and FB2 above.

Let us verify the fiber contraction properties of Ξ_1, Ξ_2.

Fiber contraction property of Ξ_1. — We first show that for fixed f, g with $f = \widetilde{f}_{i_j}$ and $g : B_j^u \to B_j^s$ a given Lipschitz map of Lipschitz constant no larger than 1, $R_1(f, g, \cdot)$ maps $\mathcal{G}_{1j}$ into $\mathcal{G}_{1,j+1}$ and is a contraction.

Since the graph of g is in Q, the C^0 size of $\Gamma(f, g)$ is no larger than 1. This, and the overflowing property of f on B_j gives that $\Gamma(f, g)$ is a map from B_{j+1}^u to B_{j+1}^s.

Let $\mathrm{Lip}(\psi)$ be the Lipschitz constant of a map ψ.

As above, let $u(x) = [f_1 \circ (1, g)]^{-1}$.

Then,

$$\mathrm{Lip}(u) \leq \frac{1}{|f_{1x}|(1 - \varepsilon_{12})}$$

Using $\Gamma(f, g) = f_2 \circ (1, g) \circ [f_1 \circ (1, g)]^{-1}$, and the fact that $\mathrm{Lip}(g) \leq 1$, we get

$$\begin{aligned}
\mathrm{Lip}(\Gamma(f, g)) &\leq \mathrm{Lip}(f_2 \circ (1, g)) \,\mathrm{Lip}(u) \\
&\leq (|f_{2x}| + |f_{2y}|) \,\mathrm{Lip}(u) \\
&\leq \frac{\varepsilon_{21}}{1 - \varepsilon_{12}} + \frac{\varepsilon_{22}}{1 - \varepsilon_{12}} \\
&\leq \frac{\varepsilon_{21} + \varepsilon_{22}}{1 - \varepsilon_{12}} \leq 1
\end{aligned}$$

by (35). Thus, $\Gamma(f, g) \in \mathcal{G}_{1,j+1}$.

If $H, \widetilde{H} \in \mathcal{G}_{1j}$, we have

$$\begin{aligned}
|R_1(f, g, H) - R_1(f, g, \widetilde{H})| &\leq |(f_{2x} + f_{2y} H_{ux}) D_1(u, H) - (f_{2x} + f_{2y} \widetilde{H}_{ux}) D_1(u, \widetilde{H})| \\
&\leq |f_{2y}||D_1(u, H)||H - \widetilde{H}| \\
&\quad + (|f_{2x}| + |f_{2y}||\widetilde{H}_{ux}|)|D_1(u, H) - D_1(u, \widetilde{H})| \\
&\leq \frac{\varepsilon_{22} \cdot}{1 - \varepsilon_{12}} |H - \widetilde{H}| \\
&\quad + (|f_{2x}| + |f_{2y}||\widetilde{H}_{ux}|)|D_1(u, H) - D_1(u, \widetilde{H})|
\end{aligned}$$

To compute $|D_1(u, H) - D_1(u, \widetilde{H})|$, we use the formula

$$|G_1^{-1} - G_2^{-1}| \leq |G_1^{-1}||G_2^{-1}||G_1 - G_2|$$

which follows immediately from the formula

$$|G_2^{-1} G_2 G_1^{-1} - G_2^{-1} G_1 G_1^{-1}| \leq |G_2^{-1}||G_2 - G_1||G_1^{-1}|$$

Thus,

$$\begin{aligned}
|D_1(u, H) - D_1(u, \widetilde{H})| &\leq |D_1(u, H)||D_1(u, \widetilde{H})||f_{1y}||H - \widetilde{H}| \\
&\leq \frac{\varepsilon_{12}}{|f_{1x}|(1 - \varepsilon_{12})^2} |H - \widetilde{H}|
\end{aligned}$$

Putting the above inequalities together, and using the fact that $|\widetilde{H}| \le 1$, we get

$$|R_1(f,g,H) - R_1(f,g,\widetilde{H})| \le \frac{\varepsilon_{22}}{1-\varepsilon_{12}}|H-\widetilde{H}| + \left[\frac{\varepsilon_{21}\varepsilon_{12}}{(1-\varepsilon_{12})^2} + \frac{\varepsilon_{22}\varepsilon_{12}}{(1-\varepsilon_{12})^2}\right]|H-\widetilde{H}|$$

Now, the fact that R_1 contracts the fibers follows from the estimates for ε_{ij} already given above.

The fiber norm of $R_2(f,g,H,J)$ and fiber contractions of $R_2(f,g,H,J)$ are obtained in the same way. We just write down the final estimates and leave the computations to the reader.

We have

$$\begin{aligned}|R_2(f,g,H,J)| &\le \frac{4|D^2f|}{|f_{1x}|^2(1-\varepsilon_{12})^2} + \frac{(\varepsilon_{21}+\varepsilon_{22})4|D^2f|}{(1-\varepsilon_{12})^3|f_{1x}|^2} \\ &\quad + \left[\frac{(\varepsilon_{21}+\varepsilon_{22})\varepsilon_{12}}{(1-\varepsilon_{12})^3|f_{1x}|} + \frac{\varepsilon_{22}}{(1-\varepsilon_{12})^2|f_{1x}|}\right]|J| \\ &\le \widetilde{A}_1 + \widetilde{A}_2|J|\end{aligned}$$

and

$$|R_2(f,g,H,J) - R_2(f,g,H,\widetilde{J})| \le \frac{(\varepsilon_{21}+\varepsilon_{22})\varepsilon_{12}}{(1-\varepsilon_{12})^3|f_{1x}|}|J-\widetilde{J}| + \frac{\varepsilon_{22}}{(1-\varepsilon_{12})^2|f_{1x}|}|J-\widetilde{J}|$$

Let us summarize the conditions we need to get the required properties of R_1, R_2.

$$\frac{\varepsilon_{21}+\varepsilon_{22}}{1-\varepsilon_{12}} < 1 \tag{39}$$

$$\frac{\varepsilon_{22}}{1-\varepsilon_{12}} + \frac{\varepsilon_{21}\varepsilon_{12}}{(1-\varepsilon_{12})^2} + \frac{\varepsilon_{22}\varepsilon_{12}}{(1-\varepsilon_{12})^2} < 1 \tag{40}$$

$$\frac{(\varepsilon_{21}+\varepsilon_{22})\varepsilon_{12}}{(1-\varepsilon_{12})^3|f_{1x}|} + \frac{\varepsilon_{22}}{(1-\varepsilon_{12})^2|f_{1x}|} < 1 \tag{41}$$

Since $\varepsilon_{12} < 1/4$ and $K_0 > 4$, inequalities (39), (40), and (41) hold. Also, $|\widetilde{A}_2| < 1$. So, if we let $\widetilde{K} > 1/(1-\widetilde{A}_2)$ and $K_2 = \widetilde{K}\widetilde{A}_1$, we have

$$|J| \le K_2 \Longrightarrow |R_2(f,g,H,J)| \le K_2.$$

Hence, this K_2 is sufficient to define the space $\mathcal{G}_{2j}$.

Proof of continuous dependence of the unstable manifolds $W_{\boldsymbol{i}}^u$ on the itineraries $\boldsymbol{i}$

We have already noted that it suffices to prove that the functions $g_{\boldsymbol{i}}$ depend C^2 continuously on $\boldsymbol{i}$.

It is clear that the maps

$$(f,g) \to \Gamma(f,g), \quad (f,g,H) \to R_1(f,g,H), \quad (f,g,H,J) \to R_2(f,g,H,J)$$

are continuous. Since the spaces $\mathcal{G}_{0j}$, $\mathcal{G}_{1j}$, $\mathcal{G}_{2j}$ are bounded and the metrics on $\mathcal{L}_0$, $\mathcal{L}_1$, $\mathcal{L}_2$ give the product topologies, it follows that the maps Φ_0, Ξ_1, Ξ_2 are continuous. Also our previous estimates give that, using the non-positive itineraries $\boldsymbol{i}$ as parameters, the family $(\Phi_0)_{\boldsymbol{i}}$ is a uniform family of contractions. Similarly, the families $(\Phi_1)_{\boldsymbol{i}}, (\Phi_2)_{\boldsymbol{i}}$ are families of uniform fiber contractions. Thus, the continuous dependence of $g_{\boldsymbol{i}}$ (and hence $W^u_{\boldsymbol{i}}$) follows from Propositions 5.1 and 5.2. □

7. Fluctuation of Derivatives

We need to estimate quotients of the form

$$\frac{|D(f_{i_1} \circ \cdots \circ f_{i_n})_z(v_z)|}{|D(f_{i_1} \circ \cdots \circ f_{i_n})_w(v_w)|} \tag{42}$$

where z, w are in a K^u_α curve γ and v_z, v_w are the unit tangent vectors to γ at z, w, respectively.

The domains of the compositions $f_{i_1} \circ \cdots \circ f_{i_n}$ become narrow and possibly very non-convex. Since we wish to use the Mean Value Theorem in these domains, it will be convenient to choose certain star-shaped subdomains. This will be done in the next section. Here we present a useful Lemma.

Recall that a set E is star-shaped relative to a point $z \in E$ if for any $w \in E$, the line segment joining z to w lies in E.

For a point $z \in E$ let $\delta_z(E)$ denote the diameter of the intersection of the horizontal line through z and E.

Writing f for one of the compositions above, assume that Df maps the cone K^u_α, into itself, expands it by at least $K_0 > 1$, and that Df^{-1} maps the cone K^s_α into itself and expands it by at least K_0 as well.

For a subset E of the domain of f and $z \in E$, define

$$\Theta_z(f, E) = \sup_{w \in E} \frac{|D^2 f(w)|}{|f_{1x}(w)|} \delta_z(E)$$

where

$$|D^2 f(w)| = \max\{|f_{ijk}(w)| : i = 1, 2 \ (j, k) = (x, x), (x, y), (y, y)\}.$$

Lemma 7.1. — *Let E be a subset of the domain of f which contains z and is star-shaped relative to z. Let γ be a C^2 curve in E parametrized in the form $\gamma : x \to (x, g(x))$ where g is a C^2 function such that $|Dg(x)| \leq \alpha$ and $|D^2 g(x)| \leq K_3$ for all x. Suppose $z, w \in \gamma, w \in E$, and v_z, v_w are the unit tangent vectors to γ at z, w, respectively. Let $\Theta = \Theta_z(f, E)$ and $\delta = \delta_z(E)$.*

Then, there is a constant $C = C(\alpha, K_3) > 0$ such that

$$\frac{|Df_z(v_z)|}{|Df_w(v_w)|} \leq \exp\left(C \exp(C\Theta) \frac{|z - w|}{\delta}\right). \tag{43}$$

Proof. — We use the max norm $|(v_1, v_2)| = \max(|v_1|, |v_2|)$.

Let $z = (x, g(x)), w = (y, g(y)), v_z = (v_{1z}, v_{2z}), v_w = (v_{1w}, v_{2w})$.

Then, $|v_{1z}| = |v_{1w}| = 1$. Also, since $Df_z(v_z), Df_w(v_w)$ are in the cone K^u_α, we have

$$|f_{1_x}(z)v_{1z} + f_{1_y}(z)v_{2z}| = |Df_z(v_z)|$$

and

$$|f_{1_x}(w)v_{1w} + f_{1_y}(w)v_{2w}| = |Df_w(v_w)|$$

So,

$$\begin{aligned}|Df_w(v_w)| &= |f_{1_x}(w)v_{1w} + f_{1_y}(w)v_{2w}| \\ &= |f_{1_x}(w)v_{1w}|\Big(1 - \frac{|f_{1_y}(w)v_{2w}|}{|f_{1_x}(w)v_{1w}|}\Big) \\ &\geq |f_{1_x}(w)|(1-\alpha^2)\end{aligned}$$

and

$$\frac{|Df_z(v_z)|}{|Df_w(v_w)|} = 1 + \frac{|Df_z(v_z)| - |Df_w(v_w)|}{|Df_w(v_w)|} \tag{44}$$

$$\leq \exp(\frac{|Df_z(v_z) - Df_w(v_w)|}{|Df_w(v_w)|}) \tag{45}$$

$$\leq \exp(\frac{A_1(z,w) + A_2(z,w)}{(1-\alpha^2)}) \tag{46}$$

where

$$A_1(z,w) = \frac{|Df_z(v_z) - Df_z(v_w)|}{|f_{1_x}(w)|} \tag{47}$$

and

$$A_2(z,w) = \frac{|Df_z(v_w) - Df_w(v_w)|}{|f_{1_x}(w)|} \tag{48}$$

We consider the two terms $A_1(z, w)$ and $A_2(z, w)$ separately.

We have

$$|Df_z(v_z) - Df_z(v_w)| =$$
$$\max(|f_{1x}v_{1z} + f_{1y}v_{2z} - f_{1x}v_{1w} - f_{1y}v_{2w}|, |f_{2x}v_{1z} + f_{2y}v_{2z} - f_{2x}v_{1w} - f_{2y}v_{2w}|)$$

where the partial derivatives are all evaluated at z.

From Lemma 4.1 an upper bound for this last quantity is

$$|f_{1x}(z)|(1 + 2\alpha + \frac{1}{K_0^2} + \alpha^2)|v_z - v_w|$$

and this gives

$$A_1(z,w) \le C(\alpha)\frac{|f_{1x}(z)|}{|f_{1x}(w)|}|v_z - v_w|$$

Now, $|v_z - v_w|$ is bounded above by the product of the maximum curvature of γ and $|z-w|$. An upper bound for the curvature is the quantity K_3.

Let us use $C = C(\alpha, K_3)$ for possibly different values of C below.

As in the proof of Lemma 4.2, we get

$$\frac{|f_{1x}(z)|}{|f_{1x}(w)|} \le \exp(C\Theta)$$

So,

$$A_1(z,w) \le C\exp(C\Theta)|z-w| \le C\exp(C\Theta)\frac{|z-w|}{\delta}. \tag{49}$$

Proceding similarly, the numerator of $A_2(z,w)$ is bounded above by

$$\max_{i=1,2}(|f_{ix}(z) - f_{ix}(w)||v_{1w}| + |f_{iy}(z) - f_{iy}(w)||v_{2w}|)$$
$$\le 2 \max_{i=1,2,j=x,y} |f_{ij}(z) - f_{ij}(w)|$$

Now,

$$|f_{ix}(z) - f_{ix}(w)| \le |f_{ixx}(\tau)||z-w| + |f_{ixy}(\tau)||z-w|$$

and

$$|f_{iy}(z) - f_{iy}(w)| \le |f_{iyx}(\tau_1)||z-w| + |f_{iyy}(\tau_1)||z-w|$$

for suitable τ, τ_1, which implies that

$$A_2(z,w) \le C\Theta\exp(C\Theta)\frac{|z-w|}{\delta}. \tag{50}$$

Using $C\Theta \le \exp(C\Theta)$, (49), (50) and a different C, we see that the proof of Lemma 7.1 is complete. □

8. Distortion for compositions

In view of Lemma 7.1, to estimate quotients of the form (42), we will need to control the distortions of the compositions $\Theta_z(f_{i_1} \circ \cdots \circ f_{i_n})$ on appropriate sets.

Let $\boldsymbol{i} \in \Sigma$, and let $z \in W^s_{\mathrm{loc}}(\pi\boldsymbol{i})$ be a point in the local stable manifold of $\pi(\boldsymbol{i})$. Write $i_j = i_j(z)$ for the j-th entry in the itinerary of z, and write $F^n(z) = f_{i_{n-1}} \circ \cdots \circ f_{i_1} \circ f_{i_0}(z)$ so that $F^n(z) \in E_{i_n(z)}$ for all n.

For a curve γ, and $z, w \in \gamma$, let v_z, v_w denote the unit tangent vectors to γ at z, w, respectively.

As in section 2, let

$$E_{i_0\ldots i_n} = E_{i_0} \cap f_{i_0}^{-1}(E_{i_1\ldots i_n})$$

***Proposition 8.1* (Bounded distortion of compositions).** — *There is a constant $K_4 > 0$ such that for any $\boldsymbol{i} \in \Sigma$, any full width K^u_α curve γ in E_{i_0}, and any $n > 0$, we have*

$$\frac{|DF^n_z(v_z)|}{|DF^n_w(v_w)|} \leq K_4 \tag{51}$$

for any $z, w \in E_{i_0\ldots i_n} \cap \gamma$.

To prove this proposition, it will be convenient to cover the images $F^j(\gamma \cap E_{i_0\ldots i_n})$ by small parallelograms in which the distortions $\Theta(F)$ become small, and to make use of affine coordinates as in section 6.

Let E^s_z be the tangent space to $W^s_{\mathrm{loc}}(\pi \boldsymbol{i})$ at z, and let E^u_z be the tangent space to γ at z. Writing z_j for $F^j z$, $j \geq 0$, we translate these subspaces along the forward orbit of z by defining

$$E^s_{z_j} = DF^j_z(E^s_z), \; E^u_{z_j} = DF^j(E^u_z), \quad j \geq 0$$

This gives us a splitting of $T\mathbf{R}^2$ along the forward orbit of z and the angles between the subspaces $E^s_{z_j}, E^u_{z_j}$ are uniformly bounded away from 0 by a constant that depends on α.

Using these splittings, we can define affine coordinates along the forward orbit of z, giving local coordinate representatives $\widetilde{f}_{i_j}$ of f_{i_j}, and small parallelograms $B_j = B^u_j \times B^s_j$ with sides parallel to the subspaces $E^u_{z_j}, E^s_{z_j}$ satisfying conditions analogous to those in B1, B2 following Lemma 6.3. As we have already noted, in view of Lemmas 4.3 and 6.3, we also can arrange for the conditions (35) to hold where $c > 117$.

In these affine coordinates, the subspaces $E^u_{z_j}, E^s_{z_j}$ become horizontal and vertical, repectively. As in section 6 we use the max norm in these coordinates, so each small ε-ball $B_\varepsilon(z_j) = B(z_j, \varepsilon)$ will be a square of side length 2ε centered at z_j.

If E is any subset of B_j, and $z \in E$, let $C(z, E)$ denote the connected component of E containing z. As in section 6, we may assume that

$$B_j \subset \bigcup_{w \in E_{i_j}} B(w, \overline{K}\delta_w(E_{i_j}))$$

where $\overline{K} > 0$ is a fixed constant.

For the remainder of this section we identify f_{i_j} with its local coordinate representative $\widetilde{f}_{i_j}$.

Thus, we may assume, for $w \in B_j$,

$$|f_{i_j 1x}(w)| \geq K_0 > 117 \tag{52}$$

$$\frac{|D^2 f_{i_j}(w)|}{|f_{i_j 1x}(w)|}\delta_{z_j}(E_{i_j}) < C_0 \tag{53}$$

$$\max(\varepsilon_{12}(w), \varepsilon_{21}(w)) < \varepsilon_0, \quad \varepsilon_{22}(w) < \varepsilon_0 \tag{54}$$

$$\overline{K}\delta_{z_j}(E_{i_j}) > \operatorname{diam}(B_j) > C_1\delta_{z_j}(E_{i_j}) \tag{55}$$

where C_0, C_1, ε_0 are positive constants, $C_1 < C_0$, and $\varepsilon_0 < 1/4$.

We also may assume that $\gamma_j \equiv C(z_j, F^j\gamma \cap B_j)$ is a $K^u_{\varepsilon_0}$ curve in B_j.

Let $\varepsilon_1 \in (0, \min(C_1/2, 1))$ be small enough so that

$$\exp(156\varepsilon_1 C_0)\frac{16}{15} < 2 \tag{56}$$

Let $B_{j,\varepsilon_1} = B_j \cap B(z_j, \frac{\varepsilon_1}{2}\delta_{z_j}(E_{i_j}))$.

The definition of B_{j,ε_1} implies that

$$\Theta_{z_j}(f_{i_j}, E) \le \varepsilon_1 C_0$$

for any subset $E \subset B_{j,\varepsilon_1}$.

We use ∂B to denote the boundary of a set B.

Since f_{i_j} maps E_{i_j} to a full-width rectangle in Q, there is a constant $K > 0$ such that

$$\delta_{z_j}(E_{i_j}) > K\left|f_{i_j 1x}(z_j)\right|^{-1}$$

Therefore, since $\delta_{z_j}(B_{j,\varepsilon_1}) = \varepsilon_1\delta_{z_j}(E_{i_j})$, Lemma 4.2 provides a constant $K_5 > 0$ such that

$$\operatorname{dist}(f_{i_j}(z_j), \partial f_{i_j}(\gamma_j \cap B_{j,\varepsilon_1})) \ge K_5\varepsilon_1 \tag{57}$$

For $z_j \in E_{i_j}$, let

$$\widetilde{B}_j = \begin{cases} B_{j,\varepsilon_1} & \text{if } \frac{1}{2}\varepsilon_1\delta_{z_j}(E_{i_j}) < \frac{K_5\varepsilon_1}{2K_0} \\ B(z_j, K_5\varepsilon_1/2K_0) & \text{if } \frac{1}{2}\varepsilon_1\delta_{z_j}(E_{i_j}) \ge \frac{K_5\varepsilon_1}{2K_0} \end{cases}$$

Thus, each $\widetilde{B}_j \subseteq B_{j,\varepsilon_1}$.

Since f_{i_j} expands horizontal distances by at least K_0, we have that

$$\operatorname{dist}(f_{i_j}(z_j), \partial f_{i_j}(\gamma_j \cap \widetilde{B}_j)) \ge \frac{K_5\varepsilon_1}{2}$$

so

$$f_{i_j}(C(z_j, \gamma_j \cap \widetilde{B}_j)) \supset C(f_{i_j}z_j, f_{i_j}\gamma_j \cap \widetilde{B}_{j+1})$$

The set $\widetilde{B}_{j,n} = \widetilde{B}_{i_j} \cap F^{-1}\widetilde{B}_{i_{j+1}} \cap \cdots \cap F^{-(n-1-j)}\widetilde{B}_{i_{n-1}}$ is a narrow curvilinear rectangle around z_j.

Let

$$\alpha_{j,n} = \operatorname{dist}(z_j, \gamma_j \cap \partial\widetilde{B}_{j,n})$$

Let $E_{j,n}$ be the curvilinear rectangle whose left and right boundary curves are pieces of the left and right boundaries of $\widetilde{B}_{j,n}$ and whose top and bottom boundary curves are horizontal line segments each of whose distance from z_j is $\alpha_{j,n}$.

Lemma 8.2. — *The curvilinear rectangle $E_{j,n}$ is star-shaped relative to z_j.*

Proof. — Let $\boldsymbol{b}_1, \boldsymbol{b}_2$ denote the left and right boundary curves of $E_{j,n}$ and let ℓ_1, ℓ_2 denote the top and bottom boundary curves (which are horizontal line segments).

Let $w \in E_{j,n}$, let $\ell_{j,w}$ denote the line segment joining z_j to w, and let $\partial_{vert} E_{j,n}$ denote the union $\boldsymbol{b}_1 \bigcup \boldsymbol{b}_2$.

Since z_j, w lie between the horizontal lines through ℓ_1, ℓ_2, any intersection of $\ell_{j,w}$ and the boundary of $E_{j,n}$, must be in $\partial_{vert} E_{j,n}$. Thus, to show that $\ell_{j,w}$ is contained in $E_{j,n}$ it suffices to show that

$$(\ell_{j,w} \setminus \{z_j, w\}) \cap \partial_{vert} E_{j,n} = \varnothing. \tag{58}$$

Assuming (58) fails we will get a contradiction.

By construction, $\boldsymbol{b}_1, \boldsymbol{b}_2$ are $K^s_{\varepsilon_0}$-curves. This and the assumption that $\varepsilon_0 < 1/4$, imply that any line segment joining z_j to a point $\overline{z}$ in $\partial_{vert} E_{j,n}$ must have slope no larger than $4/3$.

But since z_j and w lie in $E_{j,n}$, if (58) fails there is a point $\overline{z} \in \partial_{vert} E_{j,n}$ such that the line $\ell_{j,w}$ is parallel to the tangent vector to $\partial_{vert} E_{j,n}$ at $\overline{z}$. However, $\boldsymbol{b}_1, \boldsymbol{b}_2$ are $K^s_{\varepsilon_0}$-curves, so their tangent vectors have slope no smaller than 4 which is our contradiction, proving Lemma 8.2. □

Lemma 8.3. — *Fix $j \geq 0$. Then, for each $n > j$, we have*

$$\Theta_{z_j}(F^{n-j}, E_{j,n}) \leq 13\varepsilon_1 C_0. \tag{59}$$

Proof. — The proof is by induction on $n - j$. Clearly (59) holds for $n - j = 1$. Assuming it holds for $n - j$, we show it holds for $n + 1 - j$.

Let $z = z_j$, $f = F$, $g = F^{n-j}$, $E_f = E_{n,n+1} = \widetilde{B}_{gz}$, $E_g = E_{j,n}$, $h = f \circ g$, and $E_h = E_{j,n+1}$.

Let $\Delta f = \delta_{z_n}(E_f)$, $\Delta g = \delta_{z_j}(E_g)$, $\Delta h = \delta_{z_j}(E_h)$.

We use $\Theta(f) = \Theta_{gz}(f, E_f)$, $\Theta(g) = \Theta_z(g, E_g)$, $\Theta(h) = \Theta_z(h, E_h)$.

Consider the quotient

$$\frac{g_{1x}(w)}{g_{1x}(z)}$$

where $w \in E_h$.

Since the left and right boundary curves of E_h are $K^s_{\varepsilon_0}$ curves, and $\varepsilon_0 < 1/4$ we have that $|w - z| \leq 3\Delta h$.

Since both $g_{1x}(z), g_{1x}(w)$ have absolute value greater than 1, they have the same sign. Replacing g by $-g$ if necessary, we may assume these signs are positive.

By the mean value theorem,

$$\begin{aligned}
\left|\log g_{1x}(w) - \log g_{1x}(z)\right| &\leq \frac{|g_{1xx}(\tau)|}{|g_{1x}(\tau)|}|z - w| + \frac{|g_{1xy}(\tau)|}{|g_{1x}(\tau)|}|z - w| \\
&\leq 6\Theta(g)\frac{\Delta h}{\Delta g}
\end{aligned}$$

so,

$$\exp\Big(-6\Theta(g)\frac{\Delta h}{\Delta g}\Big) \le \frac{|g_{1x}(w)|}{|g_{1x}(z)|} \le \exp\Big(6\Theta(g)\frac{\Delta h}{\Delta g}\Big). \tag{60}$$

Further, setting $\zeta = \exp(12\Theta(g)\Delta h/\Delta g)$, we have, for any $w, \tau \in E_h$,

$$\begin{aligned} \frac{|g_{1x}(w)|}{|g_{1x}(\tau)|} &= \frac{|g_{1x}(w)|}{|g_{1x}(z)|}\frac{|g_{1x}(z)|}{|g_{1x}(\tau)|} \\ &\le \zeta. \end{aligned} \tag{61}$$

Similarly, if $\tau_1, \tau_2 \in E_g$, then

$$\frac{|g_{1x}(\tau_2)|}{|g_{1x}(\tau_1)|} \le \exp(12\Theta(g)) < \exp(156\varepsilon_1 C_0) < 2 \tag{62}$$

Also note that if ℓ_0 is the full width horizontal line segment through z in E_h, then $g(\ell_0)$ is a full width $K^u_{\varepsilon_0}$ curve in E_f, and there is a $\tau \in \ell_0$ for which $|g_{1x}(\tau)|\Delta h =$ length $(g(\ell_0)) \le \frac{5}{4}\Delta f$.

This gives

$$\frac{\Delta h}{\Delta f} \le \frac{5}{4|g_{1x}(\tau)|}. \tag{63}$$

Observe that it follows from the definition of E_g and (57) that, for some $\tau_1 \in E_g$,

$$|g_{1x}(\tau_1)|\Delta g \ge K_5\varepsilon_1.$$

So, by (62),

$$\begin{aligned} \frac{\Delta h}{\Delta g} &\le \frac{5\Delta f}{4|g_{1x}(\tau)|}\frac{|g_{1x}(\tau_1)|}{K_5\varepsilon_1} \\ &\le \frac{5\Delta f}{2K_5\varepsilon_1} \\ &\le \frac{3}{K_0} \end{aligned} \tag{64}$$

Let us estimate

$$\Theta(h) = \max_{w\in E_h} \frac{|D^2h(w)|}{|h_{1x}(w)|}\Delta h \tag{65}$$

Let

$$\eta = \frac{1}{1-\varepsilon_{21}(g)\varepsilon_{12}(f)} \le \frac{16}{15}.$$

and

$$\zeta = \exp(12\Theta(g)\frac{\Delta h}{\Delta g}) \le \exp\Big(156\varepsilon_1 C_0\frac{3}{K_0}\Big)$$

Recall that K_0 and ε_1 were chosen so that

$$K_0 > 117 \tag{66}$$

$$\eta\zeta < 2 \tag{67}$$

Recall the Chain Rule formulas (15)–(18).

By (15), we have

$$\begin{aligned}|h_{1_x}(w)| &= |f_{1_x}g_{1_x} + f_{1_y}g_{2_x}| \\ &\geq |f_{1_x}g_{1_x}|(1 - \varepsilon_{21}(g)\varepsilon_{12}(f)) \\ &= |f_{1_x}g_{1_x}|\eta^{-1}.\end{aligned}$$

Write $\varepsilon_2(f) = \max(\varepsilon_{12}(f), \varepsilon_{22}(f))$.

From (16) we get

$$\begin{aligned}\left|\frac{h_{i_{xx}}(w)}{h_{1_x}(w)}\Delta h\right| \leq \eta \Big[\Theta(f)|g_{1_x}(w)|\frac{\Delta h}{\Delta f} + 2\Theta(f)|g_{2_x}(w)|\frac{\Delta h}{\Delta f} \\ + \Theta(f)\varepsilon_{21}(g)|g_{2_x}(w)|\frac{\Delta h}{\Delta f} + \Theta(g)\max(1, \varepsilon_{21}(f))(1 + \varepsilon_2(f))\frac{\Delta h}{\Delta g}\Big]\end{aligned}$$

Now, using (61), (63), (64), (66),(67), we get

$$\begin{aligned}\left|\frac{h_{i_{xx}}(w)}{h_{1_x}(w)}\Delta h\right| &\leq \eta\left[\Theta(f)3\zeta(1 + 2\varepsilon_{21}(g) + \varepsilon_{21}(g)^2) + 2\Theta(g)\frac{\Delta h}{\Delta g}\right] \\ &\leq 6\eta\zeta\Theta(f) + 3\Theta(g)\frac{3}{K_0} \\ &\leq 12\Theta(f) + \frac{1}{13}\Theta(g)\end{aligned}$$

Similarly,

$$\begin{aligned}\left|\frac{h_{i_{xy}}(w)}{h_{1_x}(w)}\Delta h\right| &\leq \eta\Big[3\Theta(f)\zeta(\varepsilon_{12}(g) + \varepsilon_{22}(g) + \\ &\qquad +\varepsilon_{12}(g)\varepsilon_{21}(g) + \varepsilon_{22}(g)\varepsilon_{21}(g)) + 2\Theta(g)\frac{\Delta h}{\Delta g}\Big] \\ &\leq 3\Theta(f) + 3\Theta(g)\frac{3}{K_0} \\ &\leq 3\Theta(f) + \frac{1}{13}\Theta(g)\end{aligned}$$

and

$$\begin{aligned}\left|\frac{h_{i_{yy}}(w)}{h_{1_x}(w)}\Delta h\right| &\leq \eta\left[3\Theta(f)\zeta(\varepsilon_{12}(g)^2 + 2\varepsilon_{22}(g)\varepsilon_{12}(g) + \varepsilon_{22}(g)^2)\right] \\ &\leq 2\Theta(f) + 3\Theta(g)\frac{3}{K_0} \\ &\leq 2\Theta(f) + \frac{1}{13}\Theta(g)\end{aligned}$$

In all cases we have

$$\Theta(h) \leq 12\Theta(f) + \frac{1}{13}\Theta(g) \leq 13\varepsilon_1 C_0$$

proving Lemma 8.3. □

Proof of Proposition 8.1. — The curvilinear rectangles $\widetilde{B}_{j,n}$ are determined by the orbit segment $\{z_j = F^j(z)\}_{j=0}^n$. We write this as

$$\widetilde{B}_{j,n} = \widetilde{B}_{z,j,n}$$

There are analogous sets

$$\widetilde{B}_{w,j,n} = \widetilde{B}_{F^j w} \cap \cdots \cap F^{-(n-1-j)} \widetilde{B}_{F^{n-1} w}$$

where $\widetilde{B}_{F^\ell w}$ is a suitable small parallelogram centered at $F^\ell w$ for any $w \in E_{i_0 \ldots i_n} \cap \gamma$.

Let γ, z, w, v_z, v_w be as in the hypotheses of the Proposition, and consider $F^n z$, $F^n w \in \gamma \cap E_{i_n}$.

We can connect these points by a chain $F^n z = F^n(w_1)$, $F^n(w_2)$, $\ldots$, $F^n(w_k) = F^n(w)$ with $k \leq C_1(\alpha, \varepsilon)$ such that, for every $\ell = 1, \ldots k-1$, and every $0 \leq j \leq n$,

$$F^j(w_{\ell+1}) \in \widetilde{B}_{w_\ell, j, n}$$

then, it follows from Lemmas 7.1 and 8.3 that, for some constant $C_2(\alpha, \varepsilon)$, we have

$$\frac{|DF^n_{w_\ell}(v_{w_\ell})|}{|DF^n_{w_{\ell+1}}(v_{w_{\ell+1}})|} \leq C_2(\alpha, \varepsilon_1) \tag{68}$$

in the special affine coordinates centered at w_ℓ. Changing back to the standard coordinates on Q simply makes (68) hold with a different constant $C_2 = C_2(\alpha, \varepsilon_1)$.

Then,

$$\frac{|DF^n_z(v_z)|}{|DF^n_w(v_w)|} = \prod_{\ell=1}^{k-1} \frac{|DF^n_{w_\ell}(v_{w_\ell})|}{|DF^n_{w_{\ell+1}}(v_{w_{\ell+1}})|} \leq C_2^{k-1}$$

proving Proposition 8.1. □

9. Sinai Local Measures

For two points z_1, z_2 in an unstable manifold $W^u_{\boldsymbol{i}}$ and unit tangent vectors v_1, v_2 to $W^u_{\boldsymbol{i}}$ at z_1, z_2, respectively, let $D^u F(z_i) = |DF_{z_i}(v_i)|$ denote the Jacobian of F at z_i along $W^u_{\boldsymbol{i}}$. We know that $W^u_{\boldsymbol{i}}$ is a full-width K^u_α curve in E_{i_0}. Also, the curve $f_{i_0} W^u_{\boldsymbol{i}}$ is a full width K^u_α curve in Q.

Proposition 9.1. — *Suppose $\boldsymbol{i} = (\ldots i_{-n} \ldots i_0)$ is an arbitrary infinite non-positive itinerary and let $W^u_{\boldsymbol{i}}$ denote its unstable manifold. Write $f_{i_0} W^u_{\boldsymbol{i}} = \operatorname{graph} g_{\boldsymbol{i}}$ where $g_{\boldsymbol{i}} : I \to I$ is the C^2 function given in Theorem 6.1. Suppose $x_1, x_2 \in I$ and $z_1 = (x_1, g_{\boldsymbol{i}} x_1)$, $z_2 = (x_2, g_{\boldsymbol{i}} x_2)$. Then, the infinite product*

$$\xi(x_1, x_2, \boldsymbol{i}) = \prod_{s=1}^{\infty} \frac{D^u F(F^{-s} z_1)}{D^u F(F^{-s} z_2)} \tag{69}$$

converges and depends continuously on $(x_1, x_2, \boldsymbol{i})$.

Moreover, there is a constant $K_6 > 0$ *independent of* $(x_1, x_2, \boldsymbol{i})$ *such that*

$$K_6^{-1} < \xi(x_1, x_2, \boldsymbol{i}) < K_6 \tag{70}$$

Proof. — It clearly suffices to prove the upper bound in (70) since interchanging z_1 and z_2 would then give the lower bound.

Let $\overline{z}_1 = f_{i_0}^{-1} z_1, \overline{z}_2 = f_{i_0}^{-1} z_2$ so that $\overline{z}_1, \overline{z}_2 \in W_{\boldsymbol{i}}^u$.

We use the local coordinates $\widetilde{f}_{i_j}$ and rectangles $\widetilde{B}_j$ of the previous section. To avoid confusion, we will use $\widetilde{F}^s(z) = (\widetilde{f}_{i_{-1}} \circ \cdots \circ \widetilde{f}_{i_{-s}})(z)$ and $\widetilde{F}^{-s} = (\widetilde{f}_{i_{-1}} \circ \cdots \circ \widetilde{f}_{i_{-s}})^{-1}$ instead of identifying F, f_{i_j} with $\widetilde{F}, \widetilde{f}_{i_j}$ as in the preceding section. We use $\widetilde{B}_z$ for the affine neighborhood centered at z.

In our local coordinates, with $z \in E_{i_0}, W_{\boldsymbol{i}}^u \cap \widetilde{B}_z$ becomes a $K_{\varepsilon_0}^u$ curve. Also, there is a sequence $\overline{z}_1 = w_1, \ldots, w_k = \overline{z}_2$ of points in $W_{\boldsymbol{i}}^u$ such that

$$d(w_{j+1}, w_j) < \frac{\varepsilon_1 K_5}{2} \tag{71}$$

and

$$k \leq \left[\frac{2}{\varepsilon_1 K_5}\right] + 1. \tag{72}$$

Recall that $\widetilde{F}^{-1} w_j = \widetilde{f}_{i_{-1}}^{-1} w_j$.

Further, $\widetilde{f}_{i_{-1}}(\widetilde{f}_{i_{-1}}^{-1} W_{\boldsymbol{i}}^u \cap \widetilde{B}_{\widetilde{F}^{-1} w_j})$ contains the intersection of $W_{\boldsymbol{i}}^u$ with the ball of radius $\varepsilon_1 K_5/2$ about w_j. Since w_{j+1} is in this latter set, we have $\widetilde{F}^{-1} w_{j+1} \in \widetilde{B}_{\widetilde{F}^{-1} w_j}$.

Analogously, we have $\widetilde{F}^{-s} w_{j+1} \in \widetilde{B}_{\widetilde{F}^{-s} w_j}$ for every $s \geq 1$.

Now, there is a constant $C_6 > 0$ such that

$$\begin{aligned}
\prod_{s=1}^{\infty} \frac{D^u F(F^{-s} z_1)}{D^u F(F^{-s} z_2)} &\leq C_6 \prod_{s=1}^{\infty} \frac{D^u F(F^{-s} \overline{z}_1)}{D^u F(F^{-s} \overline{z}_2)} \\
&= \prod_{s=1}^{\infty} \prod_{j=1}^{k-1} \frac{D^u F(F^{-s} w_j)}{D^u F(F^{-s} w_{j+1})}
\end{aligned}$$

so, to prove Proposition 9.1 it suffices to show

$$\prod_{s=1}^{\infty} \frac{D^u F(F^{-s} w_j)}{D^u F(F^{-s} w_{j+1})} \leq K_7 \tag{73}$$

for some $K_7 > 0$ and any j.

Since the angles between E_z^u and E_z^s are bounded by a constant depending on α, the linear maps $D\widetilde{F}(\widetilde{F}^{-s}(w_j))$ and $DF(F^{-s}(w_j))$ are conjugate by a linear map whose images on unit vectors are bounded above and below by constants which depend only on α. A similar statement holds replacing w_j by w_{j+1}. Hence, there is a constant $C_5 = C_5(\alpha)$ such that, for any $s \geq 1$ and any j,

$$C_5^{-1} \frac{D^u \widetilde{F}(\widetilde{F}^{-s} w_j)}{D^u \widetilde{F}(\widetilde{F}^{-s} w_{j+1})} \leq \frac{D^u F(F^{-s} w_j)}{D^u F(F^{-s} w_{j+1})} \leq C_5 \frac{D^u \widetilde{F}(\widetilde{F}^{-s} w_j)}{D^u \widetilde{F}(\widetilde{F}^{-s} w_{j+1})} \tag{74}$$

so, it suffices to find $K_7 > 0$ such that

$$\prod_{s=1}^{\infty} \frac{D^u\widetilde{F}(\widetilde{F}^{-s}w_j)}{D^u\widetilde{F}(\widetilde{F}^{-s}w_{j+1})} \leq K_7 \tag{75}$$

By Lemmas 7.1 and 8.3, there is a constant $K_1 > 0$ such that

$$\frac{\left|D\widetilde{F}_z^N(v_z)\right|}{\left|D\widetilde{F}_{\overline{z}}^N(v_{\overline{z}})\right|} \leq K_1 \tag{76}$$

for any $N > 1, \overline{z} \in E_{z,N}$ and unit vectors $v_z, v_{\overline{z}}$ tangent to $W^u(z), W^u(\overline{z})$, respectively.

Let N be large enough so that

$$\tau = \frac{K_1}{K_0^N K_5 \varepsilon_1} < 1. \tag{77}$$

By definition, $\widetilde{F}^{N-1}(E_{\widetilde{F}^{-sN}w_j,N})$ is a full-width subrectangle of $\widetilde{B}_{\widetilde{F}^{N-1-sN}w_j}$. So, the $\widetilde{F}$ image of a full-width horizontal line segment in $\widetilde{B}_{\widetilde{F}^{N-1-sN}w_j}$ contains a curve of horizontal width at least $\varepsilon_1 K_5$.

Thus, setting $\widetilde{w}_j = \widetilde{F}^{-1}w_j$ and $\delta_{i_s} = \delta_{\widetilde{F}^{-sN}\widetilde{w}_j}(E_{\widetilde{F}^{-sN}\widetilde{w}_j,N})$, we have

$$\frac{\varepsilon_1 K_5}{\left|D^u\widetilde{F}^N(\tau_N)\right|} \leq \delta_{i_s} \leq \frac{1}{K_0^N} \tag{78}$$

for some $\tau_N \in W^u(\widetilde{F}^{-sN}\widetilde{w}_j) \cap E_{\widetilde{F}^{-sN}\widetilde{w}_j,N}$.

Then,

$$\begin{aligned}
\left|\widetilde{F}^{-sN}\widetilde{w}_{j+1} - \widetilde{F}^{-sN}\widetilde{w}_j\right| &= \left|\widetilde{F}^{-sN+N}\widetilde{w}_{j+1} - \widetilde{F}^{-sN+N}\widetilde{w}_j\right| \frac{1}{\left|D^u\widetilde{F}^N(\widetilde{\tau_N})\right|} \\
&\leq \frac{\varepsilon_1 K_5}{\left|D^u\widetilde{F}^N(\tau_N)\right|} \frac{\left|D^u\widetilde{F}^N(\tau_N)\right|}{\varepsilon_1 K_5\left|D^u\widetilde{F}^N(\widetilde{\tau_N})\right|} \\
&\qquad \cdot \left|\widetilde{F}^{-sN+N}\widetilde{w}_{j+1} - \widetilde{F}^{-sN+N}\widetilde{w}_j\right| \\
&\leq \delta_{i_s} \frac{K_1}{\varepsilon_1 K_5}\left|\widetilde{F}^{-sN+N}\widetilde{w}_{j+1} - \widetilde{F}^{-sN+N}\widetilde{w}_j\right| \\
&\leq \delta_{i_s}\delta_{i_{s-1}} \ldots \delta_{i_1} \left(\frac{K_1}{\varepsilon_1 K_5}\right)^s \left|\widetilde{w}_{j+1} - \widetilde{w}_j\right|
\end{aligned}$$

giving

$$\begin{aligned}
\frac{\left|\widetilde{F}^{-sN}\widetilde{w}_{j+1} - \widetilde{F}^{-sN}\widetilde{w}_j\right|}{\delta_{i_s}} &\leq \left(\frac{K_1}{\varepsilon_1 K_5}\right)^s \frac{1}{K_0^{N(s-1)}}\left|\widetilde{w}_{j+1} - \widetilde{w}_j\right| \\
&\leq \frac{K_1}{\varepsilon_1 K_5}\tau^{s-1}.
\end{aligned} \tag{79}$$

Hence,

$$\begin{aligned}\prod_{s=1}^{\infty} \frac{D^u\widetilde{F}(\widetilde{F}^{-s}w_j)}{D^u\widetilde{F}(\widetilde{F}^{-s}w_{j+1})} &= \prod_{s=1}^{\infty} \frac{D^u\widetilde{F}^N(\widetilde{F}^{-sN-1}w_j)}{D^u\widetilde{F}^N(\widetilde{F}^{-sN-1}w_{j+1})} \\ &= \prod_{s=1}^{\infty} \frac{D^u\widetilde{F}^N(\widetilde{F}^{-sN}\widetilde{w}_j)}{D^u\widetilde{F}^N(\widetilde{F}^{-sN}\widetilde{w}_{j+1})} \\ &\leq \exp\Big(\sum_{s=1}^{\infty} C_7\tau^s\Big) \equiv K_7\end{aligned}$$

using Lemma 7.1 and (79).

Since the functions g_i depend continuously on i in the C^2 topology, the continuity statement in 9.1 follows from the fact that given $\varepsilon > 0$, there is an $N_0 > 0$ such that if $N \geq N_0$, we have

$$\left| \prod_{j=N_0}^{\infty} \frac{D^uF(F^{-s}z_1)}{D^uF(F^{-s}z_2)} - 1 \right| < \varepsilon$$

which is immediate from the proof just given.

This proves Proposition 9.1. □

For a C^2 curve γ in Q, let ρ_γ denote the Riemannian measure on γ.

From Proposition 9.1 we get the existence of the following limit

$$\lim_{n\to\infty} \prod_{s=1}^{n} \frac{D^uF(F^{-s}z_1)}{D^uF(F^{-s}z_2)} = \xi(z_1, z_2) = \xi_i(z_1, z_2) \tag{80}$$

for any two points $z_1, z_2 \in W^u_i$. Letting γ denote W^u_i, we can use ρ_γ and the ratios $\xi(z_1, z_2)$ obtained in the preceding limits to get special measures on the unstable manifolds. More precisely, following Sinai in **[13]**, Lecture 16, we define

$$\nu_{z_1,\gamma}(A) = \int_A \xi(z_1, z_2)d\rho_\gamma(z_2).$$

It is easy to see that if z_3 is another point in γ, then $\nu_{z_3,\gamma}(A) = \xi(z_3, z_1)\nu_{z_1,\gamma}(A)$, so the measures $\nu_{z_1,\gamma}$ and $\nu_{z_3,\gamma}$ are simply rescalings of each other. In particular, if $A, B \subset \gamma$ and $\nu_{z_1,\gamma}(B) < \infty$, then $\nu_{z_1,\gamma}(A)/\nu_{z_1,\gamma}(B)$ is independent of z_1.

For $z_1 \in \gamma \cap \widetilde{Q}$, let E_{i_1} be the element of $\{E_i\}$ containing Fz_1, and let $\gamma_1 = W^u_{\sigma i}$.

The family of measures $\{\nu_{z_1,\gamma}\}$ is invariant in the sense that if $A, B \subset \gamma$, $F(A)$, $F(B) \subset \gamma_1$, and $\nu_{z_1,\gamma}(B) < \infty$, then $\nu_{z_1,\gamma}(A)/\nu_{z_1,\gamma}(B) = \nu_{Fz_1,\gamma_1}(FA)/\nu_{Fz_1,\gamma_1(FB)}$.

We call the family of measures $\nu_{z_1,\gamma}$ *Sinai local measures* or just *local measures.*

10. Absolute Continuity of the Stable Foliation

We know that for each non-negative itinerary $\boldsymbol{a} = (a_0, a_1, \dots)$ there is a $C^1 K^s_\alpha$ curve $W^s(\boldsymbol{a}) = \bigcap_{n\geq 0} E_{a_0\dots a_{n-1}}$ of full height in Q.

Note that two points in $\widetilde{Q}$ with different forward itineraries have disjoint stable manifolds since the interiors of the $E_i's$ are disjoint. Thus, the set $\{W^s(\boldsymbol{a}) : \pi\boldsymbol{a} \in \widetilde{Q}\}$ is a foliation of its union. We call this the *stable foliation*. Let $\mathcal{W} = \{W^s(\boldsymbol{a}) : \pi\boldsymbol{a} \in \widetilde{Q}\}$ denote this foliation. We denote the union $\bigcup\{W^s(\boldsymbol{a}) : \pi\boldsymbol{a} \in \widetilde{Q}\}$ by $\mathcal{W}^+$. Note that $\mathcal{W}^+$ is a Borel subset of Q of full two-dimensional Lebesgue measure in Q. For any two full width C^2 K^u_α curves γ, η, let $\pi_{\gamma\eta}$ be the holonomy projection from γ to η along the foliation $\mathcal{W}$. That is, for $z \in \gamma \cap \mathcal{W}^+$, and $W^s(\boldsymbol{a})$ the leaf of $\mathcal{W}$ which contains z, $\pi_{\gamma\eta}(z)$ is the unique point of intersection of $W^s(\boldsymbol{a})$ and η. As above, for any C^2 K^u_α-curve γ, let ρ_γ denote the Riemannian measure on γ. Recall that the foliation $\mathcal{W}$ is called *absolutely continuous* if

(AC-1) each full-width C^2 K^u_α-curve γ meets $\mathcal{W}^+$ in a set of positive ρ_γ measure and

(AC-2) the image measure $\pi_{\gamma\eta\star}\rho_\gamma$ is equivalent to the measure ρ_η.

Proposition 10.1. — *The foliation $\mathcal{W}$ is absolutely continuous*

Before we can prove Proposition 10.1, we need a couple of Lemmas.

The next Lemma is well-known and elementary. Since the proof is short, we include it for completeness.

Lemma 10.2. — *Suppose that $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ are sequences of numbers in the open unit interval $(0,1)$ such that*

$$-\sum_{i\geq 1} x_i \log y_i < \infty. \tag{81}$$

For $\varepsilon > 0$ and non-negative integer n, let $D_n = \{i : y_i \leq \exp(-\varepsilon n)\}$.
Then,

$$\sum_{n\geq 1}\sum_{i\in D_n} x_i < \infty. \tag{82}$$

Proof. — For each $n \geq 1$, let

$$E_n = \{i : \exp(-\varepsilon(n+1)) < y_i \leq \exp(-\varepsilon n)\}$$

Then, $D_n = \bigsqcup_{j\geq n} E_j$, where $\bigsqcup$ denotes disjoint union, so

$$\sum_{n\geq 1}\sum_{i\in D_n} x_i = \sum_{n\geq 1}\sum_{j\geq n}\sum_{i\in E_j} x_i.$$

Letting $c_j = \sum_{i\in E_j} x_i$, this last sum is just

$$\begin{array}{cccc} c_1 & + \; c_2 & + & c_3 + \ldots \\ & + \; c_2 & + & c_3 + \ldots \\ & & + & c_3 + \ldots \\ & & \vdots & \\ & = \; \sum_{j\geq 1} j c_j & & \end{array}$$

Now, $i \in E_j$ implies that $-\log y_i \geq \varepsilon j$ or $-x_i \log y_i \geq \varepsilon j x_i$ which gives

$$\begin{aligned} -\sum_{i \in E_j} x_i \log y_i &\geq \sum_{i \in E_j} \varepsilon j x_i \\ &= \varepsilon j c_j \end{aligned}$$

Hence,

$$\varepsilon \sum_{j \geq 1} j c_j \leq \sum_{j \geq 1} \sum_{i \in E_j} -x_i \log y_i \leq -\sum x_i \log y_i < \infty$$

which implies that $\sum_{j \geq 1} j c_j < \infty$. □

In the next lemma, we will use the geometric condition $G3$. Each $z \in \widetilde{Q}$ has a unique forward itinerary $(a_0(z), a_1(z), \dots)$ with $F^n(z) \in \operatorname{int} E_{a_n(z)}$.

Lemma 10.3. — *Let γ be a C^2 K^u_α-curve of full-width in Q such that $\rho_\gamma(\gamma \cap \widetilde{Q}) > 0$. Let $\varepsilon > 0$. For ρ_γ-almost all points $z \in \gamma \cap \widetilde{Q}$, there is a positive integer $n(z) > 0$ such that if $n \geq n(z)$, then*

$$\delta_{F^n(z)}(E_{a_n(z)}) > \exp(-\varepsilon n).$$

Proof. — For ease of notation, if A is a subset of γ, let us write $|A|$ for $\rho_\gamma(A)$.

Let $D_n = \{i \geq 1 : \delta_{i,\min} < e^{-\varepsilon n}\}$.

In view of lemma 10.2, the condition G3 implies that

$$\sum_{n \geq 1} \sum_{i \in D_n} \delta_{i,\max} < \infty. \tag{83}$$

Let $V_n = \{z \in \gamma \cap \widetilde{Q} : \delta_{F^n(z)}(E_{a_n(z)}) \leq e^{-\varepsilon n}\}$.

We will show

$$\sum_{n \geq 1} |V_n| < \infty. \tag{84}$$

Once this is done, the Borel-Cantelli Lemma gives that ρ_γ-almost all points of γ lie in at most finitely many of the $V_n's$ which proves Lemma 10.3.

Let $\mathcal{A}_n$ be the set of finite itineraries $(a_0, \dots, a_{n-1})$ which occur for points in $\widetilde{Q}$.

For a given finite sequence $a_0, a_1, \dots, a_{n-1} \in \mathcal{A}_n$, let

$$V_n(a_0, \dots, a_{n-1}) = \{z \in V_n : F^i z \in E_{a_i} \text{ for } 0 \leq i < n\}.$$

Then,

$$V_n(a_0, \dots, a_{n-1}) = \bigcup_{i \geq 0} \left(\gamma \cap E_{a_0 \dots a_{n-1} i} \cap \widetilde{Q}\right)$$

and this last union is disjoint.

Also, V_n is the disjoint union of the $V_n(a_0, \dots, a_{n-1})$ as these finite itineraries vary in $\mathcal{A}_n$.

The bounded distortion of compositions (Proposition 8.1) gives us a constant $K > 0$ such that for $(a_0, \ldots, a_{n-1}i) \in \mathcal{A}_{n+1}$, and $z \in \gamma \cap E_{a_0 \ldots a_{n-1}i}$,

$$\frac{|\gamma \cap E_{a_0 \ldots a_{n-1}i}|}{|\gamma \cap E_{a_0 \ldots a_{n-1}}|} \leq K\delta_{F^n(z)}(E_i)$$

Also, the definition of $V_n(a_0, \ldots, a_{n-1})$ gives us that $\delta_{a_n(z),\min} \leq e^{-\varepsilon n}$; *i.e.*, that $a_n(z) \in D_n$.

Thus,

$$V_n(a_0, \ldots, a_{n-1}) \subseteq \bigsqcup_{i \in D_n} \gamma \cap E_{a_0, \ldots, a_{n-1}i} \cap \widetilde{Q}$$

This gives

$$\begin{aligned} |V_n| &\leq \sum_{(a_0 \ldots a_{n-1}) \in \mathcal{A}_n} \sum_{i \in D_n} |\gamma \cap E_{a_0, \ldots, a_{n-1}i}| \\ &= \sum_{a_0 \ldots a_{n-1}} \sum_{i \in D_n} \frac{|\gamma \cap E_{a_0, \ldots, a_{n-1}i}|}{|\gamma \cap E_{a_0 \ldots a_{n-1}}|} |\gamma \cap E_{a_0 \ldots a_{n-1}}| \\ &\leq \sum_{a_0 \ldots a_{n-1}} \sum_{i \in D_n} K\delta_{i,\max} |\gamma \cap E_{a_0 \ldots a_{n-1}}| \\ &\leq \sum_{i \in D_n} K\delta_{i,\max} \end{aligned}$$

Hence, (84) is a consequence of (83). □

Lemma 10.4. — *For any full-width K^u_α curve γ,*

$$\rho_\gamma(\gamma \cap \widetilde{Q}) = 1$$

Proof. — The curve γ cannot meet both the upper and lower boundaries of Q. For definiteness, we suppose that γ does not meet the lower boundary of Q. The other case is similar.

Then, there are constants $\alpha_1 > 0, \alpha_2 > 0$ and a C^1 diffeomorphism ϕ from Q onto a curvilinear subrectangle Q_1 of Q such that

(1) ϕ maps the upper boundary of Q onto γ and maps the lower boundary of Q onto itself.
(2) $D\phi(K^u_\alpha) \subset K^u_{\alpha_1}$
(3) $D\phi^{-1}(K^s_\alpha) \subset K^s_{\alpha_2}$.

Let $\widetilde{\gamma} = \phi^{-1}(\gamma)$ denote the upper boundary of Q.

Since a subset A of γ has full ρ_γ measure if and only if $\phi^{-1}(A)$ has full $\rho_{\widetilde{\gamma}}$ measure, it suffices to prove that

$$\rho_{\widetilde{\gamma}}(\phi^{-1}(\gamma \cap \widetilde{Q})) = 1 \tag{85}$$

Let $\widetilde{E}_i = \phi^{-1}(E_i)$,

$$\widetilde{\delta}_{i,\max} = \max_{z \in \widetilde{E}_i} \delta_z(\widetilde{E}_i)$$

and

$$\widetilde{\delta}_{i,\min} = \min_{z \in \widetilde{E}_i} \delta_z(\widetilde{E}_i)$$

The properties of ϕ guarantee that

$$-\sum_i \widetilde{\delta}_{i,\max} \log \widetilde{\delta}_{i,\min} < \infty \tag{86}$$

Now, $Q_1 \cap \widetilde{Q}$ has full Lebesgue measure in Q_1, so, $\phi^{-1}(\widetilde{Q})$ has full Lebesgue measure in Q. Thus, for almost all horizontal lines ℓ in Q, we have that $\ell \cap \phi^{-1}(\widetilde{Q})$ has full Riemannian measure.

To complete the proof of Lemma 10.4, we will prove that $\rho_\ell(\ell \cap \phi^{-1}(\widetilde{Q}))$ varies continuously with ℓ.

This is a consequence of the following.

For any $\varepsilon > 0$, there is an $N = N(\varepsilon) > 0$ such that for any horizontal full-width line segment ℓ,

$$\rho_\ell(\ell \cap \bigcup_{i \geq N} \widetilde{E}_i) < \varepsilon$$

which is, in turn, a consequence of

$$\sum_{i \geq N} \operatorname{diam}(\ell \cap \widetilde{E}_i) < \varepsilon.$$

Since the vertical boundaries of the $\widetilde{E}_i's$ are $K^s_{\alpha_2}$ curves, there is a constant $C(\alpha_2) > 0$ such that, for all $i, m(\widetilde{E}_i) > C(\alpha_2)(\widetilde{\delta}_{i,\max})^2$. So, $\widetilde{\delta}_{i,\max} \to 0$ as $i \to \infty$.

By (86) and Lemma 10.2 with $x_i = \widetilde{\delta}_{i,\max}, y_i = \widetilde{\delta}_{i,\min}$, given $\varepsilon > 0$, we can find $n_0 > 0$ such that

$$\sum_{\widetilde{\delta}_{i,\min} < 2^{-n_0}} \widetilde{\delta}_{i,\max} < \varepsilon$$

Now, take N such that $i \geq N$ implies that $\widetilde{\delta}_{i,\min} < 2^{-n_0}$.

This gives $\sum_{i \geq N} \operatorname{diam}(\ell \cap \widetilde{E}_i) \leq \sum_{\widetilde{\delta}_{i,\min} < 2^{-n_0}} \widetilde{\delta}_{i,\max} < \varepsilon$ as required . □

For future use let us observe that the argument in the last proof actually works for all K^u_α curves uniformly to prove

Lemma 10.5. — *Given $\varepsilon > 0$, there is an integer $N(\varepsilon) > 0$ such that for every K^u_α curve γ, we have*

$$\rho_\gamma(\gamma \cap (\bigcup_{i \geq N} E_i)) < \varepsilon.$$

Proof of Proposition 10.1. — We use $\nu \ll \mu$ for ν is absolutely continuous with respect to μ, and $\nu \sim \mu$ for $\nu \ll \mu$ and $\mu \ll \nu$.

Let γ, η be two C^2 full-width K^u_α curves.

In what follows we restrict our measures to $\widetilde{Q}$. Thus, when we write $\rho_\gamma(A)$ we mean $\rho_\gamma(A \cap \widetilde{Q})$.

We will show that

$$\pi_{\gamma\eta\star}\rho_\gamma \ll \rho_\eta \tag{87}$$

Once this is done, interchanging γ and η, we have $\pi_{\eta\gamma\star}\rho_\eta \ll \rho_\gamma$.

So, $\rho_\eta = \pi_{\gamma\eta\star}(\pi_{\eta\gamma\star}\rho_\eta) \ll \pi_{\gamma\eta\star}\rho_\gamma$ or $\rho_\eta \sim \pi_{\gamma\eta\star}\rho_\gamma$ as required for the proof of Proposition 10.1.

We know that $\rho_\gamma(\gamma \cap \widetilde{Q}) = 1$.

Let $B \subset \gamma \cap \widetilde{Q}$ be such that $\rho_\gamma(B) > 0$.

Let $K_1 \in (1, K_0)$.

By Lemma 10.3, for almost all $z \in \gamma$, there is an $n(z) > 0$ such that $n \geq n(z)$ implies

$$\delta_{F^n(z)}(E_{a_n(z)}) > K_1^{-n}. \tag{88}$$

From standard measure theory, we can take a compact set $A \subset B$ such that $\rho_\gamma(A) > 0$ and there is an $n(A) > 0$ such that (88) holds for all $n \geq n(A)$ and all $z \in A$.

We will show that there is a constant $K > 0$ such that

$$\rho_\eta(\pi_{\gamma\eta}(A)) \geq K^{-1}\rho_\gamma(A) \tag{89}$$

This, in turn gives $\rho_\eta(\pi_{\gamma\eta}(B)) > 0$ to prove (87).

Since $K_1 < K_0$, and $\operatorname{dist}(F^n(z), F^n(\pi_{\gamma\eta}(z))) \leq \text{const} \cdot K_0^{-n}$, we may assume that, for $z \in A$ and large n,

$$\frac{1}{2} < \frac{\operatorname{diam}(F^n(E_{a_0(z)\dots a_n(z)} \cap \gamma))}{\operatorname{diam}(F^n(E_{a_0(z)\dots a_n(z)} \cap \eta))} < 2 \tag{90}$$

For a unit vector v tangent to the curve γ at τ and a positive integer n, let us write $D_\gamma F^n(\tau)$ for $DF^n(v)$.

Now, for $z \in A$, there are points $\tau_n \in \gamma, \widetilde{\tau} \in \eta$ such that

$$\operatorname{diam}(\gamma \cap E_{a_0(z)\dots a_n(z)})\left|D_\gamma F^n(\tau_n)\right| = \operatorname{diam}(F^n(\gamma \cap E_{a_0(z)\dots a_n(z)}))$$

and

$$\operatorname{diam}(\eta \cap E_{a_0(z)\dots a_n(z)})\left|D_\eta F^n(\widetilde{\tau}_n)\right| = \operatorname{diam}(F^n(\eta \cap E_{a_0(z)\dots a_n(z)})).$$

We claim (AC-3) there is a constant $K = K(A) > 0$ such that for all $z \in A$ and $n \geq 0$

$$K^{-1} < \frac{\left|D_\gamma F^n(z)\right|}{\left|D_\eta F^n(\pi_{\gamma\eta}(z))\right|} < K. \tag{91}$$

Assuming (AC-3) for the moment, we see that there is a possibly different $K > 0$ such that, for all $n \geq 0, z \in A$, we have

(92)
$$K^{-1} < \frac{\operatorname{diam}(\gamma \cap E_{a_0(z)\ldots a_n(z)})}{\operatorname{diam}(\eta \cap E_{a_0(z)\ldots a_n(z)})} < K.$$

But, for large n, as z varies in A, the sets $\gamma \cap E_{a_0(z)\ldots a_n(z)}$ form a covering of A by small intervals and the sets $\eta \cap E_{a_0(z)\ldots a_n(z)}$ form a covering of $\pi_{\gamma\eta}(A)$ by small intervals. This gives (89) and concludes the proof of Proposition 10.1. □

Proof of (AC-3). — Let $z_n = F^n(z), w_n = F^n(\pi_{\gamma\eta}(z))$ for each $n \geq 0$. We use affine coordinates centered as z_n as in our earlier sections. We use the splitting $T_{z_n}\mathbf{R}^2 = E^u_{z_n} \oplus E^s_{z_n}$ in which $E^u_{z_n}$ contains $DF^n(v_z)$ and $E^s_{z_n}$ is tangent to $W^s_{\mathrm{loc}}(z_n)$ at z_n.

Let $\widetilde{F}$ denote the representative of F in these coordinates, and let $\widetilde{B}_n$ be the small parallelogram centered at z_n as before. We may and do assume that $K_0 > 3$.

Write v_{z_n}, v_{w_n} for the unit vectors tangent to $F^n(\gamma)$ at z_n and $F^n(\eta)$ at w_n, respectively.

Now,

(93)
$$\frac{D_\gamma F^n(z)}{D_\eta F^n(\pi_{\gamma\eta}(z))} \leq \text{const} \cdot \prod_{s=1}^{n-1} \frac{|D\widetilde{F}_{z_s}(v_{z_s})|}{|D\widetilde{F}_{w_s}(v_{w_s})|}$$

so, it suffices to show

(94)
$$\frac{|D\widetilde{F}_{z_n}(v_{z_n})|}{|D\widetilde{F}_{w_n}(v_{w_n})|} \leq \exp(a_n)$$

where

(95)
$$\sum_{n \geq 1} a_n < \text{const} \cdot \log K$$

to prove (AC-3).

Write δ_n for $\delta_{F^n(z)}(E_{a_n(z)})$.

In our affine coordinates, $v_{z_n} = \binom{1}{0}$.

Since $\operatorname{dist}(F^n(\pi_{\gamma\eta}(z)), F^n(z))$ is exponentially smaller than δ_n for large n and $|v_{w_n} - v_{z_n}| \to 0$ as $n \to \infty$, there is an $n_0 = n_0(A)$ such that $n \geq n_0$ implies $w_n \in \widetilde{B}_n$ and $v_{w_n} \in K^u_{\varepsilon_0}$. (Here $\varepsilon_0 < 1/4$ as in section 6).

Below, we use various constants C_s, $1 \leq s \leq 8$, which are independent of n and $z \in A$ and are defined in the first equation in which they appear.

As in the proof of lemma 7.1,

$$\frac{|D\widetilde{F}_{z_n}(v_{z_n})|}{|D\widetilde{F}_{w_n}(v_{w_n})|} \leq \exp(A_{1,n} + A_{2,n})$$

where

$$A_{1,n} \leq C_1 |v_{z_n} - v_{w_n}|$$

and

$$A_{2,n} \le C_2 \frac{|z_n - w_n|}{\delta_n}.$$

Since

$$\frac{|z_n - w_n|}{\delta_n} \le C_3 \left(\frac{K_1}{K_0}\right)^n$$

it suffices to show

$$|v_{w_n} - v_{z_n}| \le C_4 \left(\frac{K_1}{K_0}\right)^{n-1} \tag{96}$$

for all n to prove (94), (95), and (AC-3).

Writing $D\widetilde{F}_{w_{n-1}}(v_{w_{n-1}}) = (\xi_n, \eta_n)$ and $v_{w_n} = (u_n^1, u_n^2)$ we have

$$\begin{aligned}\xi_n &= \widetilde{F}_{1x}(w_{n-1})u_{n-1}^1 + \widetilde{F}_{1y}(w_{n-1})u_{n-1}^2\\ \eta_n &= \widetilde{F}_{2x}(w_{n-1})u_{n-1}^1 + \widetilde{F}_{2y}(w_{n-1})u_{n-1}^2\end{aligned}$$

and

$$|D\widetilde{F}_{w_{n-1}}(v_{w_{n-1}})| = |\xi_n|.$$

Thus,

$$u_n^1 = 1, \quad u_n^2 = \frac{\eta_n}{|\xi_n|}.$$

This gives

$$\begin{aligned}|v_{w_n} - v_{z_n}| &\le \frac{|\eta_n|}{|\xi_n|}\\ &\le \left(\frac{1}{1-\varepsilon_0^2}\right)\left(\frac{|\widetilde{F}_{2x}(w_{n-1})|}{|\widetilde{F}_{1x}(w_{n-1})|} + \frac{|\widetilde{F}_{2y}(w_{n-1})|}{|\widetilde{F}_{1x}(w_{n-1})|}|u_{n-1}^2|\right)\end{aligned}$$

Using $\widetilde{F}_{2x}(z_{n-1}) = 0$, we get

$$\begin{aligned}\frac{|\widetilde{F}_{2x}(w_{n-1})|}{|\widetilde{F}_{1x}(w_{n-1})|} &\le \frac{|\widetilde{F}_{2xx}(\tau)|}{|\widetilde{F}_{1x}(w_{n-1})|}|w_{n-1} - z_{n-1}| + \frac{|\widetilde{F}_{2xy}(\tau)|}{|\widetilde{F}_{1x}(w_{n-1})|}|w_{n-1} - z_{n-1}|\\ &\le C_5 \frac{|w_{n-1} - z_{n-1}|}{\delta_{n-1}}\end{aligned}$$

for suitable τ.

Analogously,

$$\frac{|\widetilde{F}_{2y}(w_{n-1})|}{|\widetilde{F}_{1x}(w_{n-1})|} \le \frac{|\widetilde{F}_{2y}(z_{n-1})|}{|\widetilde{F}_{1x}(w_{n-1})|} + C_6 \frac{|w_{n-1} - z_{n-1}|}{\delta_{n-1}}$$

which gives

$$\begin{aligned}\frac{|\eta_n|}{|\xi_n|} &\leq \frac{1}{(1-\varepsilon_0^2)K_0^2}|u_{n-1}^2| + C_7\frac{|w_{n-1}-z_{n-1}|}{\delta_{n-1}} \\ &\leq \frac{1}{(1-\varepsilon_0^2)K_0^2}|u_{n-1}^2| + C_8\left(\frac{K_1}{K_0}\right)^{n-1}\end{aligned}$$

Inductively, we assume

$$|u_{n-1}^2| \leq 2C_8\left(\frac{K_1}{K_0}\right)^{n-2}$$

and get

$$|u_n^2| = \frac{|\eta_n|}{|\xi_n|} \leq \frac{2C_8}{(1-\varepsilon_0^2)K_0^2}\left(\frac{K_1}{K_0}\right)^{n-2} + C_8\left(\frac{K_1}{K_0}\right)^{n-1}$$

Since $\frac{3}{K_0} < 1 < K_1$ and $\varepsilon_0 < 1/4$, we get $2/[(1-\varepsilon_0^2)K_0^2] < K_1/K_0$ and

$$|u_n^2| \leq 2C_8\left(\frac{K_1}{K_0}\right)^{n-1}$$

which proves (96). □

11. Construction of an SRB measure

We wish to use a construction analogous to that of Sinai in **[13]** to construct our SRB measure. There are several difficulties which appear.

(1) The family of unstable manifolds $\{W^u(z)\}$ does not form a measurable partition of the attractor Λ in Q.
(2) The underlying set Λ is not compact, so care has to exercised in the taking of limits of iterates of measures.

We will see that these problems can be handled by lifting the required construction to the symbolic space Σ, getting a measure there, compactifying, getting a limit measure which is supported on Σ, and projecting back into Q.

We have defined a continuous map π from Σ into Q as follows. For $\boldsymbol{a} \in \Sigma$ with $\boldsymbol{a} = (\dots a_{-1}a_0a_1\dots)$,

$$\{\pi(\boldsymbol{a})\} = \bigcap_{n\geq 0} E_{a_0\dots a_n} \cap f_{a_0}^{-1} S_{a_{-n}\dots a_0}$$

Let σ be the left shift automorphism on Σ. For each $\boldsymbol{a} \in \Sigma$, we have local stable and unstable sets defined by

$$W_{\mathrm{loc}}^s(\boldsymbol{a}) = \{\boldsymbol{b} : a_i = b_i, i \geq 0\}$$

$$W_{\mathrm{loc}}^u(\boldsymbol{a}) = \{\boldsymbol{b} : a_i = b_i, i \leq 0\}$$

We have the local stable and unstable sets in Q as well:

$$W^u_{\mathrm{loc}}(\pi\boldsymbol{a}) = \bigcap_{n\geq 0} f^{-1}_{a_0} S_{a_{-n}\dots a_0}$$

$$W^s_{\mathrm{loc}}(\pi\boldsymbol{a}) = \bigcap_{n\geq 0} E_{a_0\dots a_n}$$

Each $W^u_{\mathrm{loc}}(\pi\boldsymbol{a})$ is a K^u_α curve which has full width in E_{a_0}, and each $W^s_{\mathrm{loc}}(\pi\boldsymbol{a})$ is a K^s_α curve of full height in Q.

Note that if $\boldsymbol{a} = (\dots a_i \dots), \boldsymbol{b} = (\dots b_i \dots), \pi\boldsymbol{a}, \pi\boldsymbol{b} \in \widetilde{Q}$, and $a_i \neq b_i$ for some $i \geq 0$, then $W^s_{\mathrm{loc}}(\pi\boldsymbol{a}) \cap W^s_{\mathrm{loc}}(\pi\boldsymbol{b}) = \varnothing$.

Thus, the map $\pi : W^u_{\mathrm{loc}}(\boldsymbol{a}) \cap \pi^{-1}\widetilde{Q} \to W^u_{\mathrm{loc}}(\pi\boldsymbol{a}) \cap \widetilde{Q}$ is a one-to-one, continuous onto map for each $\boldsymbol{a} \in \pi^{-1}(\widetilde{Q})$. By standard results, it is a Borel isomorphism.

Recall the functions $\xi(z_1, z_2)$ and the Sinai local measures $\nu_{z_1,\gamma}$ defined at the end of section 9.

We now use them to define finite measures on the local unstable sets $W^u_{\mathrm{loc}}(\boldsymbol{a})$ in Σ.

Write $\widetilde{W}^u_{\mathrm{loc}}(\boldsymbol{a}) = W^u_{\mathrm{loc}}(\boldsymbol{a}) \cap \pi^{-1}\widetilde{Q}$.

If $\gamma = W^u_{\mathrm{loc}}(\pi\boldsymbol{a})$, then $\gamma \cap \widetilde{Q}$ has full Riemannian measure in γ, and the Borel isomorphism $\pi : \widetilde{W}^u_{\mathrm{loc}}(\boldsymbol{a}) \to W^u_{\mathrm{loc}}(\pi\boldsymbol{a}) \cap \widetilde{Q}$ allows us to transfer the Riemannian measure ρ_γ from $\gamma \cap \widetilde{Q}$ up to $\widetilde{W}^u_{\mathrm{loc}}(\boldsymbol{a})$. We call this measure $\rho_{\boldsymbol{a}}$. It clearly only depends on the non-positive indices of $\boldsymbol{a}$.

For $z, w \in \widetilde{W}^u_{\mathrm{loc}}(\boldsymbol{a})$, let

$$\overline{\xi}(z, w) = \xi(\pi z, \pi w)$$

where $\xi(\cdot, \cdot)$ is the density of the Sinai local measure defined at the end of Section 9.

Next, for $z \in \widetilde{W}^u_{\mathrm{loc}}(\boldsymbol{a})$, we define a finite measure ν_z on $\widetilde{W}^u_{\mathrm{loc}}(\boldsymbol{a})$ by

$$\nu_z(A) = \int_A \overline{\xi}(z, w)d\rho_{\boldsymbol{a}}(w)$$

These measures have the following properties

(1) For $z_1, z_2 \in \widetilde{W}^u_{\mathrm{loc}}(\boldsymbol{a})$, and $A \subset \widetilde{W}^u_{\mathrm{loc}}(\boldsymbol{a})$

$$\nu_{z_1}(A) = \overline{\xi}(z_1, z_2)\nu_{z_2}(A)$$

(2) If $A, B \subset \widetilde{W}^u_{\mathrm{loc}}(\boldsymbol{a}), z_1 \in \widetilde{W}^u_{\mathrm{loc}}(\boldsymbol{a})$, $\nu_{z_1}(B) > 0$, and $\sigma(A), \sigma(B) \subset \widetilde{W}^u_{\mathrm{loc}}(\sigma\boldsymbol{a})$, then $\nu_{\sigma z_1}(\sigma B) > 0$ and

$$\frac{\nu_{\sigma z_1}(\sigma A)}{\nu_{\sigma z_1}(\sigma B)} = \frac{\nu_{z_1}(A)}{\nu_{z_1}(B)}$$

It follows from these facts that if $\nu_{z_1}(B) > 0$, for some z_1, then $\nu_{z_2}(B) > 0$ for any z_2, and the normalized measure $\nu_B(A) = \nu_{z_1}(A \cap B)/\nu_{z_1}(B)$ is independent of the choice of $z_1 \in \widetilde{W}^u_{\mathrm{loc}}(\boldsymbol{a})$. Moreover, the normalized measures are σ-invariant in the following sense: if A and B are as in 2 above, then $\sigma_\star\nu_B(\sigma(A)) = \nu_{\sigma(B)}(\sigma(A))$. We will call the measures ν_z, *local measures* or *Sinai measures*.

For a point $\boldsymbol{a} \in \Sigma$, with local unstable set $\widetilde{W}^u_{\text{loc}}(\boldsymbol{a})$, let $\nu_{\boldsymbol{a},norm}$ be its normalized local measure. Thus,

$$\nu_{\boldsymbol{a},norm}(A) = \frac{\nu(A)}{\nu_{\boldsymbol{a}}(\widetilde{W}^u_{\text{loc}}(\boldsymbol{a}))}$$

for every $A \subset \widetilde{W}^u_{\text{loc}}(\boldsymbol{a})$.

For each $i \geq 1$, let $V_i = \{\boldsymbol{a} \in \Sigma : a_0 = i\}$, and fix a local stable set $S_i \subset V_i$. Thus, $S_i = W^s_{\text{loc}}(z_i)$ where z_i is a particular point in V_i. Let $\mathcal{M}_i$ be the partition of V_i into local unstable sets. The quotient set $V_i/\mathcal{M}_i$ is in one-to-one correspondence with S_i, so the partition $\mathcal{M}_i$ is measurable with respect to any complete Borel probability measure on V_i. Let $\mathcal{M} = \bigcup_i \mathcal{M}_i$. Since Σ is a countable disjoint union of the $V_i' s$, $\mathcal{M}$ is a measurable partition of Σ for any complete Borel probability measure.

For convenience, we will say that a Borel partition $\mathcal{M}$ is *measurable* with respect to a Borel Probability measure μ, if it is equal mod zero to a measurable Borel partition of the Borel completion of the measure μ. This allows us to discuss systems of conditional measures, etc, with respect to arbitrary measurable Borel partitions of Borel probability measures.

Now fix an element $z_0 \in \pi^{-1}(\widetilde{Q})$, and let $\widetilde{W}^u_{\text{loc}}(z_0)$ be its local unstable set. Let ν_0 be the associated normalized Sinai measure.

Theorem 11.1. — *The sequence of averages*

$$\nu_n = \frac{1}{n}\sum_{k=0}^{n-1} \sigma^k_\star \nu_0$$

converges weakly to a measure $\overline{\mu}$ on Σ which is σ-invariant, ergodic, and the conditional measures of $\overline{\mu}$ with respect to the partition $\mathcal{M}$ coincide with the normalized Sinai measures on elements of $\mathcal{M}$.

The proof will require several steps.

Let $\mathbf{N}$ be the set of positive integers, and let $\overline{\mathbf{N}} = \mathbf{N}\bigcup\{\infty\}$ be its one-point compactification. We put a metric on $\overline{\mathbf{N}}$ making it isometric to $\{0, 1, 1/2, 1/3\ldots\} \subset \mathbf{R}$ with the standard metric. Let $\overline{\Sigma} = \overline{\mathbf{N}}^{\mathbf{Z}}$ with the product topology and let $\overline{\sigma} : \overline{\Sigma} \to \overline{\Sigma}$ be the shift. The set Σ is a dense $\overline{\sigma}$-invariant subset of $\overline{\Sigma}$.

We take a subsequence $\{\nu_{n_k}\}$ of $\{\nu_n\}$ which converges to a measure $\overline{\mu}$ on $\overline{\Sigma}$.

Claim 1. — *The measure $\overline{\mu}$ is supported on Σ. That is,*

$$\overline{\mu}(\overline{\Sigma} \setminus \Sigma) = 0.$$

Proof. — A point $\boldsymbol{a} \in \overline{\Sigma} \setminus \Sigma$ has $a_i = \infty$ for some i. Fixing i, let

$$J_i = \{\boldsymbol{a} \in \overline{\Sigma} : a_i = \infty\}.$$

We will show that, given $\varepsilon > 0$, there is an open neighborhood U_i of $J_i \setminus \Sigma$ such that for all $n \geq 1 - i$,

$$\sigma_\star^n(\nu_0)(U_i) \leq \varepsilon \tag{97}$$

This will imply that $\overline{\mu}(J_i \setminus \Sigma) = 0$. Since this holds for every i, Claim 1 follows.

Let $\varepsilon_1 > 0$ be a small number to be chosen later.

From Lemma 10.5, there is an $N > 0$, such that for every K_α^u curve γ,

$$\rho_\gamma\left(\bigcup_{j \geq N} E_j\right) < \varepsilon_1 \tag{98}$$

Let $\rho_0 = \rho_{z_0}$ be the lift to $\widetilde{W}^u_{\mathrm{loc}}(z_0)$ of the Riemannian measure on $W^u_{\mathrm{loc}}(\pi z_0) \cap \widetilde{Q}$, and let $\ell_0 = \rho_0(\widetilde{W}^u_{\mathrm{loc}}(z_0))$.

By Proposition 9.1, for any $E \subset \widetilde{W}^u_{\mathrm{loc}}(z_0)$,

$$K_6^{-2}\,\frac{\rho_0(E)}{\ell_0} \leq \nu_0(E) \leq K_6^2\,\frac{\rho_0(E)}{\ell_0} \tag{99}$$

Given a non-negative itinerary $\boldsymbol{a} = (a_0 a_1 \dots)$, let

$$V_{a_0 \dots a_n} = \{\boldsymbol{b} \in \widetilde{W}^u_{\mathrm{loc}}(z_0) : b_i = a_i, i = 0, \dots, n\}$$

By Proposition 8.1, for any $n \geq 1$, if $\gamma_n = F^n(W^u_{\mathrm{loc}}(\pi z_0))$, then

$$K_4^{-1}\,\rho_{\gamma_n}(E_{a_n}) \leq \frac{\rho_0(V_{a_0 \dots a_n})}{\rho_0(V_{a_0 \dots a_{n-1}})} \leq K_4\,\rho_{\gamma_n}(E_{a_n}) \tag{100}$$

Setting $U_i = \{\boldsymbol{a} \in \overline{\Sigma} : a_i \geq N\}$, we see that (98) and (100) imply that, if $n + i \geq 1$, then

$$\rho_0(V_{a_0 \dots a_{n+i-1}} \cap \sigma^{-n} U_i) \leq K_4 \varepsilon_1 \rho_0(V_{a_0 \dots a_{n+i-1}}) \tag{101}$$

Also, $\widetilde{W}^u_{\mathrm{loc}}(z_0) \cap \sigma^{-n}(U_i)$ is the disjoint union

$$\bigsqcup_{a_0 \dots a_{n+i-1}} V_{a_0 \dots a_{n+i-1}} \cap \sigma^{-n} U_i$$

So,

$$
\begin{aligned}
(\sigma_\star^n \nu_0)(U_i) &= \nu_0(\sigma^{-n}(U_i)) \\
&= \nu_0(\widetilde{W}^u_{\mathrm{loc}}(z_0) \cap \sigma^{-n} U_i) \\
&= \sum_{a_0 \dots a_{n+i-1}} \nu_0(V_{a_0 \dots a_{n+i-1}} \cap \sigma^{-n} U_i) \\
&\le \sum_{a_0 \dots a_{n+i-1}} \frac{K_6^2}{\ell_0} \rho_0(V_{a_0 \dots a_{n+i-1}} \cap \sigma^{-n} U_i) \\
&= \sum_{a_0 \dots a_{n+i-1}} \frac{K_6^2}{\ell_0} \frac{\rho_0(V_{a_0 \dots a_{n+i-1}} \cap \sigma^{-n} U_i)}{\rho_0(V_{a_0 \dots a_{n+i-1}})} \cdot \rho_0(V_{a_0 \dots a_{n+i-1}}) \\
&\le \sum_{a_0 \dots a_{n+i-1}} \frac{K_6^2}{\ell_0} K_4 \varepsilon_1 \rho_0(V_{a_0 \dots a_{n+i-1}}) \\
&= \varepsilon_1 \frac{K_6^2}{\ell_0} K_4 \ell_0 \\
&= \varepsilon_1 K_6^2 K_4
\end{aligned}
$$

Hence, if we set $\varepsilon_1 = \varepsilon/(K_6^2 K_4)$, we get (97), and Claim 1 is proved. □

The measure $\overline{\mu}$ is clearly invariant under the shift σ.

We extend the partition $\mathcal{M}$ of Σ to $\overline{\Sigma}$ by adding the element $\overline{\Sigma} \setminus \Sigma$. We will also use the letter $\mathcal{M}$ to denote this extended partition. We let $\overline{V}_i$ denote the closure of V_i in $\overline{\Sigma}$, and let $\mathcal{M}_i$ denote the restriction of $\mathcal{M}$ to $\overline{V}_i$.

Let $\widetilde{\pi} : \overline{\Sigma} \to \overline{\Sigma}/\mathcal{M}$ be the natural projection. Let $\widetilde{\mu} = \widetilde{\pi}_\star \overline{\mu}$ be the induced measure on $\overline{\Sigma}/\mathcal{M}$.

There is a system of conditional measures $\overline{\mu}_C$ on $C \in \mathcal{M}$ defined for $\widetilde{\mu}$-almost all $C \in \mathcal{M}$.

Claim 2. — *For $\widetilde{\mu}$-almost all C, $\overline{\mu}_C = \nu_C$.*

Proof. — Let us use $\overline{A}$ for the closure of a subset $A \subset \overline{\Sigma}$ in $\overline{\Sigma}$.

Let $\phi : \overline{\Sigma} \to \mathbf{R}$ be a continuous function supported in $\overline{V}_i$ for some i.

For each $n \ge 0$, the measure $\sigma_\star^n \nu_0$ is supported on countably many $C's$ in $\mathcal{M}$, and these $C's$ are local unstable sets.

The conditional measure $(\sigma_\star^n \nu_0)_C$ is then just the restriction of $\sigma_\star^n \nu_0$ to C normalized.

But, the invariance property of quotients of the Sinai measures gives, for $A \subset C$,

$$
\begin{aligned}
\frac{(\sigma_\star^n \nu_0)(A)}{(\sigma_\star^n \nu_0)(C)} &= \frac{\nu_0(\sigma^{-n} A)}{\nu_0(\sigma^{-n} C)} \\
&= \frac{\nu_C(A)}{\nu_C(C)} = \nu_C(A)
\end{aligned}
$$

Thus,

(*) the conditional measure $(\sigma_\star^n \nu_0)_C$ is equal to the normalized Sinai measure ν_C when C is a local unstable set in $\sigma^n(\widetilde{W}^u_{\mathrm{loc}}(z_0))$.

and this implies

(**) the conditional measure $(\nu_n)_C$ equals ν_C on each local unstable set C in Σ such that $\nu_n(C) > 0$.

Let S_i be the stable set of $z_i \in V_i$. Its closure $\overline{S}_i$ is the local stable set of z_i in $\overline{\Sigma}$. This is a compact subset of $\overline{\Sigma}$ and may be identified with $\overline{V}_i/\mathcal{M}_i$.

Thus, we may think of the projection $\widetilde{\pi}$ as a map from $\overline{V}_i \to \overline{S}_i$.

Let $\overline{K} > 0$ be such that $|\phi(z)| \leq \overline{K}$ for all $z \in \overline{V}_i$.

The function

$$h(z) = \begin{cases} \int_{\widetilde{\pi}^{-1}(z)} \phi(w) d\nu_{\widetilde{\pi}^{-1}(z)}(w) & \text{for } z \in S_i \\ 0 & \text{for } z \in \overline{S}_i \setminus S_i \end{cases}$$

is then bounded and measurable and its restriction to S_i is continuous. Also, $|h(z)| \leq \overline{K}$ for all $z \in \overline{S}_i$.

Let $\overline{\mu}^i$ be the normalized restriction of $\overline{\mu}$ to $\overline{V}_i$.

We assert

$$\int_{\overline{S}_i} h(z) d(\widetilde{\pi}_\star \overline{\mu}^i) = \int_{\overline{V}_i} \phi d\overline{\mu}^i \tag{102}$$

Since, $\overline{\mu}(\overline{\Sigma} \setminus \Sigma) = 0$, this tells us that the conditional measures of $\overline{\mu}$ with respect to $\mathcal{M}$ are the ν_C as required for Claim 2.

To prove (102), we let $\varepsilon > 0$ be arbitrary, and we show

$$\left| \int_{\overline{S}_i} h(z) d(\widetilde{\pi}_\star \overline{\mu}^i) - \int_{\overline{V}_i} \phi d\overline{\mu}^i \right| \leq 5\varepsilon \overline{K} \tag{103}$$

Let $\nu^i_{n_k}$ be the normalized restriction of ν_{n_k} to $\overline{V_i}$.

Since $\overline{V}_i$ is open and closed in $\overline{\Sigma}$, we have $\nu^i_{n_k} \to \overline{\mu}^i$ as $k \to \infty$.

Since $\widetilde{\pi} : \overline{V}_i \to \overline{S}_i$ is continuous, we get $\widetilde{\pi}_\star \nu^i_{n_k} \to \widetilde{\pi}_\star \overline{\mu}^i$.

By (97), there is a compact subset $A_i \subset S_i$ such that, for large $k \geq 0$,

$$\nu^i_{n_k}(\overline{V}_i \setminus \widetilde{\pi}^{-1}(A_i)) < \varepsilon. \tag{104}$$

Since h restricted to A_i is continuous, we can use the Tietze extension theorem to find a continuous map $\overline{h} : \overline{S}_i \to \mathbf{R}$ such that $|\overline{h}(z)| \leq \overline{K}$ for all $z \in \overline{S}_i$, and $\overline{h}(z) = h(z)$ for $z \in A_i$.

Then,

$$\int_{\overline{S}_i} \overline{h} d(\widetilde{\pi}_\star \nu^i_{n_k}) \to \int_{\overline{S}_i} \overline{h} d(\widetilde{\pi}_\star \overline{\mu}^i).$$

By construction of $\overline{h}$, we then get, for large k,

$$\left| \int_{A_i} h d(\widetilde{\pi}_\star \nu^i_{n_k}) - \int_{A_i} h d(\widetilde{\pi}_\star \overline{\mu}^i) \right| \leq 3\varepsilon \overline{K}.$$

By $(**)$,

$$\int_{A_i} hd(\widetilde{\pi}_\star \nu^i_{n_k}) = \int_{\widetilde{\pi}^{-1}(A_i)} \phi d(\nu^i_{n_k}),$$

The right side of this last equality differs from $\int_{\overline{V}_i} \phi d\nu^i_{n_k}$ by no more than $\varepsilon\overline{K}$, and

$$\left| \int_{A_i} hd(\widetilde{\pi}_\star \overline{\mu}^i) - \int_{S_i} hd(\widetilde{\pi}_\star \overline{\mu}^i) \right| \leq \varepsilon \overline{K}.$$

Putting all these together gives (103) and completes the proof of Claim 2. □

Claim 3. — *$\overline{\mu}$ is ergodic.*

Proof. — This is a variant of the standard Hopf argument for geodesic flows in negatively curved Riemannian manifolds.

Let $\phi : \overline{\Sigma} \to \mathbf{R}$ be continuous. We show that $\overline{\mu}$-almost all forward time averages

$$\frac{1}{n} \sum_{k=0}^{n-1} \phi(\sigma^k z)$$

approach the same value.

Let

$$\phi_{for}(z) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \phi(\sigma^k z)$$

and

$$\phi_{bac}(z) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \phi(\sigma^{-k} z)$$

be the forward and backward limiting time averages of a point z.

From the Ergodic Theorem and standard arguments, there is a set $A_1 \subset \overline{\Sigma}$ of full $\overline{\mu}$-measure such that $z \in A_1$ implies $\phi_{for}(z), \phi_{bac}(z)$ exist and are equal. Also, since ϕ is continuous, ϕ_{for} is constant on stable sets and ϕ_{bac} is constant on unstable sets.

For each $z \in \Sigma$, let

$$W^s(z) = \bigcup_{n\geq 0} \sigma^{-n} W^s_{\mathrm{loc}}(z)$$

be the global stable set of z.

Now, $\overline{\mu}$-almost any local unstable set C is such that $\nu_C(A_1 \cap C) = 1$. Pick one such C and let $\widetilde{S}$ be the union of the global stable sets of points in $A_1 \cap C$. By the topological transitivity of the shift, the absolute continuity of the stable foliation $\mathcal{W}$ in $\widetilde{Q}$, and the fact that the push forwards by π of the conditional measures of $\overline{\mu}$ with respect to $\mathcal{M}$ are equivalent to the Riemannian measures on the local F-unstable manifolds, we get that $\nu_{C_1}(\widetilde{S}) = 1$, for every local unstable set C_1. Hence, $\overline{\mu}(\widetilde{S}) = 1$.

For any two points $z_1, z_2 \in \widetilde{S}$, there are points $w_1, w_2 \in A_1 \cap C$ such that $z_1 \in W^s(w_1), z_2 \in W^s(w_2)$.

Then, $\phi_{for}(z_1) = \phi_{for}(w_1) = \phi_{bac}(w_1) = \phi_{bac}(w_2) = \phi_{for}(w_2) = \phi_{for}(z_2)$. This proves Claim 3. □

Claim 4. — $\lim_{n\to\infty} \nu_n = \overline{\mu}$.

Proof. — Let $\overline{\mu}_1$ be another subsequential limit of the sequence $\{\nu_n\}$. Substituting $\overline{\mu}_1$ for $\overline{\mu}$ in the preceding arguments gives that $\overline{\mu}_1$ is ergodic, shift invariant and $\overline{\mu}_1(\Sigma) = 1$.

Let $G_{\overline{\mu}}$ be the set of $\overline{\mu}$-generic points, and let $G_{\overline{\mu}_1}$ be the set of $\overline{\mu}_1$-generic points. Thus, for any continuous function $\phi : \overline{\Sigma} \to \mathbf{R}$,

$$\boldsymbol{a} \in G_{\overline{\mu}} \Longrightarrow \frac{1}{n}\sum_{k=0}^{n-1} \phi(\sigma^k \boldsymbol{a}) \to \int \phi d\overline{\mu} \tag{105}$$

and

$$\boldsymbol{a} \in G_{\overline{\mu}_1} \Longrightarrow \frac{1}{n}\sum_{k=0}^{n-1} \phi(\sigma^k \boldsymbol{a}) \to \int \phi d\overline{\mu}_1 \tag{106}$$

Ergodicity implies that $\overline{\mu}(G_{\overline{\mu}} \cap \Sigma) = 1 = \overline{\mu}_1(G_{\overline{\mu}_1} \cap \Sigma) = 1$.

If we show that

$$G_{\overline{\mu}} \cap G_{\overline{\mu}_1} \cap \Sigma \neq \varnothing \tag{107}$$

then, in view of (105) and (106), we get

$$\int \phi d\overline{\mu} = \int \phi d\overline{\mu}_1$$

for all continuous ϕ, and Claim 4 follows.

For a given set $A \subset \Sigma$, let

$$W^s(A) = \bigcup_{\boldsymbol{a}\in A} W^s_{\mathrm{loc}}(\boldsymbol{a})$$

We call A *stably saturated* if $W^s(A) = A$. It is easy to see that both $G_{\overline{\mu}}$ and $G_{\overline{\mu}_1}$ are stably saturated.

The arguments in the proof of Claim 3 show that if $\overline{\mu}(A) = 1$ and $A \subset \Sigma$, then, for any local unstable set C, with Sinai measure ν_C, we have $\nu_C(W^s(A)) = 1$. In particular,

$$\nu_C(G_{\overline{\mu}} \cap \Sigma) = \nu_C(W^s(G_{\overline{\mu}} \cap \Sigma)) = 1$$

Replacing $\overline{\mu}$ by $\overline{\mu}_1$ in the arguments of Claim 3 gives $\nu_C(G_{\overline{\mu}_1} \cap \Sigma) = 1$, as well. Thus, $\nu_C(G_{\overline{\mu}} \cap G_{\overline{\mu}_1} \cap \Sigma) = 1$ for any C and (107) holds.

This completes the proof of Theorem 11.1. □

The construction of the SRB measure μ. — Let $\mu = \pi_\star \overline{\mu}$.

The measure μ is clearly an F-invariant and ergodic measure on Q.

There is a set $A \subset \widetilde{Q}$ of full μ measure consisting of μ-generic points; *i.e.*, $x \in A, \phi : Q \to \mathbf{R}$ continuous implies that $\frac{1}{n}\sum_{k=0}^{n-1}\phi(F^k x) \to \int \phi d\mu$.

Let S be the union of the local stable manifolds of points $x \in A$. Clearly, each $x \in S$ is μ-generic. We will show that $m(S) = 1$ (*i.e.* that S has full Lebesgue measure in Q) to prove that μ is SRB.

Now, $\pi^{-1}(S)$ has full $\overline{\mu}$-measure in Σ. Hence, for some (in fact, $\overline{\mu}$-almost any) local unstable set $C \subset \Sigma$, we have $\nu_C(\pi^{-1}S) = 1$. This gives $\pi_\star\nu_C(S) = 1$. But $\pi_\star\nu_C$ is equal to the normalized Sinai measure on the local unstable manifold containing $S \cap \pi(C)$, and, hence, is equivalent to the Riemannian measure restricted to $S \cap \pi(C)$. This implies that $S \cap \pi(C)$ has full Riemannian measure in $\pi(C)$. Then, the absolute continuity of $\mathcal{W}$ gives $\rho_\gamma(S) = 1$ for every K^u_α curve γ, so Fubini's theorem gives $m(S) = 1$.

12. Further ergodic properties and an entropy formula

In this section we will study properties of the natural extension of the ergodic system $(F, \widetilde{Q}, \mu)$. The first proposition identifies this natural extension with the system $(\sigma, \Sigma, \overline{\mu})$.

Proposition 12.1. — *The system $(\sigma, \Sigma, \overline{\mu})$ is isomorphic to the natural extension of the system $(F, \widetilde{Q}, \mu)$.*

Proof. — Since the map F on $\widetilde{Q}$ is not surjective, the meaning of this proposition is that there is a subset $\widetilde{Q}_1$ of $\widetilde{Q}$ of full μ-measure such that $F(\widetilde{Q}_1) = \widetilde{Q}_1$, and the system $(\sigma, \Sigma, \overline{\mu})$ is isomorphic (mod 0) to the natural extension of the system $(F, \widetilde{Q}_1, \mu)$.

Indeed, let $\widetilde{Q}_1$ be the set of points $x \in \widetilde{Q}$, such that there is a sequence $x_0, x_1, \ldots$ in $\widetilde{Q}$ with $x_0 = x$ and $F(x_{n+1}) = x_n$ for all $n \geq 0$. It is easy to see that F maps $\widetilde{Q}_1$ onto itself. To see that $\mu(\widetilde{Q}_1) = 1$, it suffices to show that $\overline{\mu}(\pi^{-1}\widetilde{Q}_1) = 1$, and, since $\pi^{-1}\widetilde{Q}$ has full $\overline{\mu}$ measure and $\overline{\mu}$ is σ-invariant, this follows from

$$\pi^{-1}(\widetilde{Q}_1) \supset \bigcap_{n\geq 0} \sigma^n(\pi^{-1}\widetilde{Q}) \tag{108}$$

To prove (108), let $\boldsymbol{a} \in \bigcap_{n\geq 0}\sigma^n(\pi^{-1}\widetilde{Q})$, and let $x_0 = \pi(\boldsymbol{a}), x_n = \pi\sigma^{-n}\boldsymbol{a}$.

Since, $\sigma^{-n}\boldsymbol{a} \in \pi^{-1}\widetilde{Q}$ for all $n \geq 0$, we have that $x_n = \pi\sigma^{-n}\boldsymbol{a} \in \widetilde{Q}$ for each such n. On the other hand, $Fx_{n+1} = F\pi\sigma^{-n-1}\boldsymbol{a} = \pi\sigma\sigma^{-n-1}\boldsymbol{a} = \pi\sigma^{-n}\boldsymbol{a} = x_n$ for all $n \geq 0$. This shows that $x_0 = \pi\boldsymbol{a} \in \widetilde{Q}_1$, so $\boldsymbol{a} \in \pi^{-1}\widetilde{Q}_1$ which is (108). So, $\widetilde{Q}_1$ is the required set.

The underlying set $\widehat{Q}$ of the natural extension of $(F, \widetilde{Q}_1, \mu)$ may be identified with the set of sequences $\overline{x} = (x_0, x_1, \ldots)$ in which each $x_n \in \widetilde{Q}_1$ and $Fx_{n+1} = x_n$ for all $n \geq 0$.

Let $\xi = \{E_1, E_2, \dots\}$ be the original collection of full height rectangles of Q. For any sequence $\overline{x} \in \widehat{Q}$, the element x_n is in the interior of a unique $E_{a_{-n}}$. Similarly, the point $F^n(x_0)$ is in the interior of a unique E_{a_n}. This enables us to define a map $\phi : \widehat{Q} \to \Sigma$ by

$$\phi(\overline{x}) = \boldsymbol{a}$$

where, for each $n \geq 0$, $x_n \in \operatorname{int} E_{a_{-n}}$ and $F^n x_0 \in \operatorname{int} E_{a_n}$. Now, the verification that the map ϕ induces an isomorphism (mod 0) between the system $(\sigma, \Sigma, \overline{\mu})$ and the natural extension of $(F, \widetilde{Q}_1, \mu)$ is straightforward, and we leave the details to the reader. □

Let ζ be the partition of Σ into the sets V_i; *i.e.*, the time 0 partition. Put $\eta = \bigvee_{i=-\infty}^{0} \sigma^i \zeta$. Then, the elements of η coincide with the local stable sets $W^s_{\mathrm{loc}}(\boldsymbol{a})$.

Moreover, we have that, mod zero, $\sigma\eta \succ \eta, \bigvee_n \sigma^n \eta$ is the point partition, and $\bigwedge_n \sigma^n \eta$ is the trivial partition $\{\Sigma\}$.

So, by definition, $(\sigma, \overline{\mu})$ is a K-system.

Then we state

Proposition 12.2. — *The map $(\sigma, \overline{\mu})$ is Bernoulli.*

We thank Dan Rudolph and Francois Ledrappier for useful conversations in connection with the proof of this proposition.

The following *Weak Markov* property was introduced in **[11]**. It was used to prove the Bernoulli property of Anosov flows (see **[4]**, **[11]**).

Let β be any partition,

$$\beta_k^l = \bigvee_{k \leq i \leq l} \sigma^i \beta.$$

Given a collection of sets P, let us use P^+ for its union.

Say that β is weak Markov (WM) if, for any $\varepsilon > 0$, there is an integer $N = N(\varepsilon)$, and collections $P = P(\varepsilon)$ of atoms of $\beta_0^\infty, M = M(\varepsilon)$ of atoms of $\beta_{-\infty}^0$ with the following properties.

(1) $\overline{\mu}(P^+) > 1 - \varepsilon$, and $\overline{\mu}(M^+) > 1 - \varepsilon$.

(2) For any $x_0^N \in \beta_0^N$,any $\overline{x}, \overline{y} \in P$ with $\overline{x} \bigcup \overline{y} \subset x_0^N$, and any subcollection A of M with $\overline{\mu}(A^+|\overline{y}) > 0$, one has

$$\left| \frac{\overline{\mu}(A^+|\bar{x})}{\overline{\mu}(A^+|\overline{y})} - 1 \right| < \varepsilon \tag{109}$$

The proof of Proposition 2.2 in **[11]** shows that a finite weak Markov partition in a K-system is weakly Bernoulli in the sense of Friedman and Ornstein **[5]**.

We will prove that the partition ζ is weak Markov. Then, arguments as in the proof of Proposition 2.2 in **[11]** give us that each of the finite partitions

$$\zeta_k = \{V_1, V_2, \ldots, V_k, \Sigma \setminus \bigcup_{i=1}^{k} V_i\}$$

is also weakly Bernoulli. This implies that each factor map on $\Sigma / \bigvee_i \sigma^i \zeta_k$ is Bernoulli. Then, Theorem 5 in **[9]** gives that $(\sigma, \overline{\mu})$ is Bernoulli.

Thus, to prove Proposition 12.2, it suffices to show that the partition ζ of Σ is weak Markov.

The corresponding ζ_0^∞ is the partition into local unstable sets $W^u_{\text{loc}}(\boldsymbol{a})$, and $\zeta^0_{-\infty}$ is the partition into local stable sets.

Given $\varepsilon > 0$, let $n_0(\varepsilon) >$ be large enough so that $\overline{\mu}(\bigcup_{|i| > n_0(\varepsilon)} V_i) < \varepsilon/4$.

For each i, let z_i be a point in V_i, and let $A_i \subset W^s_{\text{loc}}(z_i), B_i \subset W^u_{\text{loc}}(z_i)$ be compact subsets so that the sets

$$D_i^u = \bigcup_{z \in A_i} W^u_{\text{loc}}(z), \quad D_i^s = \bigcup_{w \in B_i} W^s_{\text{loc}}(w)$$

satisfy

$$\overline{\mu}(V_i \setminus D_i^u) < \frac{\varepsilon}{2^{|i|+2}}, \quad \overline{\mu}(V_i \setminus D_i^s) < \frac{\varepsilon}{2^{|i|+2}}$$

Then, set $P = P(\varepsilon) = \bigcup_{|i| \le n_0(\varepsilon)} D_i^u, M = M(\varepsilon) = \bigcup_{|i| \le n_0(\varepsilon)} D_i^s$.

We have that $\overline{\mu}(P^+) > 1 - \varepsilon, \overline{\mu}(M^+) > 1 - \varepsilon$. Also, the set $Z_\varepsilon = \bigcup_{|i| \le n_0(\varepsilon)} D_i^s \cap D_i^u$ is compact.

Let $\boldsymbol{a}, \boldsymbol{b} \in V_i$, and let $W^u_{\text{loc}}(\boldsymbol{a}), W^u_{\text{loc}}(\boldsymbol{b})$ be their local unstable sets. Let $\pi_{\boldsymbol{a},\boldsymbol{b}}$ be the projection from $W^u_{\text{loc}}(\boldsymbol{a})$ to $W^u_{\text{loc}}(\boldsymbol{b})$ along the local stable sets in V_i. As $\boldsymbol{a}$ approaches $\boldsymbol{b}$ in Σ, the maps $\pi_{\boldsymbol{a},\boldsymbol{b}}$ approach the identity $\pi_{\boldsymbol{b},\boldsymbol{b}}$ and the measures $\rho_{\boldsymbol{a}}$ approach $\rho_{\boldsymbol{b}}$. Also, the densities $\overline{\xi}(\boldsymbol{a}, \boldsymbol{b})$ vary continuously with $\boldsymbol{a}, \boldsymbol{b}$ in V_i. On the compact set Z_ε the convergence and continuity above are uniform. Further, each $\overline{x} \in P$ is one of the sets $W^u_{\text{loc}}(\boldsymbol{a})$ and the conditional measure $\overline{\mu}(\cdot|\overline{x})$ is just the Sinai measure $\nu_{\overline{x}}$. If $\boldsymbol{a} \in \overline{x}, \boldsymbol{b} \in \overline{y}$ and $\overline{x} \bigcup \overline{y} \subset x_0^N \in \zeta_0^N$, then $a_j = b_j$ for $-N \le j \le 0$. For N large, the measures $\overline{\mu}(\cdot|\overline{x}), \overline{\mu}(\cdot|\overline{y})$ have densities whose quotient is closer to 1 than ε. These statements imply the Weak Markov property above. This completes the proof of Proposition 12.2.

Entropy formula. — It follows from our constructions that the measures of V_i satisfy

$$c_1 \delta_{i,\min} < \overline{\mu}(V_i) < c_2 \delta_{i,\max} \tag{110}$$

for some positive constants c_1, c_2.

Since the partition ζ generates, we get

$$h_{\overline{\mu}}(\sigma) = \inf_n \frac{1}{n} H_{\overline{\mu}}(\zeta^0_{-n+1}) \le H_{\overline{\mu}}(\zeta)$$

From condition G3 and (110), the last term is finite.

For $\boldsymbol{a} \in \Sigma$, and $n \geq 1$, let $V_{a_0 \dots a_{n-1}} = V_{a_0} \cap \sigma^{-1} V_{a_1} \cap \dots \cap \sigma^{-n+1} V_{a_{n-1}}$, and let $E_{a_0 \dots a_{n-1}}$ be the full height subpost of Q defined in section 2.

Since σ is ergodic with respect to $\overline{\mu}$, the Shannon-Breiman-Macmillan theorem gives a set A with $\overline{\mu}(A) = 1$ such that $\boldsymbol{a} \in A$, implies

$$-\frac{1}{n} \log \overline{\mu}(V_{a_0 \dots a_{n-1}}) \to h_{\overline{\mu}}(\sigma) \tag{111}$$

Using that the conditional measures of $\overline{\mu}$ along local unstable sets have bounded densities relative to the measures $\rho_{\boldsymbol{a}}$, we see that there is a constant $K > 0$ such that, for $\boldsymbol{a} \in \Sigma$,

$$K^{-1} \min_{z \in W^s_{\mathrm{loc}}(\pi \boldsymbol{a})} \operatorname{diam}(\ell_z \cap E_{a_0 \dots a_{n-1}}) \leq \overline{\mu}(V_{a_0 \dots a_{n-1}}) \tag{112}$$

and

$$\overline{\mu}(V_{a_0 \dots a_{n-1}}) \leq K \max_{z \in W^s_{\mathrm{loc}}(\pi \boldsymbol{a})} \operatorname{diam}(\ell_z \cap E_{a_0 \dots a_{n-1}}). \tag{113}$$

This and Proposition 8.1 imply that, if $F^n(z) = (F_1^n(z), F_2^n(z))$, then there are a constant $K_1(\boldsymbol{a}) > 0$ and points $u_{n,1}(\boldsymbol{a}), u_{n,2}(\boldsymbol{a}) \in W^s_{\mathrm{loc}}(\pi \boldsymbol{a})$ such that

$$\overline{\mu}(V_{a_0 \dots a_{n-1}}) \left| F^n_{1x}(u_{n,1}(\boldsymbol{a})) \right| \leq K_1(\boldsymbol{a}) \tag{114}$$

and

$$K_1(\boldsymbol{a})^{-1} \leq \overline{\mu}(V_{a_0 \dots a_{n-1}}) \left| F^n_{1x}(u_{n,2}(\boldsymbol{a})) \right| \tag{115}$$

By arguments like those in the proof of estimate (91), for $\overline{\mu}$ almost all $\boldsymbol{a}$, there is a constant $K_1(\boldsymbol{a}) > 0$ such that, for $z, w \in W^s_{\mathrm{loc}}(\pi \boldsymbol{a})$, $n \geq 1$,

$$K_2(\boldsymbol{a})^{-1} \leq \frac{\left| F^n_{1x}(z) \right|}{\left| F^n_{1x}(w) \right|} \leq K_2(\boldsymbol{a}) \tag{116}$$

From (114), (115), (116) we get the existence of a constant $K_3(\boldsymbol{a}) >$, such that

$$K_3(\boldsymbol{a})^{-1} < \overline{\mu}(V_{a_0 \dots a_{n-1}}) \left| F^n_{1x}(\pi \boldsymbol{a}) \right| < K_3(\boldsymbol{a}). \tag{117}$$

Thus there is a set A with $\overline{\mu}(A) = 1$, so that if $\boldsymbol{a} \in A$, then

$$\lim_{n \to \infty} \frac{1}{n} \log \left| F^n_{1x}(\pi \boldsymbol{a}) \right| = h_{\overline{\mu}}(\sigma) \tag{118}$$

Since σ is isomorphic to the natural extension of F, we have $h_\mu(F) = h_{\overline{\mu}}(\sigma)$.

Letting $A_1 = \pi(A)$, then, for μ-almost all z in A_1, we have

$$\lim_{n \to \infty} \frac{1}{n} \log \left| F^n_{1x}(z) \right| = h_\mu(F) \tag{119}$$

Taking S to be the union of the stable manifolds of points in A_1, we get that S has full Lebesgue measure in Q and (119) holds for all $z \in S$.

But, for $z \in \widetilde{Q}$, we have

$$\left| F^n_{1x}(z) \right| = \max(\left| F^n_{1x}(z) \right|, \left| F^n_{2x}(z) \right|) \leq \left| DF^n(z) \right| \leq (1 + \alpha) \left| F^n_{1x}(z) \right|.$$

So, we have proved formula (6) and completed the proof of Theorem 3.1.

As a final remark, if $v = (v_1, v_2)$ is a unit vector in K^u_α, then

$$\begin{aligned}(1-\alpha^2)\left|F^n_{1x}\right| &\leq \left|F^n_{1x}v_1 + F^n_{1y}v_2\right| \\ &= \left|DF^n(z)v\right| \\ &\leq \left|F^n_{1x}\right|(1+\alpha^2)\end{aligned}$$

That is, for certain constants C_1, C_2, we have

$$C_1\left|F^n_{1x}\right| \leq \left|DF^n(z)(v)\right| \leq C_2\left|F^n_{1x}\right|$$

which, together with (119), implies formula (7).

References

[1] R. Abraham and J. Robbin. *Transversal Mappings and Flows.* W. A. Benjamin, New York, N.Y., 1967.

[2] R. Adler. Afterword to R. Bowen, Invariant measures for Markov maps of the interval. *Comm. Math. Phys.*, 69(1):1–17, 1979.

[3] V. M. Alekseev. Quasi-random dynamical systems, I. *Math. of the USSR, Sbornik*, 5(1):73–128, 1968.

[4] L. Bunimovich. On a class of special flows. *Math. USSR IZV.*, 8:219–232, 1974.

[5] N. Friedman and D. Ornstein. On the isomorphism of weak Bernoulli transformations. *Advances in Math.*, 5:365–394, 1970.

[6] Jack K. Hale. *Ordinary Differential Equations, 2nd Ed.* Krieger Publishing Co., 1980.

[7] M. Hirsch and C. Pugh. Stable manifolds and hyperbolic sets. *Proc. AMS Symp. Pure Math.*, 14, 1970.

[8] M. V. Jakobson and S. E. Newhouse. A two dimensional version of the folklore theorem. *American Math. Soc. Translations, Series 2*, 171:89–105, 1996.

[9] Donald S. Ornstein. *Ergodic Theory, Randomness, and Dynamical Systems.* Number 5 in Yale Mathematical Monographs. Yale University Press, New Haven, Conn., 1975.

[10] C. Pugh and M. Shub. Ergodic attractors. *Transactions AMS*, 312(1):1–54, 1989.

[11] M. Ratner. Anosov flows with Gibbs measures are also Bernoullian. *Israel J. Math.*, 17:380–391, 1974.

[12] M. Shub. *Global Stability of Dynamical Systems.* Springer-Verlag, 1987.

[13] Ya.G. Sinai. *Topics in Ergodic Theory.* Number 44 in Princeton Mathematical Series. Princeton University Press, 1994.

[14] Peter Walters. Invariant measures and equilibrium states for some mappings which expand distances. *Trans. AMS*, 236:121–153, 1978.

M. Jakobson, Mathematics Department, University of Maryland, College Park, MD 20742, USA
E-mail : mvy@math.und.edu

S. Newhouse, Mathematics Department, Michigan State University, E. Lansing, MI 48824-1027, USA • *E-mail :* sen@math.msu.edu

GENADI LEVIN

SEBASTIAN VAN STRIEN

Total disconnectedness of Julia sets and absence of invariant linefields for real polynomials

Astérisque, tome 261 (2000), p. 161-172

<http://www.numdam.org/item?id=AST_2000__261__161_0>

Astérisque
261, 2000, p. 161–172

TOTAL DISCONNECTEDNESS OF JULIA SETS AND ABSENCE OF INVARIANT LINEFIELDS FOR REAL POLYNOMIALS

by

Genadi Levin & Sebastian van Strien

Abstract. — In this paper we shall consider real polynomials with one (possibly degenerate) non-escaping critical (folding) point. Necessary and sufficient conditions are given for the total disconnectedness of the Julia set of such polynomials. Also we prove that the Julia sets of such polynomials do not carry invariant linefields. In the real case, this generalises the results by Branner and Hubbard for cubic polynomials and by McMullen on absence of invariant linefields.

1. Introduction

In a paper by Branner and Hubbard [**BH**], cubic polynomials were considered, and the problem was solved when the Julia set of such a polynomial is totally disconnected (for the history of this problem see [**BH**], Ch. 5). In the same paper, the question was raised whether this result could be extended to polynomials of higher degrees. The method and results of [**BH**] hold for polynomials P of higher degrees with all but one critical points escaping to infinity, under the condition that the unique non-escaping critical point c is *simple*: $P'(c) = 0$, $P''(c) \neq 0$, see [**BH**], Ch. 12.

On the other hand, if the non-escaping critical point c is *multiple*, *i.e.*,

$$P''(c) = \cdots = P^{(\ell-1)}(c) = 0, \quad P^{(\ell)}(c) \neq 0,$$

for some $\ell > 2$, the method of [**BH**] breaks down (see [**Dou1**] for a discussion on this). The positive integer ℓ is called the *multiplicity*, or *local degree* of the critical point c of the polynomial P.

In this paper we shall prove

Theorem 1.1. — *Let P be a polynomial with real coefficients, such that one (maybe multiple) critical point c of P of even multiplicity ℓ has a bounded orbit, and all other critical points escape to infinity. Then*

1991 ***Mathematics Subject Classification.*** — Primary 58F03; Secondary 30F60, 58F23, 32G05.
Key words and phrases. — Julia sets, local connectivity, total disconnectedness, complex bounds.

– *the filled Julia set of* P:

$$K(P) = \{z : \{P^n(z)\}_{n=0}^{\infty} \text{ is bounded}\}$$

is totally disconnected if and only if the connected component of the real trace $K(P) \cap \mathbb{R}$ *of the filled Julia set containing the critical point* c*, is equal to a point*

– *the Julia set* $J(P) = \partial K(P)$ *carries no measurable invariant linefields.*

Remark 1.1. — For the case that the multiplicity is odd, see **[LS2]**.

Remark 1.2. — The map $P\colon \mathbb{R} \to \mathbb{R}$ does not have a wandering interval (on the real line) with bounded orbit **[MS]**. Hence, the condition: "the connected component of the real trace $K(P) \cap \mathbb{R}$ of the filled Julia set which contains the critical point c, is equal to a point" is equivalent to one of the following conditions:

– the component of the filled Julia set $K(P)$ containing the non-escaping critical point, is non-periodic;
– P does not have an attracting or neutral periodic orbit, and is not renormalizable on the real line (*i.e.*, there is no interval I around $c = 0$, such that $P^i(I) \cap P^j(I) = \varnothing$ for $0 \le i < j \le q - 1$ and $P^q(I) \subset I$);
– the intersections of the critical puzzle pieces with the real line shrink to the point c (see the next Section).

Remark 1.3. — We allow escaping critical points to be non real. Of course, since P is real, the orbit of the unique non-escaping critical point is real. The theorem holds in particular for maps of the form $f(z) = z^\ell + c_1$ when c_1 is real.

The second part of the Theorem extends the main result of McMullen in **[McM]**. As usual, we say that the Julia set $J(P)$ of P carries a *measurable invariant line field* if there exists a measurable subset E of the Julia set of P and a measurable map which associates to Lebesgue almost every $x \in E$ a line $l(x)$ through x which is P-invariant in the sense that $l(P(x)) = DP(x)\, l(x)$. (So the absence of linefields is obvious if the Julia set has zero Lebesgue measure.) The absence of invariant linefields was proved by McMullen for all maps of the form $P(z) = z^\ell + c_1$, ℓ is even and c_1 is real, which are infinitely renormalizable. If P is quadratic (*i.e.*, $\ell = 2$) and only finitely often renormalizable then this holds because then the corresponding parameter c_1 lies at the boundary of the Mandelbrot set, see **[Y]**, **[H]**. (Actually, the result of Yoccoz is much stronger: local connectivity of the boundary of the Mandelbrot set at such points).

The (non-)existence of the invariant linefields is strongly related to the Density of Hyperbolicity Conjecture, see **[MSS]**. It follows from the second part of the Theorem, because of Theorem E of **[MSS]**, that for any polynomial P as in the Theorem, there exists another (maybe complex) polynomial Q of the same degree and with the same multiplicity ℓ of the non-escaping critical point $c = 0$, which is as close to P as we wish and such that Q is hyperbolic: every critical point of Q tends either to infinity, or to an attracting periodic orbit. (One can assume that $P(z) = z^\ell \cdot p(z) + t$, where $p(z) = z^m + \cdots + p_m$ is a monic polynomial of the degree $m \geq 0$, and $p_m \neq 0$. Then the polynomial Q as above is of the same form, and P and Q are considered as points of the space $\mathbb{C}^m \times \mathbb{C}$). In fact, for $\ell = 2$ a much stronger statement is true, since the density of hyperbolicity within real quadratic maps implies that one chooses Q as above real.

The proof of the Theorem is postponed until Section 3, and is based on Propositions 2.1- 2.4 of the next section.

Propositions 2.1 and 2.3 give sufficient conditions for the total disconnectedness of Julia set and absence of invariant linefields. They could be applied to complex polynomials as well. Nevertheless, we can only show that this condition is satisfied in the case considered in the theorem: see Proposition 2.4 and Section 3.

A similar problem exists for odd multiplicities. If P is a polynomial as in the Theorem, but with ℓ an odd number (say, cubic one), then the Theorem is easy if there are no other critical points (*i.e.*, $P(z) = z^\ell + c_1$, where c_1 is real and ℓ is odd), because then the map is monotone on the real line. On the other hand, if other (escaping) critical points exist, we can still apply Propositions 2.1, but the implementation of it (a statement like Proposition 2.4) uses different methods, see **[Le]** and **[LS2]**.

Before giving the proofs, let us make a remark about the non-minimal case (*i.e.*, when the postcritical set $\omega(c)$ contains c, but the system restricted to $\omega(c)$ is not minimal). Such system is relatively simple when it is *real* (see Proposition 3.2 of **[LS1]** and Proposition 2.5 below, or see **[Ly]**). But this is definitely not the case for complex parameters:

Remark 1.4. — In each of Douady's examples of an infinitely renormalizable quadratic map f with a non-locally connected Julia set, the postcritical set is non-minimal. Indeed, according **[P-M]** there exists an invariant Cantor set on which the map is injective. If this Cantor set does not intersect the postcritical set, then according to a well-known result of Mañé **[Ma]** the map is expanding on this set. This is impossible since f is injective on this set, see **[Dou2]**.

Acknowledgements. — The authors thank the referee for carefully reading the manuscript and particularly for pointing out that the case of odd multiplicity is non-trivial, see **[Le]** and **[LS2]**.

2. Associated mappings and complex bounds

Let $G\colon \cup_{i=0}^{i_0} \Omega^i \to \Omega$ be an *ℓ-polynomial-like mapping.* As in **[DH]**, **[LM]**, **[LS1]** this means that all Ω^i are open topological discs with pairwise disjoint closures which are compactly contained in the topological disc Ω, the map $G\colon \Omega^0 \to \Omega$ is ℓ-to-one holomorphic covering with a unique critical point $c = 0 \in \Omega^0$, and that each other map $G\colon \Omega^i \to \Omega$ is a conformal isomorphism. Its *filled Julia set* is defined by

$$K(G) = \{z : G^n(z) \in \cup_{i=0}^{i_0} \Omega^i, n = 0, 1, \ldots\}.$$

The boundary $\partial K(G)$ is called the *Julia set* $J(G)$ of G.

The *puzzle* (see **[BH]**) of the map G is said to be the set of connected components of all preimages $G^{-k}(\Omega)$, $k = 0, 1, \ldots$. A *piece* of level $k \geq 0$ is a connected component of $G^{-k}(\Omega)$. A piece is *critical* if it contains the critical point c of G. We call another ℓ-polynomial-like mapping $G'\colon \cup_{i=0}^{i'_0} \Omega'^i \to \Omega'$ *associated to* or *induced by* G if G' restricted to each Ω'^i is some iterate $G^{j(i)}$ of G. We also call G *real* iff all topological discs Ω^i, Ω are symmetric w.r.t. the real axis, and $\overline{G(z)} = G(\overline{z})$, for any $z \in \cup_{i=0}^{i_0} \Omega^i$. In particular, this implies that the postcritical set of the unique critical point $c = 0 \in \Omega_0$ is real.

Proposition 2.1. — *Fix a real ℓ-polynomial-like mapping G. Assume there exists an infinite sequence $G(j)\colon \cup_i \Omega^i(j) \to \Omega(j)$ of real ℓ-polynomial-like mappings associated to the mapping G with $\omega(c)$ minimal such that the critical point $c = 0 \in \Omega^0(j)$ of $G(j)$ does not escape the domain of $G(j)$ under iterations of $G(j)$. Assume moreover that*

(1) *each $\Omega^i(j) \cap \mathbb{R}$ coincides with the intersection of some piece of G with the real line;*
(2) *when $G^i(c) \in \Omega^0(j)$ then $G^i(c)$ is an iterate of c under $G(j)$ (we call this* the first return condition for G on $\Omega^0(j)$*);*
(3) *the modulus of the annuli $\Omega(j) \setminus \Omega^0(j)$ is uniformly bounded away from zero by a constant $m > 0$ which does not depend on j;*
(4) *the diameter of $\Omega(j)$ tends to zero as $j \to \infty$.*

Then the filled Julia set of G is totally disconnected.

Remark 2.1. — Conditions (1) and (4) obviously imply the third condition of Remark 1.2: the traces of the critical pieces on the real line shrink to the point.

Remark 2.2. — The proposition also holds for a complex (*i.e.*, not real) map G, if one replaces (1) by the following condition:

for each j, one can find another ℓ-polynomial-like mapping $R(j)$ having as its range a critical piece $P(j)$ of G, so that the two mappings $G(j)$ and $R(j)$ satisfy the conditions of Proposition 2.1 from **[LS1]**.

Proof. — Let us first observe that the first return condition also holds for the first return map of $G(j)$ to $\Omega^0(j)$. Indeed, consider the first return map of $G(j)$ to $\Omega^0(j)$ along the iterates of the critical point. This is again a real ℓ-polynomial-like mapping $\tilde{G}(j)\colon \cup_i \tilde{\Omega}^i(j) \to \Omega^0(j)$. Obviously, if $G^i(c) \in \Omega^0(j)$ then $G^i(c)$ is an iterate of c under $\tilde{G}(j)$. In addition, the modulus of $\Omega^0(j) \setminus \tilde{\Omega}^0(j)$ is greater or equal to $1/\ell$ mod $(\Omega(j) \setminus \Omega^0(j))$. So we may replace $G(j)$ by the first return map to $\Omega^0(j)$ and therefore in the remainder of the proof we can and will assume that the first return condition even holds on $\Omega(j)$ (by renaming everything). Note, that this condition is crucial in the proof of Claim 1 below.

Fix j and let $P(j)$ be the open piece of G based on $\Omega(j) \cap \mathbb{R}$. Let $P^i(j)$ be a piece of G based on $\Omega^i(j) \cap \mathbb{R}$. Since $G(j)$ is associated to G, each restriction of $G(j)$ to $\Omega^i(j) \cap \mathbb{R}$ is an iterate of G. This gives another real ℓ-polynomial-like mapping $G'(j)\colon \cup_i P^i(j) \to P(j)$, which is also associated with G and coincides with $G(j)$ on the real line.

Now we are going to use the following statement from **[LS1]** (called Proposition 2.3 in that paper).

Proposition 2.2. — *Let*

$$G_1\colon \Omega_1^0 \cup \Omega_1^1 \cup \cdots \cup \Omega_1^r \longrightarrow \Omega_1, \qquad G_2\colon \Omega_2^0 \cup \Omega_2^1 \cup \cdots \cup \Omega_2^r \longrightarrow \Omega_2$$

be two real ℓ-polynomial-like mappings, with the common critical point $c \in \Omega_1^0 \cap \Omega_2^0$, and assume that the following conditions hold:

(1) *$G_1(z) = G_2(z)$ whenever $z \in \cup_{i=0}^r \Omega_1^i \cap \Omega_2^i$.*

(2) *Denoting $I_k = \Omega_k \cap \mathbb{R}$ and $I_k^i = \Omega_k^i \cap \mathbb{R}$, one has $I_2 \subseteq I_1$, $I_2^i \subseteq I_1^i$.*

Under these conditions, the Julia sets of G_1 and G_2 coincide. If, additionally, the point c lies in the Julia set of G_1 (and, hence of G_2), then there exists a component of a preimage $G_1^{-n}(\Omega_1)$, which contains c and is contained in Ω_2.

Applying this proposition, we conclude that the Julia sets of $G(j)$ and $G'(j)$ coincide for every j. On the other hand, the Julia set is the intersection of the full preimages of the range. Hence, there exists a large integer N such that the full preimage $G'(j)^{-N}(P(j))$ is inside the domain of definition $\cup_i \Omega^i(j)$ of $G(j)$. Note that $G'(j)^{-N}(P(j))$ consists of finitely many (open) pieces of G. In particular, since the critical point c does not escape under the map $G(j)$, we obtain, that there is a critical piece of G inside the domain $\Omega(j)$. As the diameters of $\Omega(j)$ tend to zero, the intersections of the critical pieces is the point.

Let us consider the pieces of $G'(j)^{-N}(P(j))$ inside the central domain $\Omega^0(j)$, *i.e.*,

$$P'(j) = G'(j)^{-N}(P(j)) \cap \Omega^0(j).$$

If z is in the Julia set but $\omega(z)$ does not hit some critical piece then $\{z\}$ is a component of $J(G)$ by **[BH]**. So choose a point z from the Julia set of G so that the forward orbit of z hits every critical piece. Then there exists a *minimal* $K = K(j)$ such that $G^K(z) \in P'(j)$. In particular, the point $G^K(z)$ belongs to one of the pieces inside $\Omega^0(j)$. Let B_j be the branch of G^{-K} which maps a neighbourhood of $G^K(z)$ to a neighbourhood of z.

Claim 1. — *The map B_j extends to a holomorphic map on $\Omega(j)$.*

Proof of the claim. — Assume the contrary. We then get for some minimal $r < K$ that $G^{-r}(j)(\Omega(j))$ (along the same orbit) meets the critical value $c_1 = G(c)$.

This means that the branch G^{-r} follows the points $c_{r+1} = G^r(c_1) \in \Omega$, $c_r = G^{r-1}(c_1), \dots, c_2 = f(c_1), c_1$. Among these iterates of c_1, let us mark all those $c_{j_1}, c_{j_2}, \dots, c_{j_m}$, where $j_1 < j_2 < \dots < j_m$, which hit the domain $\Omega(j)$. Because of *the first return condition* there exist integers $k(1) < k(2) < \dots$ such that $c_{j_1} = G^{k(1)}(c)$, $c_{j_2} = G^{k(2)-k(1)}(c_{j_1}) = G^{k(2)}(c), \dots, c_{j_m} = G^{k(m)}(c)$. It follows, that

$$G^{-r} = G^{-(s-1)} \circ G^{-(k(m)-1)},$$

where $G^{-(s-1)}$ is the branch corresponding to the restriction of $G(j)$ on $\Omega^0(j)$ (so $G(j)|\Omega^0(j) = G^{s-1} \circ G$). Hence,

$$G^{-(r+1)}(\Omega(j)) \subset G^{-k(m)}(j)(\Omega(j)) \subset \Omega^0(j)$$

and

$$G^{K-r-1}(z) \in G^{-(r+1)}(P'(j)) = (G(j)|_{\Omega^0(j)})^{-1} \circ G(j)^{-k(m)+1}(P'(j)) \subset P'(j).$$

This contradicts the minimality of K and proves the claim. □

Let $P_j(z) = B_j(P'(j))$. We want to show that the Euclidean diameters of $P_j(z)$ tend to zero as $j \to \infty$. For this, let us consider the domain $M'_j \subset \Omega(j)$ bounded by the core curve of the annulus $\Omega(j) \setminus \Omega^0(j)$. Then

$$\frac{\max_{y \in \partial M'_j} |G^K(z) - y|}{\min_{y \in \partial M'_j} |G^K(z) - y|} \le C$$

for all $j = 1, 2, \dots$ where C only depends on the uniform bound for the moduli of $\Omega(j) \setminus \Omega^0(j)$ (see *e.g.* **[McM]**). Define $E_j = B_j(M'_j)$. Since the modulus of the annulus $\Omega(j) \setminus M'_j$ is half the modulus of $\Omega(j) \setminus \Omega^0(j)$, by the Koebe Distortion Theorem,

$$\frac{\max_{y \in \partial E_j} |z - y|}{\min_{y \in \partial E_j} |z - y|} \le C_1, \quad j = 1, 2, \dots.$$

If we assume by contradiction that $\operatorname{diam}(P_j(z)) \ge d > 0$ for $j = 1, 2, \dots$, then

$$\min_{y \in \partial E_j} |z - y| \ge d/2C_1 = r > 0,$$

i.e., the disc $D_z(r) \subset E_j$. Hence, $G^K(D_z(r)) \subset M'_j$, for $j \to \infty$ and $K = K(j) \to \infty$. This contradicts the assumption that z is in the boundary of the filled Julia set of G. Thus, $\bigcap_{j>0} P_j(z) = \{z\}$ and so $\{z\}$ is a component of $J(G)$. □

We shall derive the absence of invariant linefields from the following statement, which is very similar to the previous one.

Proposition 2.3. — *Under the conditions of Proposition 2.1, the Julia set of G carries no invariant linefield.*

Remark 2.3. — In fact, the proof of this statement is a purely complex one, and, therefore, holds for every (complex) map G, such that $\omega(c)$ is minimal and conditions (2)-(4) of Proposition 2.1 are satisfied.

Proof of Proposition 2.3. — We will follow the main idea of Theorem 10.3 in [**McM**] (absence of invariant line field for infinitely renormalizable quadratic polynomial with complex apriori bounds). Let us first outline the differences with the proof in [**McM**].

First, our renormalizations are ℓ-(*i.e.*, generalised) polynomial-like maps; hence their Julia sets are not connected, and the number of the components in the domain of definition of these maps can increase. To overcome this, we consider the dynamics only on the central domains. The second problem is that the central domain $\Omega_0(j)$ can become smaller and smaller compared to the range $\Omega(j)$, so that the range of the limit dynamics can be the whole plane (after rescaling $\Omega_0(j)$ to a definite size). To avoid this, we shall rescale $\Omega_0(j)$ and $\Omega(j)$ by different factors. The third difference is that in our setting the critical value of the 'limit dynamics' can escape to the boundary of the range. For this, we extend the dynamics passing to the first return maps. Fourthly, in [**McM**] a contradiction against the existence of a measurable invariant linefield (defined on a set E) is obtained through a univalent map from the range of the renormalization to a neighbourhood of a point of density of the set E. In our setting we cannot argue like that, and instead we use two consecutive first return maps. So let us prove the proposition:

1. Fix for a moment a mapping $G = G(j)$, and consider the first return to its central domain. We obtain in this way another mapping $G' = G'(j)$ (we drop the index j). Its central branch $G_0'\colon \Omega_0' \to \Omega_0$ extends to an ℓ-covering $G_0'\colon \widetilde{\Omega}_0' \to \Omega$ (we keep the same notation for the extension), where $\widetilde{\Omega}_0' \subset \Omega_0$, and

$$\operatorname{mod}\,(\Omega_0 \setminus \Omega_0') > m_0 = \frac{m}{\ell},$$

where $m > 0$ is the number introduced in Proposition 2.1. Also, as in the proof of the previous proposition, the condition 2 holds for the new map. Now, let us *replace the initial sequence of mappings $G(j)$ by the sequence of the first return maps $G'(j)$ replacing the notations as well* (so forget about the initial sequence).

2. Let us assume by contradiction that G admits an invariant line field, *i.e.*, there is a G-invariant Beltrami differential μ supported on $J(G)$. Then we can fix a point $x \in J(G)$ of almost continuity of μ. We may assume from the beginning that $\omega_G(x)$ contains c since the set of the points of $J(G)$ without this property has Lebesgue measure zero (this follows from the well-known fact that for almost all point x one has that $\omega(x) \subset \omega(c)$, see [**Ly2**], [**McM**], and from the minimality of $\omega(c)$).

3. Let us consider a mapping $G(j)$ so that the point x is outside of the range $\Omega(j)$. Then there is minimal $k = k(j) > 0$ so that $y(j) = G^k(x) \in \Omega_0(j)$. Due to the condition 2 of Proposition 2.1, we can apply *Claim 1* from the proof of the previous proposition: there exists a branch F_j of G^{-k} univalent in the range $\Omega(j)$ of the map $G(j)$ and such that $F_j(y(j)) = x$. Let us consider a domain M_j (containing $\Omega_0(j)$ and contained in $\Omega(j)$) bounded by the core curve of the annulus $\Omega(j) \setminus \Omega_0(j)$, so that $\operatorname{mod}\,(\Omega(j) \setminus M_j) = \operatorname{mod}\,(M_j \setminus \Omega_0(j)) > m_0/2$. By the Koebe Distortion Theorem, the image $F_j(M_j)$ is roughly a disc around x (it means that it contains a disc centred

at x of radius r and it is contained in a disc centred at x of radius R so that R/r is less than a constant depending of m_0 only). Since x belongs to the Julia set, the domains $F_j(M_j)$ shrink to the point x as $j \to \infty$.

4. Let us consider the first return map $G'(j)$ of the mapping $G(j)$ to its central domain $\Omega_0(j)$ (as we did in Step 1 with respect to the old $G(j)$). By Step 1, the central branch $G'_0(j)\colon \Omega'_0(j) \to \Omega_0(j)$ extends to an ℓ-covering onto $\Omega(j)$, and again any iterate of c by G entering the $\Omega_0(j)$ is an iterate of c under the first return map $G'(j)$. Let $m = m(j) > 0$ be minimal so that $z(j) = G^m(x) \in \Omega'_0(j)$. By Step 3, there exists a branch F'_j of G^{-m} univalent in the domain $\Omega_0(j)$ and such that $F'_j(z(j)) = x$. Define a domain M'_j as the preimage of M_j under $G'_0(j)$. Then $\mod (\Omega_0(j) \setminus M'_j) = \mod (M'_j \setminus \Omega'_0(j)) > m_0/2\ell$. Repeating the argument from Step 3, we get that the domains $F'_j(M'_j)$ are roughly discs around x and shrink to this point.

5. Finally, let us rescale the dynamical system $G'_0(j)\colon M'_j \to M_j$ by the maps $A_j(z) = (z - z(j))/\operatorname{diam}(M'_j)$ and $B_j(z) = (z - y(j))/\operatorname{diam}(M_j)$ in the domain of the definition and in the range of $G'_0(j)$ respectively. Denote the new system by

$$g_j\colon D'_j \to D_j, \text{ where}$$

$$D'_j = A_j(M'_j) \text{ and } D_j = B_j(M'_j)$$

are approximately Euclidean discs centred at zero with diameter 1 and

$$g_j = B_j \circ G'_0(j) \circ A_j^{-1}$$

is an ℓ-covering with a unique critical point $c_j = A_j(c) \in U_j = A_j(\Omega'_0(j))$ and critical value $g_j(c_j) \in V_j = B_j(\Omega_0(j))$.

The map g_j takes the line field $\mu'_j = (F'_j \circ A_j^{-1})^*(\mu)$ in D'_j onto the line field $\mu_j = (F_j \circ B_j^{-1})^*(\mu)$ in D_j (we use here that all maps $F'_j, F_j, G'_0(j)$ are iterates of G or inverses of iterates).

Now we are in a position to apply general theorems on sequences of invariant line fields and covering maps as in **[McM]**. By Theorem 5.2 in **[McM]** the sequences $(D'_j, 0)$ and $(D_j, 0)$ are pre-compact in the Caratheodory topology. Since their diameters are 1, we find two limit domains D' and D respectively, which are roughly discs around 0. Since the critical value $g_j(c_j) \in V_j$, and $\mod (D_j \setminus V_j) > m_0/2\ell$, by Theorem 5.6(3) **[McM]**, there is a limit map $g\colon D' \to D$, which is a branched degree ℓ-covering with a unique critical point $q \in D'$. On the other hand, by our construction and Theorem 5.16 **[McM]**, some subsequences of the line fields μ'_j and μ_j converge in measure to univalent line fields μ'_* and μ_* on D' and D respectively, and g takes μ'_* to μ_*. (A linefield is said to be univalent if it is a univalent pullback of the horizontal linefield). Since g has a critical point, this is a contradiction. $\square$

In order to use Propositions 2.1 and 2.3, we need to construct a sequence of ℓ-polynomial-like mappings as in the propositions. This is the content of the following statement, in which we assume that all critical points of a real polynomial are real.

Proposition 2.4 (**[GS]**, **[Ly1]**, **[LS1**, Theorem C**]**). — *Let f be a polynomial with the real coefficients and so that all critical points of f are real. Moreover, assume that all critical points escape to infinity, except for the critical point c of an even multiplicity ℓ. Assume that f is not renormalizable on the real axis and $\omega(c)$ is minimal. Then there exists a sequence of topological discs $\Omega_n \ni c$ such that $\Omega_n \cap \mathbb{R}$ is equal to the real trace of a puzzle piece, and such that* $\operatorname{diam}(\Omega_n) \to 0$ *so that the first return map to Ω_n along the points of the set $\omega(c) \cap \Omega_n$ is an ℓ-polynomial-like map $R_n : \cup_i \Omega_n^i \to \Omega_n$ and so that the modulus of the annulus between the range Ω_n and the central domain Ω_n^0 is bounded away from zero by a constant which only depends on f.*

This result was proved in **[Ly1]** and **[GS]** for the real quadratic polynomials, and adapted in **[LS1]**[Theorem C] and in **[GS1]** for real unimodal polynomials. For the polynomials as in the proposition still minor modifications of the proofs are needed, see the Remark after Theorem C in **[LS1]**, and also the next Section.

The following (in fact, well known) proposition settles the case when $\omega(c)$ is not minimal.

Proposition 2.5. — *Assume that G is a real ℓ-polynomial-like mapping, the map G restricted to the real line has no attracting or neutral periodic orbit, is non-renormalizable and that $\omega(c)$ is not minimal. Then $J(G)$ is totally disconnected and has zero Lebesgue measure.*

Proof. — If $\omega(c)$ is not minimal then it contains a point x whose forward orbit avoids some critical piece P_N. Here we use that the traces of the critical pieces on the real line tend to zero in diameter, because the real map G has no wandering interval. (In general, it is possible that a point remains outside a neighbourhood of c and still visits every critical piece (if they do not shrink to zero in diameter)). In particular, this forward orbit lies in a hyperbolic set. Therefore the puzzle-pieces $P_n(x)$ containing x shrink down in diameter to zero. The puzzle-pieces $P_n(x)$ are mapped by some iterates of G onto a fixed critical piece P_N: there is a fixed critical piece P_N and a sequence of critical pieces P_{n_k} with $n_k \to \infty$ so that each map $f^{n_k - N}\colon P_{n_k} \to P_N$ is ℓ-covering. Since there are no points of the postcritical set in a neighbourhood of the boundary of P_N, this proves the proposition. The statement that the Lebesgue measure is zero, follows from this as well. □

3. Proof of the Theorem

Let again $\ell \geq 2$ be the multiplicity (*i.e.*, the local degree) of the critical point c of the polynomial P, and assume ℓ is even. Let us first prove the total disconnectedness result. If $\ell = 2$ (the critical point c is simple), then it is proved in **[BH]** (even for the complex P). Let $\ell > 2$. If the critical point is not recurrent, then the proof is already given in **[BH]**. When the limit set $\omega(c)$ of the critical point c is not minimal, then the proof is also easy (though we use the reality of the maps: see Proposition 2.5). So assume $\omega(c)$ is minimal. We may assume $c = 0$. As a first step, we construct an initial ℓ-polynomial-like mapping as follows. Fix a level curve Γ of the Green function of the polynomial P, such that Γ is connected and not critical, *i.e.*, all preimages $P^{-n}(\Gamma)$ are smooth curves. By the condition of the theorem, one can pick up a full preimage $\gamma = P^{-k}(\Gamma)$ such that γ bounds finitely many topological discs $V_0, V_1, \dots$ such that $P : V_0 \to P(V_0)$ is ℓ-covering, and all other $P : V_i \to P(V_i)$ are one-to-one (this is still not a polynomial-like mapping because the images $P(V_i)$ could be different). Denote $V_0 = \Omega$ (the only domain containing the critical point $c = 0$ of P). The desired ℓ-polynomial-like mapping $G\colon \cup_{i=0}^{i_0} \Omega^i \to \Omega$ will be the *first return map* of the points of the set $\omega(c) \cap \Omega$ to Ω. Since $\omega(c)$ is minimal, the definition makes sense.

G obeys (by construction) the following *first return property*:

$$\text{if } P^n(c) \in \Omega, \text{ then } P^n(c) \text{ is an iterate of } G.$$

Moreover, the map G is real since it is the first return of a real map. We are going to apply Proposition 2.4. By the Straightening Theorem for polynomial-like maps **[DH]**, **[LM]**, **[LS1]**, $G\colon \cup_{i=0}^{i_0} \Omega^i \to \Omega$ can be quasi-conformally conjugated to a real polynomial f with all critical points real. (After all, we only need to move the escaping critical points.) To this end, let us consider the restriction $G|_{\mathbb{R}}$ of the map $G\colon \cup_{i=0}^{i_0} \Omega^i \to \Omega$ to the real axis. Then $G|_{\mathbb{R}}$ is unimodal on the central interval $T_0 = \Omega^0 \cap \mathbb{R} \ni c$, and maps each other interval $T_i = \Omega^i \cap \mathbb{R}$, $i = 1, \dots, i_0$ diffeomorphically onto T. Let us now add (finitely many) components to the domain of definition of G in such a way, that the new map $\widehat{G}$ is again a real ℓ-polynomial-like map, with an advantage that the graph of the new map on the real axis "looks like" a graph of a polynomial with all critical points being real, *i.e.*, the monotone increasing and monotone decreasing branches alternate each other. Now we can use the Straightening Theorem to conjugate the polynomial-like map $\widehat{G}$ with a polynomial f, so that all critical points of f are real. The existence of the sequence of maps induced by f follows now from Proposition 2.4. Moreover, the quasi-conformal conjugacy between f and $\widehat{G}$ transfers any polynomial-like structure induced by f to a one induced by $\widehat{G}$. On the other hand, since the induced maps we consider are the first returns along the postcritical set, and since the $\widehat{G}$-orbit of the critical point visits only the branches of the original map G, the maps induced by $\widehat{G}$ are, in fact, the ones, induced by G. Thus the statement of the theorem follows from Proposition 2.1.

To prove that no invariant linefields exist we consider two complementary cases: P is not renormalizable on the real line, or P is renormalizable. In the first case, we apply Propositions 2.3 and 2.4, if $\omega(c)$ is minimal, and Proposition 2.5 otherwise. In the second case, some iterate of P restricted to an appropriate neighbourhood of the critical point c is quasi-conformally conjugate to a polynomial f of the form $f(z) = z^\ell + c_1$, where c_1 is real, and the critical point $c = 0$ of f does not escape to infinity under the iterates of f. Again, there are two possibilities. If f is finitely many times renormalizable, we consider the last renormalization and again apply

Propositions 2.3 and 2.4, if $\omega_f(c)$ is minimal, or Proposition 3.2 of **[LS1]** (similar to Proposition 2.5) otherwise. On the other hand, if f is infinitely renormalizable, the result is proved already in **[McM]**. $\square$

References

[BH] B. Branner and J.H. Hubbard, *The iteration of cubic polynomials II,* Acta Math. **169**, 1992, 229–325.

[Dou1] A. Douady, *Description of compact sets in* $\mathbb{C}$, In: "Topological Methods in Modern Mathematics, A Symposium in Honor of John Milnor's 60th Birthday", Publish or Perish, 1992.

[Dou2] A. Douady, *Systèmes dynamiques holomorphes*, Séminaire Bourbaki, 35e année, 1982-83, exposé 599, novembre 1982.

[DH] A. Douady and J.H. Hubbard, *On the dynamics of polynomial-like mappings,* Ann. Sc. Éc. Norm. Sup. 4e série **18**, 1985, 287–343.

[GS] J. Graczyk and G. Świątek, *Induced expansion for quadratic polynomials*, Ann. Sc. Éc. Norm. Sup. 4e série **29**, 1996, 399–482.

[GS1] J. Graczyk and G. Świątek, *Polynomial-like mappings induced by real polynomials,* WSSIAA 4, 1995, 315–328.

[H] J.H. Hubbard *Local connectivity of Julia sets and bifurcation loci: three theorems of J.-C. Yoccoz*, preprint IHES, 1992

[Le] G. Levin, *Bounds for maps of an interval with one reflecting critical point. I.* Fundamenta Math **157**, 1998, 287–298.

[Ly] Oral communication by M. Lyubich, September 1996.

[Ly1] M.Yu. Lyubich, *Combinatorics, geometry and attractors of quasi-quadratic maps.* Annals of Math. **140**, 1994, 347–404.

[Ly2] M.Yu. Lyubich, *On typical behavior of trajectories of a rational mapping of the sphere.* Soviet Math. Dokl. **27**, 1983, 22–25.

[LM] M. Lyubich and J. Milnor, *The unimodal Fibonacci map,* Journal of the A.M.S. **6**, 1993, 425–457.

[LS1] G. Levin and S. van Strien, *Local connectivity of the Julia set of real polynomials.* Annals of Math. **147**, 1998, 471–541.

[LS2] G. Levin and S. van Strien, *Bounds for maps of an interval with one critical point of inflection type. II.*, Preprint.

[MS] W. de Melo and S. van Strien, *One-dimensional dynamics,* Ergebnisse Series **25**, Springer Verlag, (1993).

[Ma] R. Mañé, *On a theorem of Fatou*, Bol. Soc. Bras. Mat., **24**, 1993, 1–11.

[McM] C. McMullen, *Complex dynamics and renormalization,* Ann. Math. Studies, vol. 135, Princeton University Press, 1994.

[MSS] R. Mañe, P. Sad and D. Sullivan, *On the dynamics of rational maps,* Ann. Sc. Éc. Norm. Sup. 4e série **16**, 1983, 193–217.

[P-M] R. Perez-Marco, *Topology of Julia sets and hedgehogs,* Preprint, Orsay 1994.

[S] D. Sullivan, *Bounds, quadratic differentials, and renormalization conjectures, 1990,* In AMS Centennial Publications. **2**: Mathematics into Twenty-first Century.

[Y] J.C. Yoccoz, *MLC*, Manuscript.

G. LEVIN, Institute of Mathematics, Hebrew University, Givat Ram, 91904 Jerusalem, Israel
E-mail : `levin@math.huji.ac.il`

S. VAN STRIEN, Mathematics Department, University of Warwick, Coventry CV4 7AL, England
E-mail : `strien@maths.warwick.ac.uk`

MIKHAIL LYUBICH

Dynamics of quadratic polynomials, III : parapuzzle and SBR measures

Astérisque, tome 261 (2000), p. 173-200
<http://www.numdam.org/item?id=AST_2000__261__173_0>

Astérisque
261, 2000, p. 173–200

DYNAMICS OF QUADRATIC POLYNOMIALS, III PARAPUZZLE AND SBR MEASURES

by

Mikhail Lyubich

Dedicated to 60th birthday of A. Douady

Abstract. — This is a continuation of notes on the dynamics of quadratic polynomials. In this part we transfer our previous geometric result [L3] to the parameter plane. To any parameter value c (outside the main cardioid and the little Mandelbrot sets attached to it) we associate a "principal nest of parapuzzle pieces". We then prove that the moduli of the annuli between two consecutive pieces grow at least linearly. This implies, using Martens & Nowicki (*cf.* this volume) geometric criterion for existence of an absolutely continuous invariant measure together with [L2], that Lebesgue almost every real quadratic polynomial is either hyperbolic, or has a finite absolutely continuous invariant measure, or is infinitely renormalizable. In the further papers [L5,L7] we show that the latter set has zero Lebesgue measure, which completes the measure-theoretic picture of the dynamics in the real quadratic family.

You first plow in the dynamical plane
and then harvest in the parameter plane.
Adrien Douady

1. Introduction

This is a continuation of notes on dynamics of quadratic polynomials. In this part we transfer the geometric result of **[L3]** to the parameter plane. To any parameter value $c \in M$ in the Mandelbrot set (which lies outside of the main cardioid and satellite Mandelbrot sets attached to it) we associate a "principal nest of parapuzzle pieces"

$$\Delta^0(c) \supset \Delta^1(c) \supset \cdots$$

corresponding to the generalized renormalization type of c. Then we prove:

1991 ***Mathematics Subject Classification.*** — 58F23, 58F03.
Key words and phrases. — Mandelbrot set, puzzle, invariant measure.

This work was partially supported by the NSF.

Theorem A. — *The moduli of the parameter annuli* $\mathrm{mod}(\Delta^l(c) \setminus \Delta^{l+1}(c))$ *grow at least linearly.*

(See §4 for a more precise formulation.)

This result was announced at the Colloquium in honor of Adrien Douady (July 1995), and in the survey **[L4]**, Theorem 4.8. The main motivation for this work was to prove the following:

***Theorem B* (joint with Martens and Nowicki).** — *Lebesgue almost every real quadratic* $P_c : z \mapsto z^2 + c$ *which is non-hyperbolic and at most finitely renormalizable has a finite absolutely continuous invariant measure.*

More specifically, Martens and Nowicki **[MN]** have given a geometric criterion for existence of a finite absolutely continuous invariant measure (acim) in terms of the "scaling factors". Together with the result of **[L2]** on the exponential decay of the scaling factors in the quasi-quadratic case this yields existence of the acim once "the principal nest is eventually free from the central cascades". Theorem A above implies that this condition is satisfied for almost all real quadratics which are non-hyperbolic and at most finitely renormalizable (see Theorem 5.1). Note that Theorem A also implies that this condition is satisfied on a set of positive measure, which yields a new proof of Jacobson's Theorem **[J]** (see also Benedicks & Carleson **[BC]**).

A measure μ will be called SBR (Sinai-Bowen-Ruelle) if

$$\frac{1}{n}\sum_{k=0}^{n-1} \delta_{f^k x} \to \mu \tag{1.1}$$

for a set of x of positive *Lebesgue* measure. It is known that if an SBR measure exists for a real quadratic map $f = P_c$, $c \in [-2, 1/4]$, on its invariant interval I_c, then it is unique and (1.1) is satisfied for Lebesgue almost all $x \in I_c$ (see Introduction of **[MN]** for a more detailed discussion and references). Theorem B yields

Corollary. — *For almost all* $c \in [-2, 1/4]$, *the quadratic polynomial* P_c *has a unique SBR measure on its invariant interval* I_c.

Another consequence of our geometric results is concerned with the shapes of little Mandelbrot copies (see **[L3]**, §2.5, for a discussion of little Mandelbrot copies). Let us say that a Mandelbrot set M' has a (K, ε)*-a bounded shape* if the straightening $\chi : M' \to M$ admits a K-quasi-conformal extension to an (ε diam M')-neighborhood of M'. We say that the little Mandelbrot sets of some family *have bounded shape* if a bound (K, ε) can be selected uniform over the family.

A Mandelbrot copy M' is called *maximal* if it is not contained in any other copy except M itself. It is called *real* if it is centered at the real line.

A little Mandelbrot copy encodes the combinatorial type of the corresponding renormalization. In **[L3]** we dealt with diverse numerical functions of the combinatorial type. For real copies a crucial information is encoded by the *essential period* $p_e(M')$ (see **[L3**, §8.1], **[LY]**).

For a definition of Misiurewicz wakes see §3.3 of this paper.

Theorem C. — *For any Misiurewicz wake O, the maximal Mandelbrot copies contained in O have bounded shape. In particular, all maximal real Mandelbrot copies, except the doubling one, have a bounded shape. Moreover, a real copy M' has a (K,ε)-bounded shape, where $K \to 1$ and $\varepsilon \to \infty$ as $p_e(M') \to \infty$.*

In §6 we will refine this statement and will comment on its connection with the MLC problem and the renormalization theory.

Let us now take a closer look at Theorem A. It nicely fits to the general philosophy of correspondence between the dynamical and parameter plane. This philosophy was introduced to holomorphic dynamics by Douady and Hubbard **[DH1]**. Since then, there have been many beautiful results in this spirit, see Tan Lei **[TL]**, Rees **[R]**, Shishikura **[Sh]**, Branner-Hubbard **[BH]**, Yoccoz (see **[H]**).

In the last work, special tilings into "parapuzzle pieces" of the parameter plane are introduced. Its main geometric result is that the tiles around at most finitely renormalizable points shrink. It was done by transferring, in an ingenious way, the corresponding dynamical information into the parameter plane.

In **[L3]** we studied the rate at which the dynamical tiles shrink. The main geometric result of that paper is the linear growth of the moduli of the principal dynamical annuli. Let us note that the way we transfer this result to the parameter plane (Theorem A) is substantially different from that of Yoccoz. Our main conceptual tool is provided by holomorphic motions whose transversal quasi-conformality is responsible for commensurability between the dynamical and parameter pictures (compare Shishikura **[Sh]**). To make it work we exploit existence of uniform quasi-conformal pseudo-conjugacy between the generalized renormalizations **[L3]**.

The properties of holomorphic motions are discussed in §2. In §3 we describe the principal parameter tilings according to the generalized renormalization types of the maps. In §4 we prove Theorem A. In §5, we derive the consequence for the real quadratic family (Theorem B). In the last section, §6, we prove Theorem C on the shapes of Mandelbrot copies.

Let us finally draw the reader's attention to the work of LeRoy Wenstrom **[W]** which studies in detail parapuzzle geometry near the Fibonacci parameter value.

Remark. — We have recently proven that the set of infinitely renormalizable real parameter values has zero linear measure. Together with Theorem B this implies that almost every real quadratic has either an attracting cycle or an absolutely continuous invariant measure **[L7]**.

2. Background

2.1. Notations and terminology

$$\begin{aligned} \mathbb{D}_r(p) &= \{z : |z-p| < r\}; \\ \mathbb{D}_r &\equiv \mathbb{D}_r(0); \\ \mathbb{D} &\equiv \mathbb{D}_1; \\ \mathbb{T}_r &= \{z : |z| = r\}; \\ \mathbb{A}(r,R) &= \{r < |z| < R\}. \end{aligned}$$

The closed and semi-closed annuli are denoted accordingly: $\mathbb{A}[r,R]$, $\mathbb{A}(r,R]$, $\mathbb{A}[r,R)$.

By a *topological disc* we will mean a simply connected domain $D \subset \mathbb{C}$ whose boundary is a Jordan curve.

Let π_1 and π_2 denote the coordinate projections $\mathbb{C}^2 \to \mathbb{C}$. Given a set $\mathbb{X} \subset \mathbb{C}^2$, we denote by $X_\lambda = \pi_1^{-1}\{\lambda\}$ its vertical cross-section through λ (the "fiber" over λ). Vice versa, given a family of sets $X_\lambda \subset \mathbb{C}$, $\lambda \in D$, we will use the notation:

$$\mathbb{X} = \cup_{\lambda \in D} X_\lambda = \{(\lambda, z) \in \mathbb{C}^2 : \lambda \in D, z \in X_\lambda\}.$$

Let us have a discs fibration $\pi_1 : \mathbb{U} \to D$ over a topological disc $D \subset \mathbb{C}$ (such that the sections U_λ are topological discs, and the closure of $\mathbb{U}$ in $D \times \mathbb{C}$ is homeomorphic to $D \times \overline{\mathbb{D}}$ over D). In this situation we call $\mathbb{U}$ an (open) *topological bidisc* over D. We say that this fibration admits an extension to the boundary ∂D if the closure $\overline{\mathbb{U}}$ of $\mathbb{U}$ in $\mathbb{C}^2$ is homeomorphic over $\overline{D}$ to $\overline{D} \times \overline{\mathbb{D}}$. The set $\overline{\mathbb{U}}$ is called a (closed) bidisc. We keep the notation $\mathbb{U}$ for the fibration of *open* discs over the closed disc $\overline{D}$ (it will be clear from the context over which set the fibration is considered).

If $U_\lambda \ni 0$, $\lambda \in D$, we denote by $\mathbf{0}$ the zero section of the fibration.

Given a domain $\Delta \subset D$, let $\mathbb{U}|\Delta = \mathbb{U} \cap \pi_1^{-1}\Delta$. This is a bidisc over Δ.

If the fibration π_1 admits an extension over the boundary ∂D, we define the *frame* $\delta\mathbb{U}$ as the topological torus $\cup_{\lambda \in \partial D} \partial U_\lambda$. A section $\mathbf{\Phi} : D \to \mathbb{U}$ is called *proper* if it is continuous up to the boundary and $\mathbf{\Phi}(\partial D) \subset \delta U$.

We assume that the reader is familiar with the theory of quasi-conformal maps (see e.g., [**A**]). We will use a common abbreviation K-qc for "K-quasi-conformal". Dilatation of a qc map h will be denoted as $\mathrm{Dil}(h)$.

Notation $a_n \asymp b_n$ means, as usual, that the ratio a_n/b_n is positive and bounded away from 0 and ∞.

2.2. Holomorphic motions. — Given a domain $D \subset \mathbb{C}$ with a base point $*$ and a set $X_* \subset \mathbb{C}$, a *holomorphic motion* $\boldsymbol{h}$ of X_* over D is a family of injections $h_\lambda : X_* \to \mathbb{C}$, $\lambda \in D$, such that $h_* = \mathrm{id}$ and $h_\lambda(z)$ is holomorphic in λ for any $z \in X_*$. We denote $X_\lambda = h_\lambda X_*$. The restriction of $\boldsymbol{h}$ to a parameter domain $\Delta \subset D$ will be denoted as $\boldsymbol{h}|\Delta$.

Let us summarize fundamental properties of holomorphic motions which are usually referred to as the λ-*lemma*. It consists of two parts: extension of the motion and transversal quasi-conformality, which will be stated separately. The consecutively improving versions of the Extension Lemma appeared in **[L1]** and **[MSS]**, **[ST]**, **[BR]**, **[Sl]**. The final result, which will be actually exploited below, is due to Slodkowsky:

Extension Lemma. — *A holomorphic motion $h_\lambda : X_* \to X_\lambda$ of a set $X_* \subset \mathbb{C}$ over a topological disc D admits an extension to a holomorphic motion $H_\lambda : \mathbb{C} \to \mathbb{C}$ of the whole complex plane over D.*

Quasi-Conformality Lemma (**[MSS]**). — *Let $h_\lambda : U_* \to U_\lambda$ be a holomorphic motion of a domain $U_* \subset \mathbb{C}$ over a hyperbolic domain $D \subset \mathbb{C}$. Then the maps h_λ are $K(r)$-quasi-conformal, where r is the hyperbolic distance between $*$ and λ in D.*

Let us define the dilatation of the holomorphic motion as

$$\mathrm{Dil}(\boldsymbol{h}) = \sup_{\lambda \in D} \mathrm{Dil}(h_\lambda).$$

It can be equal to ∞ over the whole domain D but becomes finite ($\leq K(r)$) over the hyperbolic disk of radius r.

A holomorphic motion $h_\lambda : U_* \to U_\lambda$ over D can be viewed as a complex one-dimensional foliation of the domain $\mathbb{U} = \cup_{\lambda \in D} U_\lambda$, whose leaves are graphs of the functions $\lambda \mapsto h_\lambda(z)$, $z \in U_*$. A *transversal* to the motion is a complex one dimensional submanifold of $\mathbb{C}^2$ which transversally intersects every leaf at one point (so that "transversal" will mean a *global* transversal). Given two transversals X and Y, we thus have a well-defined holonomy map $H : X \to Y$, $H(p) = q$ iff p and q belong to the same leaf.

A map $H : X \to Y$ is called *locally qc* at $p \in X$ if it is qc in some neighborhood of p. In this case the local dilatation of H at p is defined as the limit of $\mathrm{Dil}(H \mid \mathbb{D}_\varepsilon(p))$, as $\varepsilon \to 0$.

Corollary 2.1 (**Transverse qc structure**). — *Any holomorphic motion $\boldsymbol{h}$ over D is locally transversally quasi-conformal. More precisely, for any two transversals X and Y, the holonomy map $H : X \to Y$ is locally quasi-conformal. If $H(p) = q$ then the local dilatation of H at p depends only on the hyperbolic distance between the $\pi_1(p)$ and $\pi_1(q)$ in D. If* $\mathrm{Dil}(\boldsymbol{h}) < \infty$ *then the holonomy H is globally qc with* $\mathrm{Dil}(H) \leq \mathrm{Dil}(\boldsymbol{h})^2$.

Proof. — Let $p = (\lambda, \alpha)$, $q = (\mu, \beta)$. By the λ-Lemma, the map $G = h_\mu \circ h_\lambda^{-1} : U_\lambda \to U_\mu$ is quasi-conformal, with dilatation depending only on the hyperbolic distance between λ and μ in D and bounded by $\mathrm{Dil}(\boldsymbol{h})^2$. Hence a little disc $\mathbb{D}_\varepsilon(\alpha) \subset U_\lambda$ is mapped by G onto an ellipse $Q_\varepsilon \subset U_\mu$ with bounded eccentricity about β (where the bound depends only on the hyperbolic distance between λ and μ.

But the holonomy $U_\lambda \to X$ is asymptotically conformal near p. To see this, let us select a holomorphic coordinates (θ, z) near p in such a way that $p = 0$ and the leaf via p becomes the parameter axis. Let $z = \psi(\theta) = \varepsilon + \cdots$ parametrizes a nearby leaf of the foliation, while $\theta = g(z) = bz + \cdots$ parametrizes the transversal X.

Let us do the rescaling $z = \varepsilon\zeta, \theta = \varepsilon\nu$. In these new coordinates, the above leaf is parametrized by the function $\Psi(\nu) = \varepsilon^{-1}\psi(\varepsilon\nu)$, $|\nu| < R$, where R is a fixed parameter. Then $\Psi'(\nu) = \psi'(\varepsilon\nu)$ and $\Psi''(\nu) = \varepsilon\psi''(\varepsilon\nu)$. Since the family of functions $\{\psi(\theta)\}$ is normal, $\Psi''(\nu) = O(\varepsilon)$. Moreover, ψ uniformly goes to 0 as $\psi(0) \to 0$. Hence $|\Psi'(0)| = |\psi'(0)| \leq \delta_0(\varepsilon)$, where $\delta_0(\varepsilon) \to 0$ as $\varepsilon \to 0$. Thus $\Psi'(\nu) = \delta_0(\varepsilon) + O(\varepsilon) \leq \delta(\varepsilon) \to 0$ as $\varepsilon \to 0$ uniformly for all $|\nu| < R$. It follows that $\Psi(\nu) = 1 + O(\delta(\varepsilon)) = 1 + o(1)$ as $\varepsilon \to 0$.

On the other hand, the manifold X is parametrized in the rescaled coordinates by a function $\nu = b\zeta + 0(1)$. Since the transverse intersection persists, X intersects the leaf at the point $(\nu_0, \zeta_0) = (b, 1)(1 + o(1))$ (so that R should be selected bigger than b). In the old coordinates the intersection point is $(\theta_0, z_0) = (b\varepsilon, \varepsilon)(1 + o(1))$.

Thus the holonomy from U_λ to X transforms the disc of radius $|\varepsilon|$ to an ellipse with small eccentricity, which means that this holonomy is asymptotically conformal. As the holonomy from U_μ to Y is also asymptotically conformal, the holonomy $H : (X, p) \to (Y, q)$ is locally qc at p, and its local dilatation at p is the same as the local dilatation of $G : (U_\lambda, p) \to (U_\mu, q)$. Thus it depends only on the hyperbolic distance between λ and μ, and is bounded by $\mathrm{Dil}(\boldsymbol{h})^2$.

To conclude the proof, one should just remark that a map is globally qc if and only if it is locally qc with uniformly bounded local dilatations, and then the global dilatation is equal to the supremum of the local ones. $\square$

Remark. — The author thanks the referee for pointing out that the above Corollary also follows from [**DH2**, p. 327] (compare also [**Sh**, §3]).

2.3. Winding number. — Given two curves $\psi_1, \psi_2 : \partial D \to \mathbb{C}$ such that $\psi_1(\lambda) \neq \psi_2(\lambda)$, $\lambda \in \partial D$, the winding number of the former about the latter is defined as the increment of $\frac{1}{2\pi} \arg(\psi_1(\lambda) - \psi_2(\lambda))$ as λ wraps once around ∂D.

Let us have a bidisc $\mathbb{U}$ over $\overline{D}$. Given a proper section $\boldsymbol{\Phi} : D \to \mathbb{U}$ let us define its *winding number* as follows. Let us mark on the torus $\delta\mathbb{U}$ the homology basis $\{[\partial D], [\partial U_*]\}$. Then the winding number $w(\boldsymbol{\Phi})$ is the second coordinate of the curve $\boldsymbol{\Phi} : \partial D \to \delta\mathbb{U}$ with respect to this basis.

Argument Principle. — *Let us have a bidisc $\mathbb{U}$ over $\overline{D}$ and a proper holomorphic section $\boldsymbol{\Phi} : D \to \mathbb{U}$, $\phi = \pi_2 \circ \boldsymbol{\Phi}$. Let $\boldsymbol{\Psi} : \overline{D} \to \mathbb{U}$ be another continuous section holomorphic in D, $\psi = \pi_2 \circ \boldsymbol{\Psi}$. Then the number of solutions of the equation $\phi(\lambda) = \psi(\lambda)$ counted with multiplicity is equal to the winding number $w(\boldsymbol{\Phi})$.*

Proof. — Indeed, $w(\boldsymbol{\Phi})$ is equal to the winding number of ϕ around ψ, which is equal, by the standard Argument Principle, to the number of roots of the equation

$$\phi(\lambda) = \psi(\lambda). \qquad \square$$

3. Parapuzzle combinatorics

3.1. Holomorphic families of generalized quadratic-like maps. — Let us consider a topological disc $D \subset \mathbb{C}$ with a base point $* \in D$, and a family of topological bidiscs $\mathbb{V}_i \subset \mathbb{U} \subset \mathbb{C}^2$ over D (*tubes*), such that the $\mathbb{V}_i$ are pairwise disjoint. We assume that $V_{0,\lambda} \ni 0$.

Let

$$\boldsymbol{g} : \cup \mathbb{V}_i \to \mathbb{U} \tag{3.1}$$

be a fiberwise map, which admits a holomorphic extension to some neighborhoods of the $\mathbb{V}_i$ (warning: these extensions don't fit), and whose fiber restrictions

$$\boldsymbol{g}(\lambda, \cdot) \equiv g_\lambda : \cup_i V_{i,\lambda} \to U_\lambda, \quad \lambda \in D,$$

are generalized quadratic-like maps with the critical point at $0 \in V_\lambda \equiv V_{0,\lambda}$ (see **[L3]**, §3.7 for the definition). We will assume that the discs U_λ and $V_{i,\lambda}$ are bounded by piecewise smooth quasi-circles.

Let us also assume that there is a holomorphic motion $\boldsymbol{h}$ over $(D, *)$,

$$h_\lambda : (\overline{U}_*, \cup_i \partial V_{i,*}) \to (\overline{U}_\lambda, \cup_i \partial V_{i,\lambda}), \tag{3.2}$$

which respects the boundary dynamics:

$$h_\lambda \circ g_*(z) = g_\lambda \circ h_\lambda(z) \quad for \quad z \in \cup \partial V_{i,*}. \tag{3.3}$$

A *holomorphic family* $(\boldsymbol{g}, \boldsymbol{h})$ *of (generalized) quadratic-like maps over* D is a map (3.1) together with a holomorphic motion (3.2) satisfying (3.3). We will sometimes reduce the notation to $\boldsymbol{g}$. In case when the domain of $\boldsymbol{g}$ consists of only one tube $\mathbb{V}_0$, we refer to $\boldsymbol{g}$ as *DH quadratic-like family* (for "Douady and Hubbard", compare **[DH2]**).

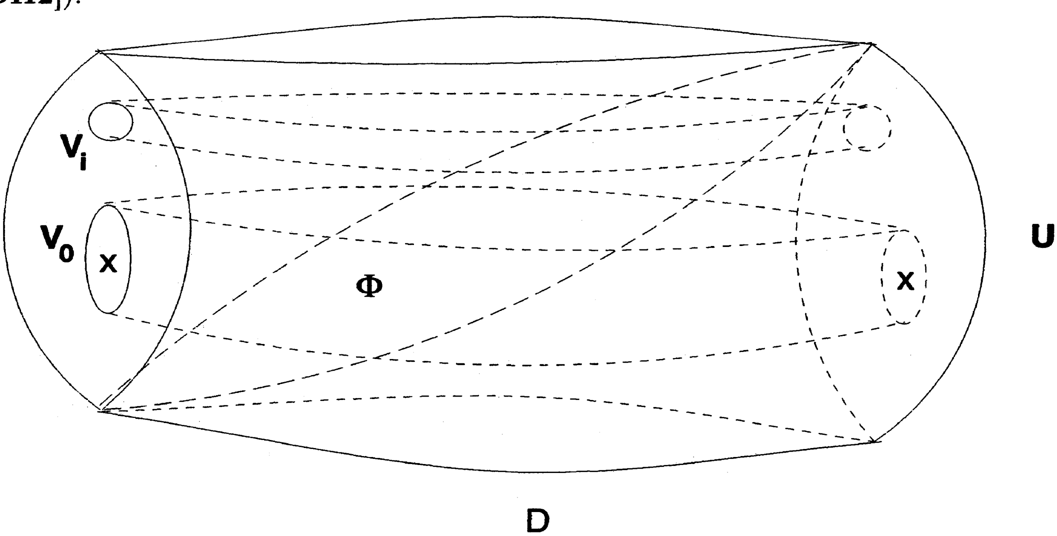


Figure 1. Generalized quadratic-like family.

Remark. — It would be more consistent to call just $\boldsymbol{g}$ a holomorphic family, while to call the pair $(\boldsymbol{g},\boldsymbol{h})$, say, an *equipped* holomorphic family. However, in this paper we will assume that the families are equipped, unless otherwise is explicitly stated.

Let us now consider the critical value function $\phi(\lambda) \equiv \phi_{\boldsymbol{g}}(\lambda) = g_\lambda(0)$, $\boldsymbol{\Phi}(\lambda) \equiv \boldsymbol{\Phi}_{\boldsymbol{g}}(\lambda) = \boldsymbol{g}(\lambda, 0) \equiv (\lambda, \phi(\lambda))$. Let us say that $\boldsymbol{g}$ is a *proper* (or *full*) holomorphic family if the fibration $\pi_1 : \mathbb{U} \to D$ admits an extension to the boundary $\overline{D}$, $\overline{\mathbb{V}}_i \subset \mathbb{U}$, and $\boldsymbol{\Phi} : D \to \mathbb{U}$ is a proper section. Note that the fibration $\pi_1 : \mathbb{V}_0 \to D$ cannot be extended to $\overline{D}$, as the domains $V_{\lambda,0}$ pinch to figure eights as $\lambda \to \partial D$.

Given a proper holomorphic family $\boldsymbol{g}$ of generalized quadratic-like maps, let us define its *winding number* $w(\boldsymbol{g})$ as the winding number of the critical value $\phi(\lambda)$ about the critical point 0. By the Argument Principle, it is equal to the winding number of the critical value about any section $\overline{D} \to \mathbb{U}$.

We will also face the situation when $\boldsymbol{g}$ does not map every tube $\mathbb{V}_i$ onto the whole tube $\mathbb{U}$ but still satisfies the following Markov property: $\boldsymbol{g}\mathbb{V}_i$ either contains $\mathbb{V}_j$ or disjoint from it (and all the rest properties listed above are still valid, see §3.3). Then we call $\boldsymbol{g}$ a holomorphic family of Markov maps.

Let $\mathrm{mod}(\boldsymbol{g}) = \inf_{\lambda \in D} \mathrm{mod}(U_\lambda \smallsetminus V_{0,\lambda})$.

3.2. Douady & Hubbard quadratic-like families. — Let us consider a proper holomorphic family $\boldsymbol{f} : \mathbb{V} \to \mathbb{U}$ of DH quadratic-like maps. The Mandelbrot set $M(\boldsymbol{f})$ is defined as the set of $\lambda \in D$ such that the Julia set $J(f_\lambda)$ is connected. We will assume that $* \in M(\boldsymbol{f})$.

Since the U_λ and $V_{i,\lambda}$ are bounded by quasi-circles, there is a qc straightening $\omega_* : \mathrm{cl}(U_* \smallsetminus V_*) \to \mathbb{A}[2,4]$ conjugating $f_* : \partial V_* \to \partial U_*$ to $z \mapsto z^2$ on $\mathbb{T}_2$. The holomorphic motion $\boldsymbol{h}$ on the "condensator" $\mathbb{U} \smallsetminus \mathbb{V}$ spreads this straightening over the whole parameter region D. We obtain a family of quasi-conformal homeomorphisms

$$\omega_\lambda : \mathrm{cl}(U_\lambda \smallsetminus V_\lambda) \to \mathbb{A}[2,4] \tag{3.4}$$

conjugating $f_\lambda | \partial V_\lambda$ to $z \mapsto z^2$ on $\mathbb{T}_2$. Pulling them back, we obtain for every f_λ the straightening $\omega_\lambda : \Omega_\lambda \to \mathbb{A}(\rho_\lambda, 4)$ well-defined up to the critical point level $\rho_\lambda = |\omega_\lambda(0)|$ (so that for $\lambda \in M(\boldsymbol{f})$ it is well-defined on the whole complement of the Julia set). This determines *external coordinates* of points $z \in \Omega_\lambda$, radius r and angle θ, defined as the polar coordinates of $\omega_\lambda(z)$.

Note that if $\mathrm{Dil}(\boldsymbol{h}) < \infty$ then the straightenings ω_λ are uniformly L-qc with $L = \mathrm{Dil}(\boldsymbol{h}) \cdot \mathrm{Dil}(\omega_*)$. Note also that $\mathrm{Dil}(\omega_*)$ depends only on the qc dilatation of the quasi-circle ∂U_*, ∂V_* and on $\mathrm{mod}(U_* \smallsetminus V_*)$.

By the *geometry* of $(\boldsymbol{f},\boldsymbol{h})$ we will mean a triple of parameters: $(\mathrm{mod}(\boldsymbol{f}))^{-1}, \mathrm{Dil}(\boldsymbol{h})$, and the best dilatation of ω_*. If $\mathrm{mod}(\boldsymbol{f}) \to \infty$, while the $\mathrm{Dil}(\boldsymbol{h})$ and $\mathrm{Dil}(\omega_*)$ go to 1 (over some directed set of quadratic-like families), then we say that the geometry of $(\boldsymbol{f},\boldsymbol{h})$ vanishes.

By an *adjustment* of a DH quadratic-like family we will mean replacement the domains U_*, V_* with some other domains $\widetilde{U}_* \subset U_*$, $\widetilde{V}_* = f_*^{-1}\widetilde{U}_*$, spreading them around by $\boldsymbol{h}$ ($\widetilde{U}_\lambda = h_\lambda \widetilde{U}_*$, $\widetilde{V}_\lambda = h_\lambda \widetilde{V}_*$), and the corresponding shrinking of the parameter domain: $\widetilde{D} = \Phi^{-1}\widetilde{\mathbb{U}}$. It provides us with an adjusted family ($\boldsymbol{f} : \widetilde{\mathbb{V}} \to \widetilde{\mathbb{U}}$, $\boldsymbol{h}$) over $\widetilde{D}$.

We will use the following standard adjustment. Select $\widetilde{U}_*$ to be bounded by the hyperbolic geodesic Γ in the annulus $U_* \setminus V_*$. Then $\widetilde{V}_*$ is bounded by the hyperbolic geodesic Γ' in $V_* \setminus f_*^{-1}V_*$. By the Koebe Theorem, the geometry of these geodesics (i.e., their qc dilatation) depends only on $\mathrm{mod}(U_* \setminus V_*)$. Thus after this adjustment, $\mathrm{Dil}(\widetilde{\omega}_*)$ depends only on $\mathrm{mod}(U_* \setminus V_*)$. Moreover, $\mathrm{mod}(\widetilde{U}_* \setminus \widetilde{V}_*) \geq (3/4)\,\mathrm{mod}(U_* \setminus V_*)$. Thus the geometry of the adjusted family depends only on $\mathrm{mod}(U_* \setminus V_*)$ and $\mathrm{Dil}(\boldsymbol{h})$.

Moreover, if we fix $\mathrm{Dil}(\boldsymbol{h})$ and let $\mathrm{mod}(U_* \setminus V_*) \to \infty$, then $\mathrm{mod}(\widetilde{U}_\lambda \setminus \widetilde{V}_\lambda)) \to \infty$, $\mathrm{Dil}(\boldsymbol{h}|\widetilde{D}) \to 1$ (by λ-lemma), and $\mathrm{Dil}(w_*) \to 1$, so that the geometry of the adjusted family vanishes.

In what follows we will not change notations when we adjust quadratic-like families.

Let us now define a map $\xi : D \setminus M(\boldsymbol{f}) \to \mathbb{A}(1,4)$ in the following way:

$$\xi(\lambda) = \omega_\lambda(f_\lambda 0). \tag{3.5}$$

Lemma 3.1. — *Let $(\boldsymbol{f}, \boldsymbol{h})$ be a DH quadratic-like family with winding number* 1. *Then formula (3.5) determines a homeomorphism $\xi : D \setminus M(\boldsymbol{f}) \to \mathbb{A}(1,4)$. If* $\mathrm{Dil}(\boldsymbol{h}) < \infty$ *then ξ is L-qc with L depending only on the geometry of $(\boldsymbol{f}, \boldsymbol{h})$. Moreover, $L \to 1$ as the geometry of $(\boldsymbol{f}, \boldsymbol{h})$ vanishes.*

Proof. — Let us consider the critical value graph $X = \Phi(\lambda) \equiv \{(\lambda, f_\lambda 0),\ \lambda \in D\}$. By the Argument Principle, it intersects at a singe point each leaf of the holomorphic motion $\boldsymbol{h}$ on $\mathbb{U} \setminus \mathbb{V}$, so that the holonomy $\gamma : U_* \setminus V_* \to X$ is a homeomorphism onto the image R_1. Hence $A_1 \equiv \pi_1 R_1 \subset D$ is a topological annulus, and the map

$$\xi^{-1} = \pi_1 \circ \gamma \circ \omega_*^{-1} : \mathbb{A}[2,4) \to A_1$$

is a homeomorphism.

Let Γ_1 be the inner boundary of A_1, and D_1 be the topological disc bounded by Γ_1. Since the critical value $f_\lambda(0)$, $\lambda \in D_1$, does not land at the leaves of holomorphic motion $\boldsymbol{h}|D_1$, it can be lifted by $\boldsymbol{f}$ to a holomorphic motion $\boldsymbol{h}_1$ of the annulus $V_*^1 \setminus V_*^2$ over D_1, where $V_\lambda^1 \equiv V_\lambda$ and $V_\lambda^2 = f_\lambda^{-1}V_\lambda$. Since the graph X intersects every leaf belonging to $\partial\mathbb{V}^1$ at a single point, the family $(\boldsymbol{f} : \mathbb{V}^2 \to \mathbb{V}^1, \boldsymbol{h})$ is proper over D_1 and has winding number 1. Let $A_2 = \Phi^{-1}(\mathbb{V}^1 \setminus \mathbb{V}^2)$. Then the same argument as above shows that the map $\xi^{-1} : \mathbb{A}[\sqrt{2}, 2) \to A_2$ is also a homeomorphism.

Continuing in the same way, we will inductively construct a sequence of holomorphic motions $\boldsymbol{h}_n$ over nested discs D_n, and a nest of adjoint annuli $A_n = D_{n-1} \setminus D_n$ which are homeomorphically mapped by ξ onto the round annuli $\mathbb{A}[2^{2^{1-n}}, 2^{2^{2-n}})$. Altogether this shows that ξ is a homeomorphism.

Finally, assume $\boldsymbol{h}$ is K-qc. Since all further motions $\boldsymbol{h}_n$ are holomorphic lifts of $\boldsymbol{h}$ over D_n by $\boldsymbol{f}^n$, they are K-qc over their domains of definition as well. By Corollary 2.1, they are transversally K-qc. Moreover, the straightening

$$\omega_* : U_* \setminus J(f_*) \longrightarrow \mathbb{A}(1,4)$$

is qc, while the projection $\pi_1 : X \to D$ is conformal. Since ξ is the composition of the straightening, the holonomy and the projection, it is L-qc with $L = K \cdot \mathrm{Dil}(\omega_*)$. In particular, $L \to 1$ as K and $\mathrm{Dil}(\omega_*)$ go to 1. □

Note that the motions $\boldsymbol{h}_n$ over domains D_n constructed in the above proof preserve the external coordinates: $\omega_\lambda(h_\lambda z) = \omega_*(z)$, $z \in U_* \setminus f_*^{-n} V_*$. We will refer to this property by saying that $\boldsymbol{h}$ *respects the external marking*, or that $\boldsymbol{h}$ is *marked*.

***Example* (see [DH1]).** — Let us consider the Mandelbrot set M of the quadratic family $P_c : z \mapsto z^2 + c$. Let $R : \mathbb{C} \setminus M \to \mathbb{C} \setminus \mathbb{D}$ be the Riemann mapping tangent to id at ∞. Recall that parameter equipotentials and external rays are defined as the R-preimages of the round circles and radial rays. Let Ω_r be the topological disc bounded by the equipotential $R^{-1}\{re^{i\theta} : 0 \le \theta \le 2\pi\}$ of radius $r > 1$.

For every $c \in \Omega_4$, let us consider the quadratic-like map $P_c : V_c \to U_c$ where V_c and U_c are topological discs bounded by the dynamical equipotentials of radius 2 and 4 correspondingly. Then the conformal map $\omega_c : U_c \setminus V_c \to \mathbb{A}[2,4)$ conjugates $P_c|\partial V_c$ to $z \mapsto z^2$ on $\mathbb{T}_2$, so that it can serve as a straightening (3.4). With this choice of the straightening, the parameter map $\xi : D \setminus M \to \mathbb{A}(1,4)$ constructed in Lemma 3.1 is just the restriction of the Riemann map R. □

With Lemma 3.1, we can extend the notion of parameter rays and equipotentials to quadratic-like families as the ξ-preimages of the polar coordinate curves in $\mathbb{A}(1,4)$. If $\xi(\lambda) = re^{i\theta}$ then r and θ are called the *external radius* and the *external angle* of the parameter value λ. Note that ∂D becomes the equipotential of radius 4.

Before going further, let us state a general lemma about qc maps:

Gluing Lemma. — *Let us have a compact set $Q \subset \mathbb{C}$ and two its neighborhoods U and V. Let us consider two qc maps $\phi : U \to \mathbb{C}$ and $\psi : V \setminus Q \to \mathbb{C}$. Assume that these maps match on ∂Q, i.e., the map $f : V \to \mathbb{C}$ defined as ϕ on Q and as ψ on $V \setminus Q$ is continuous. Then f is quasi-conformal and* $\mathrm{Dil}(f) = \max(\mathrm{Dil}(\phi|Q), \mathrm{Dil}(\psi))$.

Proof. — See e.g., [**DH2**, Lemma 2, p. 303]. □

Recall now that every quadratic-like map $f : V \to U$ is hybrid equivalent to a quadratic polynomial $P_c : z \mapsto z^2 + c$ (The Straightening Theorem [**DH2**]). It is constructed by gluing f to $z \mapsto z^2$ on $\mathbb{C} \setminus \mathbb{D}_2$ (by means of the qc straightening $\omega : \mathrm{cl}(U \setminus V) \to \mathbb{A}[2,4]$ respecting the boundary dynamics), and pulling the standard conformal structure on $\mathbb{C} \setminus \mathbb{D}_2$ back to $U \setminus K(f)$ by iterates of f. In the case of connected Julia set $J(f)$, the parameter value $c \equiv \chi(f)$ is determined uniquely.

Given a quadratic-like family $f_\lambda : V_\lambda \to U_\lambda$ over D with winding number 1, let us consider a family of straightenings (3.4) and the corresponding family of quadratic polynomials $P_{\chi(\lambda)} : z \mapsto z^2 + \chi(\lambda)$.

Note that

$$\xi = R \circ \chi | D \setminus M(\boldsymbol{f}), \tag{3.6}$$

where ξ is defined in (3.5), and R is the Riemann mapping on the complement of the Mandelbrot set. This formula follows from the definitions of ξ and χ and the description of R given in the above Example.

Lemma 3.2. — *Under the circumstances just described, the straightening*

$$\chi : (D, M(\boldsymbol{f})) \longrightarrow (\Omega_4, M)$$

is a K-qc homeomorphism of the disc D onto a neighborhood Ω_4 of the Mandelbrot set M bounded by the parameter equipotential of radius 4. The dilatation K depends only on the geometry of $(\boldsymbol{f}, \boldsymbol{h})$.

After adjusting the family $(\boldsymbol{f}, \boldsymbol{h})$, $\mathrm{Dil}(\chi)$ *will depend only on* $\mathrm{mod}(U_* \setminus V_*)$ *and* $\mathrm{Dil}(\boldsymbol{h})$. *Moreover,* $\mathrm{Dil}(\chi) \to 1$ *and* $\mathrm{mod}(D \setminus M(\boldsymbol{f})) \to \infty$ *as* $\mathrm{mod}(U_* \setminus V_*) \to \infty$ *(with a fixed* $\mathrm{Dil}(\boldsymbol{h})$*).*

Proof. — By **[DH2]**, χ is a homeomorphism. By **[L5**, Lemma 5.4], $\chi | M(\boldsymbol{f})$ admits a local qc extension χ_λ to a neighborhood N_λ of any point $\lambda \in M(\boldsymbol{f})$, with dilatation depending only on $\mathrm{mod}(\boldsymbol{f})$. Let us select neighborhoods $W_\lambda \Subset N_\lambda$. Then let us select finitely many $\chi_i \equiv \chi_{\lambda_i}$ such that the corresponding neighborhoods $W_i \equiv W_{\lambda_i}$ cover $M(\boldsymbol{f})$. By the Gluing Lemma,

$$\mathrm{Dil}(\chi | (D \setminus M(\boldsymbol{f})) \cup W_i) \le \max(\mathrm{Dil}(\chi | D \setminus M(\boldsymbol{f})),\ \mathrm{Dil}(\chi_i)).$$

Taking into account Lemma 3.1, we conclude that

$$\mathrm{Dil}(\chi) = \max_i \mathrm{Dil}(\chi | (D \setminus M(\boldsymbol{f})) \cup W_i)$$

depends only on the geometry of $(\boldsymbol{f}, \boldsymbol{h})$.

Moreover, by **[L5]**, the $\mathrm{Dil}(\chi_i) \to 1$ as $\mathrm{mod}(\boldsymbol{f}) \to \infty$. By Lemma 3.1, after the adjustment of $(\boldsymbol{f}, \boldsymbol{h})$), $\mathrm{Dil}(\chi | D \setminus M(\boldsymbol{f})) \to 1$ as $\mathrm{mod}(f_*) \to \infty$ keeping $K \equiv \mathrm{Dil}(\boldsymbol{h})$ fixed. Hence by the Gluing Lemma, $\mathrm{Dil}(\chi) \to 1$ as $\mathrm{mod}(f_*) \to \infty$.

Finally, by transversal quasi-conformality of holomorphic motions,

$$\mathrm{mod}(D \setminus M(\boldsymbol{f})) \ge K^{-1} \,\mathrm{mod}(U_* \setminus V_*) \to \infty.$$

$\square$

We will mostly deal with equipotentials of radius $4^{1/2^n}$, the preimages of the outermost equipotential of radius 4. Let us say that the equipotential of radius $4^{1/2^n}$ has *level* n, so that the outermost equipotential has level 0, the equipotential of radius 2 has level 1, etc.

3.3. Wakes and initial Markov families. — Lemma 3.2 shows that the landing properties of the parameter rays in a quadratic-like family coincide with the corresponding properties in the quadratic family. This allows us to extend the notions of the parabolic and Misiurewicz wakes from the quadratic to the quadratic-like case. Namely, the *q/p-parabolic wake* $P_{q/p} = P_{q/p}(\boldsymbol{f})$ is the parameter region in D bounded by the external rays landing at the q/p-bifurcation point $b_{q/p}$ on the main cardioid of $M(\boldsymbol{f})$ and the appropriate arc of ∂D. Dynamically it is specified by the property that for λ in this wake there are p rays landing at the α-fixed point α_λ of $J(f_\lambda)$, and they form a cycle with rotation number q/p.

The maps

$$f_\lambda^p : V_\lambda \to U_\lambda \tag{3.7}$$

restricted to appropriate domains form a (non-equipped) quadratic-like family over the wake (see **[D]**, **[L3]**, §2.5). (The domain V_λ is a thickening of the puzzle piece $Y_\lambda^{(1+p)}$ bounded by two pairs of rays landing at the α-fixed and co-fixed (i.e., the other preimage of α) points and two equipotential arcs. The domain U_λ is a thickening of the puzzle piece $Y_\lambda^{(1)}$ bounded by two rays landing at the α-fixed point and an equipotential arc of level 1.) Note however that this family fails to be proper as the domains U_λ don't admit continuous extension at the root.

***Proposition 3.3* (see [D]).** — *Let $\boldsymbol{f}$ be a DH quadratic-like family with winding number* 1. *Then the winding number of the critical value $\lambda \mapsto f_\lambda^p(0)$ about* 0 *when λ wraps once about the boundary of the parabolic wake $\partial P_{q/p}$ is also equal to* 1.

By **[DH2, D]**, the quadratic-like family (3.7) generates a homeomorphic copy $M_{q/p} = M_{q/p}(\boldsymbol{f})$ of the Mandelbrot set attached to the bifurcation point $b_{q/p}$. Its complement $M \setminus M_{q/p}$ consists of a component containing the main cardioid and infinitely many *decorations* (using terminology of Dierk Schleicher **[Sch]**) $D_{q/p}^{\sigma,i}$, where σ is a dyadic sequence of length $|\sigma| = t-1$, $t = 1, 2, \dots$, $i = 1, \dots, p-1$. The decoration $D_{q/p}^{\sigma,i}$ touches $M_{q/p}$ at a Misiurewicz point $\mu = \mu_{q/p}^\sigma$ for which

$$f_\mu^{pk}(0) \in Y_\mu^{(1+p)},\ k = 0, \dots, t-1, \ \text{ while } \ f_\mu^{pt}(0) = \alpha'_\mu,$$

where α'_μ is the α-co-fixed point. (Such Misiurewicz points are naturally labeled by the dyadic sequences).

Every decoration $D_{q/p}^{\sigma,i}$ belongs to the *Misiurewicz wake* $\widehat{O}_{q/p}^{\sigma,i}$ *of level t* bounded by two parameter rays landing at $\mu_{q/p}^\sigma$ (there are p rays landing at this point). Let us truncate such a wake by the equipotential of level pt. We will obtain the initial puzzle pieces $O_{q/p}^{\sigma,i}$ which sometimes will also be called "Misiurewicz wakes". They can be dynamically specified in terms of the initial puzzle (see **[L3]**, §3.2). Namely, there are $p-1$ puzzle pieces $Z_i^{(1)}$, $i = 1, \dots, p-1$, attached to the co-fixed point α'. Pulling them back by $(t-1)$-st iterate of the double covering $f^p : Y^{(1+p)} \to Y^{(1)}$, we obtain 2^{t-1} puzzle pieces $Z_{\sigma,i}^{(1+(t-1)p)}$ labeled by the dyadic sequences. The wake $O_{q/p}^{\sigma,i}$ is specified by the property that $f^p 0 \in Z_{\sigma,i}^{(1+(t-1)p)}$.

The wake $O_{q/p}^{\sigma,i}$ containing a point λ will also be denoted by $O(\lambda)$.

By *tiling* we mean a family of topological discs with disjoint interiors. Let us consider the initial tiling constructed in **[L3]**, §3.2 (see Figure 2):

$$Y_\lambda^{(0)} \supset V_\lambda^0 \cup \bigcup_{k>0} \bigcup_i X_{i,\lambda}^k \cup \bigcup_{k\geq 0} \bigcup_j Z_{j,\lambda}^{(1+kp)}, \tag{3.8}$$

where $V_\lambda^0 \equiv X_{0,\lambda}^0$.

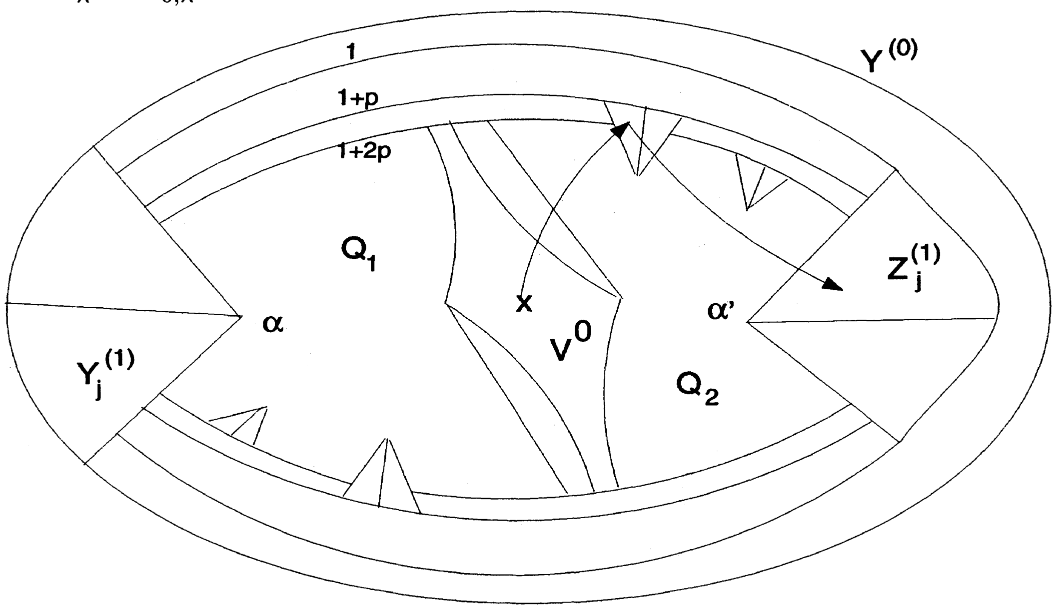


Figure 2. Initial tiling ($p = 3$, $t = 2$).

Lemma 3.4. — *The family of puzzle pieces $Y_\lambda^{(1+kp)}$ and $Z_{j,\lambda}^{(1+kp)}$, $k \leq t-1$, moves holomorphically in the region inside the parabolic wake $P_{q/p}$ bounded by the parameter equipotential of level pt with all the Misiurewicz wakes of level $\leq t-1$ removed.*

Let us recall that $\mathbb{Z}_j^{(1)}$ means $\cup_\lambda Z_{j,\lambda}^{(1)}$.

Lemma 3.5. — *Let **f** be a DH quadratic-like family with winding number* 1. *Then the initial tiling (3.8) moves holomorphically within the Misiurewicz wake $O = O_{p/q}^{\sigma,i}$. The critical value $\Phi_0(\lambda) = f_\lambda^{pt}0$, of the double covering $f_\lambda^{pt} : V_\lambda^0 \to Z_{j,\lambda}^{(1)}$ is a proper map $\Phi_0 : O \to \mathbb{Z}_j^{(1)}$ with winding number* 1 *(where $t = |\sigma| + 1$).*

Proof. — Indeed, all puzzle pieces of this initial tiling are the pullbacks of $Z_{j,\lambda}^{(1)}$. As λ ranges over the wake $O \equiv O_{q/p}^{\sigma,i}$, the corresponding iterates of 0 don't cross the boundary of $Z_{j,\lambda}^{(1)}$. It follows that the boundary of the initial tiling moves holomorphically.

Moreover, the torus $\delta\mathbb{Z}_j^{(1)}$ is foliated by the curves with the same external coordinates, and one curve corresponding to the motion of the α-co-fixed point. By definition of the Misiurewicz wake, the critical value $\mathbf{\Phi}_0(\lambda)$ intersects once every leaf of this foliation when λ wraps once around ∂O. Hence $\mathbf{\Phi}_0 : O \to \mathbb{Z}_j^{(1)}$ is a proper map with winding number 1. $\square$

Let $\boldsymbol{h}_0$ stand for the above holomorphic motion of the initial tiling.

Recall for further reference that there are two puzzle pieces $Q_{1,\lambda}$ and $Q_{2,\lambda}$ in $Y_\lambda^{(1+(t-1)p)}$ which are univalently mapped by f_λ^p onto $Y_\lambda^{(1+(t-1)p)}$ (see Figure 2). The pieces $X_{i,\lambda}^k$ are the pull-backs of $V_\lambda^0 \equiv X_{0,\lambda}^0$ under the k-fold iterate of the Bernoulli map

$$f_\lambda^p : Q_{1,\lambda} \cup Q_{2,\lambda} \to Y_\lambda^{(1+(t-1)p)}.$$

Let $\Omega_{i,\lambda}^k \supset X_{i,\lambda}^k$ denote the domain in $Y_\lambda^{(1+(t-1)p)}$ which is mapped under f_λ^{kp} onto $Y_\lambda^{(1+(t-1)p)}$; in particular, $\Omega_{0,\lambda}^0 \equiv Y_\lambda^{(1+(t-1)p)}$.

3.4. First generalized quadratic-like family. — Let us consider a proper DH quadratic-like family $\boldsymbol{f} = \{f_\lambda\}$ over D with winding number 1. Fix a Misiurewicz wake O of this family. The first generalized quadratic-like map $g_{1,\lambda} : \cup V_{i,\lambda}^1 \to V_\lambda^0$ is defined as the first return map to V_λ^0 (see **[L3]**, §3.5). The itinerary of the critical point via the elements P_i of the initial tiling (3.8) determines the parameter tiling $\mathcal{D}^1$ of a Misiurewicz wake O by the corresponding puzzle pieces. Let $\Delta^1(\lambda)$ stand for such parapuzzle piece containing λ.

More precisely, for any $\lambda \in O$, let us consider the first landing map $T_\lambda : \cup L_{\bar{i},\lambda} \to V_\lambda^0$ (see **[L3]**, §11.3). The puzzle piece $L_{\bar{i},\lambda}$ is specified by its itinerary $\bar{i} = (i_0, \dots, i_{s-1})$ through non-central pieces P_i of the initial tiling until the first landing at V_λ^0:

$$L_{\bar{i},\lambda} = \{z : G_\lambda^m z \in P_{i_k},\ m = 0, \dots, s-1,\ \ G_\lambda^m z \in V_\lambda^0\}, \tag{3.9}$$

where G_λ stands for the Markov map (3.5) from **[L3]**. These tiles are organized in tubes $\mathbb{L}_{\bar{i}}$ with holomorphically moving boundary. Moreover, the first landing map induces a diffeomorphism $\boldsymbol{T} : \mathbb{L}_{\bar{i}} \to \mathbb{V}^0$ fibered over id.

Let $\bar{i}_\lambda$ stand for the itinerary of the critical value $\phi_0(\lambda) = f_\lambda^{pt}0$ through the initial tiling, so that $f_\lambda^{pt}0 \in L_{\bar{i}_\lambda,\lambda}$. Let $\mathbb{L}_* \equiv \mathbb{L}_{i_*}$ and $\Phi_0(\lambda) = (\lambda, \phi_0(\lambda))$. Then the parapuzzle pieces of the tiling $\mathcal{D}^1$ are defined as follows:

$$\Delta^1(*) = \Phi_0^{-1}\mathbb{L}_* = \{\lambda \in O : f_\lambda^{pt}0 \in L_{*,\lambda}\}.$$

Let $\mathbb{V}_j^1$ denote the components of $\boldsymbol{f}^{-pt}(\mathbb{L}_{\bar{i}}|\Delta^1(*))$ contained in $\mathbb{V}^0$, where $\mathbb{V}_0^1 \equiv \mathbb{V}^1$ is the critical component (i.e., the one containing $\mathbf{0}$). The first return map

$$\boldsymbol{g}_1 = \boldsymbol{T} \circ \boldsymbol{f}^{pt} : \cup\mathbb{V}_j^1 \longrightarrow \mathbb{V}^0 \equiv \mathbb{U}^1$$

is the desired first generalized renormalization of $\boldsymbol{f}$.

By means of the first landing map $\boldsymbol{T}$, the holomorphic motion $\boldsymbol{h}_0$ over O can be lifted to the tubes $\mathbb{L}_{\bar{i}}$. By the λ-lemma, this lift and the motion $\boldsymbol{h}_0$ of the boundary of the initial tiling (3.8) admit a common extension $\boldsymbol{H}_0$ over O.

Since the critical value $\Phi(\lambda)$ lands at the tube $\mathbb{L}_*$ as λ ranges over $\Delta^1(*)$, $\boldsymbol{H}_0$ can be lifted to a holomorphic motion of the annulus $V^0_\lambda \smallsetminus V^1_\lambda$ over $\Delta^1(*)$. Let us extend this motion to V^1_λ by the λ-lemma. This provides us with a motion $\boldsymbol{h}_1$ which equips the generalized quadratic-like family $\boldsymbol{g}_1$.

Since the winding number of $\boldsymbol{\Phi}_0$ about $\mathbb{Z}^1_j$ over O is equal to 1 (by Lemma 3.5), the function $\boldsymbol{\Phi}_0 : \Delta^1(*) \to \mathbb{L}_*$ is proper with winding number 1. Since the first landing map $\boldsymbol{T}$ is a fiberwise diffeomorphism of every tube $\mathbb{L}_{\bar{i}}$ onto $\mathbb{V}_0$, it induces a homeomorphism between the marked tori $\delta\mathbb{L}_{\bar{j}} \to \delta\mathbb{V}_0$. Hence the function $\boldsymbol{\Phi}_1(\lambda) = (\lambda, T_\lambda \circ \phi_0(\lambda))$, $\Delta^1(\lambda) \to \mathbb{V}_0$, is also proper with winding number 1. Thus we have:

Lemma 3.6. — *Let $\boldsymbol{f}$ be a DH quadratic-like family with winding number* 1. *Then the first generalized renormalization $(\boldsymbol{g}_1 : \cup\mathbb{V}^1_j \to \mathbb{V}^0 \equiv \mathbb{U}^1, \boldsymbol{h}_1)$ is a proper family with winding number* 1 *over $\Delta^1(*)$.*

Together with the tubes (3.9) let us also consider bigger tubes $\mathbb{W}_{\bar{i}}$ over O defined as follows. Let $P_{i_r,\lambda} = X^k_{j,\lambda}$ be the first "X-pieces" in the itinerary $\{P_{i_m}\}^s_{m=0}$. Then

$$W_{\bar{i},\lambda} = \{z : G^m_\lambda z \in P_{i_m,\lambda},\ m = 0,\dots,r-1,\ \ G^r_\lambda z \in \Omega^k_{j,\lambda}\}, \tag{3.10}$$

where the domains $\Omega^k_{j,\lambda}$ are defined at the end of §3.3. Moreover,

$$G^r_\lambda : W_{\bar{i},\lambda} \longrightarrow \Omega^k_{j,\lambda}, \qquad G^{s-r}_\lambda : \Omega^k_{j,\lambda} \longrightarrow Y^{(1+(t-1)p)}_\lambda, \tag{3.11}$$

and both maps are univalent isomorphisms. Thus $\mathbf{G}^s : \mathbb{W}_{\bar{i}} \to \mathbb{Y}^{(1+(t-1)p)}$ is a fiberwise conformal diffeomorphism fibered over id.

Hence the holomorphic motion of $\mathbb{Y}^{(1+(t-1)p)}$ (see Lemma 3.4) can be lifted to holomorphic motions of the $\mathbb{W}_{\bar{i}}$. Let $\mathbb{W}_* \equiv \mathbb{W}_{\bar{i}_*}$, where $\bar{i}_*$ is the itinerary of the critical value $\phi_0(*) = f^p_* 0$ through the initial tiling. Let us introduce the following parameter domains in O:

$$\Lambda^1(*) = \Phi^{-1}\mathbb{W}_* = \{\lambda : \phi_0(\lambda) \in W_{*,\lambda}\} \supset \Delta^1(*). \tag{3.12}$$

This extension of $\Delta^1(*)$ will be used for a priori bounds on the parameter geometry (see §4).

3.5. Renormalization of holomorphic families. — Let us now consider a generalized quadratic-like family $(\boldsymbol{g} : \cup\mathbb{V}_i \to \mathbb{U}, \boldsymbol{h})$ over $(D, *)$. Let $\mathcal{I}$ stand for the labeling set of tubes $\mathbb{V}_i$. Remember that $\mathcal{I} \ni 0$ and $\mathbb{V}_0 \ni \mathbf{0}$. Let $\mathcal{I}_\#$ stand for the set of all finite sequences $\bar{i} = (i_0, \dots, i_{t-1})$ of non-zero symbols $i_k \in \mathcal{I} \setminus \{0\}$. For any $\bar{i} \in \mathcal{I}_\#$, there is a tube $\mathbb{V}_{\bar{i}}$ such that

$$\boldsymbol{g}^k\mathbb{V}_{\bar{i}} \subset \mathbb{V}_{i_k},\ k = 0,\dots,t-1 \quad \textit{and} \quad \boldsymbol{g}^t\mathbb{V}_{\bar{i}} = \mathbb{U}.$$

We call $t = |\bar{i}|$ the rank of this tube. The map $\boldsymbol{g}^t : \mathbb{V}_{\bar{i}} \to \mathbb{U}$ is a holomorphic diffeomorphism which fibers over id, that is, $g^t_\lambda V_{\bar{i},\lambda} = U_\lambda$, $\lambda \in D$.

Let us redefine the holomorphic motion $\boldsymbol{h}$ of $\mathbb{U}$ as follows:

$$g^t_\lambda(\widehat{h}_{\bar{i},\lambda} z) = h_\lambda(g^t_* z), \ z \in V_{\bar{i},*} \setminus \cup_{|\bar{j}|=t+1} V_{\bar{j},*}, \quad \text{where } t = |\bar{i}|.$$

By (3.3), $\widehat{h}_\lambda$ is correctly defined on U_* minus a Cantor set. Extend it to the whole U_* by the λ-lemma.

Consider tubes $\mathbb{L}_{\bar{i}} \subset \mathbb{V}_{\bar{i}}$ such that $\boldsymbol{g}^t \mathbb{L}_{\bar{i}} = \mathbb{V}_0$, where $t = |\bar{i}|$. The first landing map $\boldsymbol{T} : \cup \mathbb{L}_{\bar{i}} \to \mathbb{V}_0$ is defined as $\boldsymbol{T}|L_{\bar{i}} = \boldsymbol{g}^t$. By construction,

$$T_\lambda(\widehat{h}_\lambda z) = \widehat{h}_\lambda(T_* z) \quad \text{for } z \in \cup \partial L_{\bar{i},*}.$$

Let $\phi(\lambda) = g_\lambda 0$ and $\boldsymbol{\Phi}(\lambda) = (\lambda, \phi(\lambda))$. Let $\bar{i}_*$ be the itinerary of the critical value $\phi(*)$ under iterates of g_* through the domains $V_{i,*}$, until its first return to $V_{0,*}$. In other words, let $g_*(0) \in \mathbb{L}_{\bar{i}_*} \equiv \mathbb{L}_*$.

Let us now consider the following parameter region around $*$:

$$D' \equiv D'(*) = \boldsymbol{\Phi}^{-1} \mathbb{L}_*.$$

For $\lambda \in D'$, the itinerary of the critical value under iterates of g_λ until the first return back to $V_{0,\lambda}$ is the same as for g_* (that is, $\bar{i}_*$). Let us define new tubes $\mathbb{V}'_j \subset \mathbb{V}_0$ as the components of $(\boldsymbol{g}|\mathbb{V}_0)^{-1}(\mathbb{L}_{\bar{i}}|D')$. Let

$$\boldsymbol{g}' : \cup \mathbb{V}'_j \to \mathbb{V}_0 | D' \equiv \mathbb{U}' \tag{3.13}$$

be the first return map of the union of these tubes to $\mathbb{V}_0$.

For $\lambda \in D'$, the critical value $\boldsymbol{\Phi}(\lambda)$ does not intersect the boundaries of the the tubes $\mathbb{L}_{\bar{i}}$. Hence we can lift the holomorphic motion $\widehat{\boldsymbol{h}}$ on $\mathbb{U} \setminus \mathbb{L}_*$ to a holomorphic motion $\boldsymbol{h}'$ on $\mathbb{U}' \setminus \mathbb{V}_0$ over D' and extend it by the λ-lemma to the whole tube $\mathbb{U}'$. Thus we obtain a generalized quadratic-like family $(\boldsymbol{g}', \boldsymbol{h}')$ over D' which will be called the *generalized renormalization* of the family $(\boldsymbol{g}, \boldsymbol{h})$ (with base point $*$).

If $\boldsymbol{g}$ is a proper family then $\boldsymbol{g}'$ is clearly proper as well. Moreover, $w(\boldsymbol{g}') = 1$ if $w(\boldsymbol{g}) = 1$. Indeed, by the Argument Principle the curve $\boldsymbol{\Phi}|D'$ intersects once every leave of $\partial \mathbb{L}_*$. Hence it has winding number 1 about this tube. As the first landing map $\boldsymbol{T} : \mathbb{L}_* \to \mathbb{V}_0$ is a fiber bundles diffeomorphism, it preserves the winding number. Thus the new critical value $\boldsymbol{\Phi}' : D' \to \mathbb{U}'$, $\boldsymbol{\Phi}' = \boldsymbol{T} \circ \boldsymbol{\Phi}$, has also winding number 1.

Let us summarize the above discussion:

Lemma 3.7. — *Let $\boldsymbol{g} : \cup \mathbb{V}_i \to \mathbb{U}$ be a generalized quadratic-like family over $(D, *)$. Assume it is proper and has winding number* 1. *Then its generalized renormalization $\boldsymbol{g}' : \cup \mathbb{V}'_j \to \mathbb{U}'$ over D' is also proper and has winding number* 1.

3.6. Central cascades. — In this section we will describe the renormalization of a generalized quadratic-like family through a central cascade, which will be then treated as a single step in the procedure of parameter subdivisions. Let us consider a holomorphic family $(\boldsymbol{g} : \cup \mathbb{V}_i \to \mathbb{U}, \boldsymbol{h})$ of generalized quadratic-like maps over $(\Delta, *)$.

We will now subdivide Δ according to the combinatorics of the central cascades of maps g_λ (see **[L3]**, §§3.1, 3.6). To this end let us first stratify the parameter values according to the length of their central cascade. This yields a nest of parapuzzle pieces

$$\Delta \equiv D \supset D' \supset \cdots \supset D^{(N)} \supset \cdots$$

For $\lambda \in D^{(N-1)} \smallsetminus D^{(N)}$, the map g_λ has a central cascade

$$V_\lambda^{(0)} \equiv U_\lambda \supset V_\lambda \equiv V_\lambda^{(1)} \supset \cdots \supset V_\lambda^{(N)} \tag{3.14}$$

of length N, so that $g_\lambda 0 \in V_\lambda^{(N-1)} \smallsetminus V_\lambda^{(N)}$. Note that the puzzle pieces $V_\lambda^{(k)}$ are organized into the tubes $\mathbb{V}^{(k)}$ over $D^{(k-1)}$.

The intersection of these puzzle pieces, $\cap D^{(N)}$, is the little Mandelbrot set $M(\boldsymbol{g})$ centered at the superattracting parameter value $c = c(\boldsymbol{g})$ such that $g_c(0) = 0$. Let us call c the *center* of D.

Let $* \in D^{(N-1)} \smallsetminus D^{(N)}$. Let us consider the Bernoulli map

$$\boldsymbol{G} : \cup \mathbb{W}_j \to \mathbb{U} \tag{3.15}$$

associated with the cascade (3.14) (see **[L3]**, §3.6 and Figure 3). Here the tubes $\mathbb{W}_j$ over $D^{(N-1)}$ are the pull-backs of the tubes $\mathbb{V}_i | D^{(N-1)}$, $i \neq 0$, by the covering maps

$$\boldsymbol{g}^k : (\mathbb{V}^{(k)} \smallsetminus \mathbb{V}^{(k+1)}) | D^{(N-1)} \longrightarrow (\mathbb{U} \smallsetminus \mathbb{V}) | D^{(N-1)}, \quad k = 0, 1, \ldots, N-1. \tag{3.16}$$

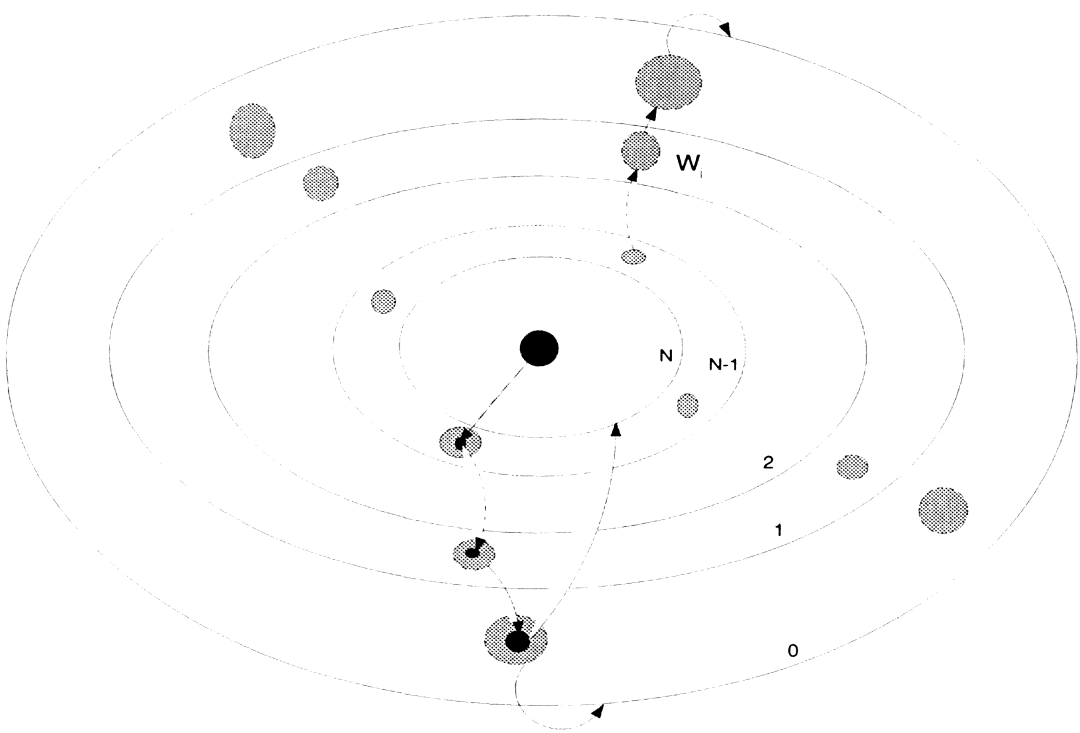


Figure 3. Solar system: Bernoulli scheme associated to a central cascade.

In the same way as in §3.5, to any string $\bar{j} = (j_0, \dots, j_{t-1})$ corresponds the tube over $D^{(N-1)}$,

$$\mathbb{W}_{\bar{j}} = \{p \in \mathbb{U}|D^{(N-1)} : \boldsymbol{G}^n p \in \mathbb{W}_{j_n}, \ n = 0, \dots, t-1\}.$$

Note that $\boldsymbol{G}^t$ univalently maps each $\mathbb{W}_{\bar{j}}$ onto $\mathbb{U}|D^{(N-1)}$. Thus $\mathbb{W}_{\bar{j}}$ contains a tube $\mathbb{L}_{\bar{j}}$ which is univalently mapped by $\boldsymbol{G}^t$ onto the central tube $\mathbb{V}^{(N)}$. These maps altogether form the first landing map to $\mathbb{V}^{(N)}$,

$$\boldsymbol{T} : \cup \mathbb{L}_{\bar{j}} \to \mathbb{V}^{(N)}. \tag{3.17}$$

Remark. — Note that

$$\operatorname{mod}(W_{\bar{j},\lambda} \smallsetminus L_{\bar{j},\lambda}) = \operatorname{mod}(U_\lambda \smallsetminus V_\lambda^{(N)}) \geq \operatorname{mod}(U_\lambda \smallsetminus V_\lambda), \tag{3.18}$$

since G_λ^t univalently maps the annulus $W_{\bar{j},\lambda} \smallsetminus L_{\bar{j},\lambda}$ onto $U_\lambda \smallsetminus V_\lambda^{(N)}$.

Let us now consider the itinerary $\bar{j}_*$ of the critical value $\phi(*) \equiv g_*(0)$ through the tubes W_j until its first return to $V^{(N)}$, so that $\boldsymbol{\Phi}(*) \in \mathbb{L}_{\bar{j}_*} \equiv \mathbb{L}_*$. Let $\mathbb{W}_* \equiv \mathbb{W}_{\bar{j}_*}$ and

$$\Delta^\diamond(*) = \boldsymbol{\Phi}^{-1}\mathbb{L}_*, \quad \Lambda^\diamond(*) = \boldsymbol{\Phi}^{-1}\mathbb{W}_*. \tag{3.19}$$

Thus the annuli $D^{(N-1)} \smallsetminus D^{(N)}$ are tiled by the parapuzzle pieces $\Delta^\diamond(\lambda)$ according as the itinerary of the critical point through the Bernoulli scheme (3.15) until the first return to $V_\lambda^{(N)}$. Altogether these tilings form the desired new subdivision of Δ. (Note however that the new tiles don't cover the whole domain Δ: the residual set consists of the Mandelbrot set $M(\boldsymbol{g})$ and of the parameter values $\lambda \in D^{(N-1)} \smallsetminus D^{(N)}$ for which the critical orbit never returns back to $V_\lambda^{(N)}$.)

The affiliated quadratic-like family over $\Delta^\diamond(*)$ is defined as the first return map to $V_\lambda^{(N)} \equiv U_\lambda^\diamond$. Its domain $\cup \mathbb{V}_i^\diamond$ is obtained by pulling back the tubes $\mathbb{L}_{\bar{j}}$ from (3.17) by the double branched covering $\boldsymbol{g} : \mathbb{V}^{(N)} \to \mathbb{V}^{(N-1)}|\Delta^\diamond(*)$, and the return map itself is just $\boldsymbol{T} \circ \boldsymbol{g}$.

The affiliated holomorphic motion is also constructed naturally. Let us first lift the holomorphic motion $\boldsymbol{h}$ from the condensator $\mathbb{U} \smallsetminus \mathbb{V}$ to the condensators $(\mathbb{V}^{(k)} \smallsetminus \mathbb{V}^{(k+1)})|D^{(N-1)}$ via the coverings (3.16). This provides us with a holomorphic motion of $(\mathbb{U} \smallsetminus \mathbb{V}^{(N)}, \cup \mathbb{W}_j)$ over $D^{(N-1)}$. Extend it through $\mathbb{V}^{(N)}$ by the λ-lemma, lift it to the tubes $(\mathbb{W}_{\bar{j}}, \mathbb{L}_{\bar{j}})$ and then extended again by the λ-lemma to the whole domain $\mathbb{U}$ over D^{N-1}. Let us denote it by $\boldsymbol{H}$.

Lifting this motion via the fiberwise analytic double covering over $\Delta^\diamond(*)$,

$$\boldsymbol{g} : \Big(\mathbb{U}^\diamond \smallsetminus \mathbb{V}^\diamond, \bigcup_{i \neq 0} \mathbb{V}_i^\diamond\Big) \longrightarrow \Big(\mathbb{V}^{(N-1)} \smallsetminus \mathbb{L}_*, \bigcup_{\bar{j} \neq \bar{j}_*} \mathbb{L}_{\bar{j}}\Big),$$

we obtain the desired motion of $(\mathbb{U}^\diamond \smallsetminus \mathbb{V}^\diamond, \bigcup_{i \neq 0} \mathbb{V}_i^\diamond)$ over $\Delta^\diamond(*)$. By the λ-lemma it extends through $\mathbb{V}_0^\diamond$.

3.7. Principal parapuzzle nest. — Let us now summarize the above discussion. Given a quadratic-like family $(\boldsymbol{f}, \boldsymbol{h})$ over $D \equiv \Delta^0$, we consider the first tiling $\mathcal{D}^1$ of a Misiurewicz wake O as described in §3.4. Each tile $\Delta \in \mathcal{D}^1$ comes together with a generalized quadratic-like family $(\boldsymbol{g}_\Delta, \boldsymbol{h}_\Delta)$ over Δ.

Now assume inductively that we have constructed the tiling $\mathcal{D}^l$ of level l. Then the tiling of the next level, $\mathcal{D}^{l+1}$ is obtained by partitioning each tile $\Delta \in \mathcal{D}^l$ by means of the cascade renormalization as described in §3.6.

Let $\Delta^l(\lambda)$ stand for the tile of $\mathcal{D}^l$ containing λ, while $\Delta^l(\lambda) \subset \Lambda^l(\lambda) \subset \Delta^{l-1}(\lambda)$ stand for the other tile defined in (3.19). Each tile $\Delta = \Delta^l(\lambda)$ contains a *central subtile* $\Pi^l(\lambda) = \boldsymbol{\Phi}_\Delta^{-1}\mathbb{V}_0$ corresponding to the central return of the critical point (here $\boldsymbol{\Phi}_\Delta(\lambda) = (\lambda, \boldsymbol{g}_\Delta(\lambda)))$. Note that $\Pi^l(\lambda)$ may or may not contain λ itself.

Let us then consider the sequence of renormalized families $(\boldsymbol{g}_{l,\lambda}, \boldsymbol{h}_{l,\lambda})$ over topological discs $\Delta^l(\lambda)$. We call the nest of topological discs $\Delta^0 \supset \Delta^1(\lambda) \supset \Delta^2(\lambda) \supset \cdots$ (supplied with the corresponding families) *the principal parapuzzle nest* of λ. This nest is finite if and only if λ is renormalizable.

Let $c_{l,\lambda} \in \Delta^l(\lambda)$ be the centers of the corresponding parapuzzle pieces. Let us call them the *principal superattracting approximations* to λ. If λ is not renormalizable, then $c_{l,\lambda} \to \lambda$ as $l \to \infty$, since $\operatorname{diam} \Delta^l(\lambda) \to 0$ (see the next section).

The $\operatorname{mod}(\Delta^l(\lambda) \smallsetminus \Delta^{l+1}(\lambda))$ are called the *principal parameter moduli* of $\lambda \in D$.

When we fix a base point $*$, we will usually skip label $*$ in the above notations, so that $\Delta^l \equiv \Delta^l(*)$, $\boldsymbol{g}_l \equiv \boldsymbol{g}_{l,*}$, $\boldsymbol{h}_l \equiv \boldsymbol{h}_{l,*}$ etc.

4. Parapuzzle geometry

The following is the main geometric result of this paper:

Theorem A. — *Let us consider a proper DH quadratic-like family $(\boldsymbol{f}, \boldsymbol{h})$ with winding number 1 over D, and a Misiurewicz wake $O \subset D$. Then for any $\lambda \in M(\boldsymbol{f}) \cap O$,*

$$\operatorname{mod}(\Delta^l(\lambda) \smallsetminus \Delta^{l+1}(\lambda)) \geq Bl, \quad \textit{and} \quad \operatorname{mod}(\Delta^l(\lambda) \smallsetminus \Pi^l(\lambda)) \geq Bl,$$

where the constant $B > 0$ depends only on O and $\operatorname{mod}(\boldsymbol{f})$.

The rest of this section will be devoted to the proof of this theorem.

4.1. Initial parameter geometry. — In this section we will give a bound on the geometry of the first level parapuzzle. Fix a quadratic-like family $(\boldsymbol{f}, \boldsymbol{h})$ and its Misiurewicz wake $O = O_{q/p}^{\sigma,i}$, $|\sigma| = t$, as in §3.3. In what follows we will use the notations of §3.3 and §3.4.

Lemma 4.1. — *There is a marked holomorphic motion of any annulus $\Omega_{i,\lambda}^k \smallsetminus X_{i,\lambda}^k$ over O with dilatation depending on the geometry of $(\boldsymbol{f}, \boldsymbol{h})$ and the choice of O only.*

Proof. — Indeed, by Lemma 3.4, the configuration

$$\Big(\mathbb{Y}^0, \bigcup_{k\leq t-1} \partial\mathbb{Y}^{(1+kp)}, \bigcup_{k\leq t-1}\bigcup_j \mathbb{Z}_j^{(1+kp)}\Big)$$

represents a holomorphic motion over the parabolic wake $P_{q/p}$ (truncated by the appropriate equipotential) with removed Misiurewicz wakes of level $\leq t-1$. Since O is compactly contained in this region, this holomorphic motion has a finite dilatation $K = K(O)$ over O (depending only on O and the geometry of $(\boldsymbol{f}, \boldsymbol{h})$).

Let us now lift this motion to the tubes $\mathbb{Z}_j^{(1+tp)}$ and $\mathbb{Q}_1$, $\mathbb{Q}_2$ by the map $\boldsymbol{f}^p$. Since this map is a fiberwise diffeomorphism over O, it preserves the dilatation of the motion (though the motion does not extend beyond O any more). Similarly we can lift the motion to all quadrilaterls between the equipotentials of level $1+(t-1)p$ and $1+tp$ left after removing Z-pieces of level $\leq t-1$ (see Figure 2). This provides us with a marked motion $\boldsymbol{h}_0$ of the annulus $Y_\lambda^{(1+(t-1)p)} \smallsetminus V_\lambda^0$ over O with the same dilatation K. This handles the case $k=0$.

For $k>0$, lift $\boldsymbol{h}_0$ by the following fiberwise diffeomorphism over O:

$$f_\lambda^{pk} : (\Omega_{i,\lambda}^k, X_{i,\lambda}^k) \longrightarrow (Y_\lambda^{(1+(t-1)p)}, V_\lambda^0).$$

□

Lemma 4.2. — *All parapuzzle pieces of the first level are well inside the corresponding wake:* $\mathrm{mod}(O \smallsetminus \Delta^1) \geq \nu > 0$. *Moreover, the holomorphic motion $\boldsymbol{h}_1$ of the condensator $\mathbb{U}^1 \smallsetminus \mathbb{V}^1$ over Δ^1 is K-qc. The constants ν and K depend only on the geometry of $(\boldsymbol{f}, \boldsymbol{h})$ and the choice of O.*

Proof. — Let us consider the tubes $\mathbb{L}_* \subset \mathbb{W}_*$ over O constructed in §3.4. By means of the fiberwise conformal diffeomorphism (3.11)

$$G_\lambda^r : W_{*,\lambda} \smallsetminus L_{*,\lambda} \to \Omega_{i,\lambda}^k \smallsetminus X_{i,\lambda}^k$$

we can lift the motion constructed in Lemma 4.1 to the condensator $\mathbb{W}_* \smallsetminus \mathbb{L}_*$. Since the dilatation of the motion under such a lift is preserved, it depends only on the geometry of $(\boldsymbol{f}, \boldsymbol{h})$ and the choice of O.

Let us now consider the parameter annulus

$$\Lambda^1 \smallsetminus \Delta^1 = \Phi^{-1}(\mathbb{W}_* \smallsetminus \mathbb{L}_*) \subset O$$

(see (3.12)). By transverse quasi-conformality of holomorphic motions (Corollary 2.1),

$$\mathrm{mod}(\Lambda^1 \smallsetminus \Delta^1) \geq K^{-1}\,\mathrm{mod}(W_* \smallsetminus L_*),$$

where $L_* \smallsetminus W_*$ is the $*$-fiber of the condensator $\mathbb{W}_* \smallsetminus \mathbb{L}_*$.

But G_*^s univalently maps $W_* \smallsetminus L_*$ onto $Y_*^{(1+(t-1)p)} \smallsetminus V_*^0$ (see (3.10)-(3.11)). Since by [**L3**, §4.1], the modulus of the latter annulus depends only on the geometry of $(\boldsymbol{f}, \boldsymbol{h})$ and the choice of O, the first statement of the lemma follows.

Moreover, the motion $\boldsymbol{h}_1$ on $\mathbb{U}^1 \setminus \mathbb{V}^1$ is a double covering of the motion $\boldsymbol{H}_0$ on $\mathbb{Z}_j^{(1)} \setminus \mathbb{L}_*$ over O (see §3.4). Since Δ^1 is well inside O, the λ-lemma implies that $\boldsymbol{H}_0$ has a bounded dilatation over Δ^1 (depending only on the geometry of $(\boldsymbol{f}, \boldsymbol{h})$ and the choice of O). Hence the dilatation of $\boldsymbol{h}_1$ on $\mathbb{U}^1 \setminus \mathbb{V}^1$ over Δ^1 is bounded as well. □

4.2. Inductive estimate of the parameter geometry

Lemma 4.3. — *Let us consider a generalized quadratic-like family $(\boldsymbol{g} : \cup \mathbb{V}_i \to \mathbb{U}, \boldsymbol{h})$ over Δ. Assume that the dilatation of $\boldsymbol{h}$ on $\mathbb{U} \setminus \mathbb{V}_0$ is bounded by K and*

$$\operatorname{mod}(U_* \setminus V_{0,*}) \geq \mu > 0, \quad \lambda \in D.$$

Then the dilatation of the cascade renormalized motion $\boldsymbol{h}^\diamond$ on $\mathbb{U}^\diamond \setminus \mathbb{V}_0^\diamond$ over $D^\diamond$ (as described in §3.6) is bounded by $K^\diamond = K^\diamond(\mu, K)$.

Proof. — We will use the notations of §3.6. We assume that f_* has a central cascade (3.14) of length N, so that $* \in D^{(N-1)} \setminus D^{(N)}$. The holomorphic motion $\boldsymbol{h}$ on $\mathbb{U} \setminus \mathbb{V}_0$ can be lifted to a motion $\boldsymbol{H}$ on $\mathbb{W}_* \setminus \mathbb{L}_*$ by a fiberwise conformal diffeomorphism $\boldsymbol{T}$ (an extension of the first landing map (3.17)). Hence this motion has dilatation K on the tube $\mathbb{W}_* \setminus \mathbb{L}_*$. By transverse quasi-conformality of holomorphic motions,

$$\mu^\diamond \equiv \operatorname{mod}(\Lambda^\diamond \setminus \Delta^\diamond) \geq K^{-1}\mu. \tag{4.1}$$

Let us extend $\boldsymbol{H}$ by the λ-lemma through $\mathbb{L}_*$ to a motion on the whole tube $\mathbb{W}_*$ over $\Lambda^\diamond$. Applying the λ-lemma again, we conclude that this motion has dilatation $K^\diamond = K^\diamond(\mu^\diamond)$ over $\Delta^\diamond$. But the motion $\boldsymbol{h}^\diamond$ on $\mathbb{U}^\diamond \setminus \mathbb{V}_0^\diamond$ is the lift of $\boldsymbol{H}$ on $\mathbb{U} \setminus \mathbb{L}_*$ over $\Delta^\diamond$ by the fiberwise conformal double covering $\boldsymbol{g}$. Hence it has the same dilatation $K^\diamond$. □

4.3. Inscribing rounds condensators. — In this section we will show that the parameter annuli have definite moduli. Given a holomorphic motion h_λ and a holomorphic family of affine maps $g_\lambda : z \mapsto a_\lambda z + b_\lambda$, we can consider an "affinely equivalent" motion $g_\lambda \circ h_\lambda$. In this way the motion can be normalized such that any two points $z, \zeta \in U_*$ don't move (that is, $h_\lambda(z) \equiv z$ and $h_\lambda(\zeta) \equiv \zeta$ for $\lambda \in D$). Let us start with a technical lemma:

Lemma 4.4. — *Let us consider a holomorphic motion $h : (U_*, V_*, 0) \to (U_\lambda, V_\lambda, 0)$ of a pair of nested topological discs over a domain D. Assume that the maps $h_\lambda : (\partial U_*, \partial V_*) \to (\partial U_\lambda, \partial V_\lambda)$ admit K-qc extensions $H_\lambda : (\mathbb{C}, U_*) \to (\mathbb{C}, U_\lambda)$ (not necessarily holomorphic in λ but with uniform dilatation K). Then there exists an $M = M(K)$ such that if* $\operatorname{mod}(U_* \setminus V_*) > M$ *then after appropriate normalization of the motion, there exists a round condensator $D \times \mathbb{A}(q, 2q)$ embedded into $\mathbb{U} \setminus \mathbb{V}$.*

Proof. — Let z_* be a point on ∂U_* closest to 0. Normalize the motion in such a way that $z_* = 1$, and this point does not move. With this normalization, $V_* \subset \mathbb{D}_\varepsilon(0)$ where $\varepsilon = \varepsilon(m) \to 0$ as $m \equiv \operatorname{mod}(U_* \setminus V_*) \to \infty$.

Since the space of normalized K-qc maps is compact, $|H_\lambda(\varepsilon e^{i\theta}|) < \delta$, where $\delta = \delta(\varepsilon, K) \to 0$ as $\varepsilon \to 0$, K being fixed, and $|H_\lambda(e^{i\theta})| > r$ where $r = r(K) > 0$. It follows that the domain $\mathbb{U} \setminus \mathbb{V}$ contains the round cylinder $D \times \mathbb{A}(\delta, r)$, and we are done. $\square$

Corollary 4.5. — *Under the circumstances of Lemma 4.4, let $\mathbf{\Phi} : D \to \mathbb{U}$ be a proper analytic map with winding number* 1. *Let $D' = \mathbf{\Phi}^{-1}\mathbb{V}$. If* $\mathrm{mod}(U_* \setminus V_*) > M = M(K)$ *then* $\mathrm{mod}(D \setminus D') \geq \log 2$.

Proof. — By Lemma 4.4, $\mathbb{U}\setminus\mathbb{V} \supset D\times A$ where $A = \mathbb{A}(q, 2q)$. Let $Q = \mathbf{\Phi}^{-1}(D\times A)$. By the Argument Principle, $\phi = \pi_2 \circ \mathbf{\Phi}$ univalently maps Q onto A, so that $\mathrm{mod}(D\setminus D') \geq \mathrm{mod}\, Q = \mathrm{mod}\, A = \log 2$. $\square$

4.4. Puzzle geometry. — Let us now recall for reader's convenience two key results of **[L3]**, which will be used below.

***Theorem 4.6* (Moduli growth [L3], Theorem III).** — *Let f be a quadratic-like map whose straightening $c = \chi(f)$ belongs to a Misiurewicz wake O. Let $n(k)$ be the non-central levels of its principal nest $V^0 \supset V^1 \supset \cdots$. Then*

$$\mathrm{mod}(V^{n(k)+1} \setminus V^{n(k)+2}) \geq Bk,$$

where B depends only on O and $\mathrm{mod}(f)$.

Remark. — A related result on moduli growth for *real* parameter values was independently proven by Graczyk & Swiatek **[GS]**. Note in this respect that the proof of our parameter result (Theorem A) needs in a crucial way the above Theorem 4.6 with complex parameter values (even if one is ultimately interested in the real case).

Let us consider a quadratic-like family $(\boldsymbol{f}, \boldsymbol{h})$ and its parameter tilings. Let $\widetilde{\lambda} \in \Delta^l(\lambda)$. Let us consider the corresponding l-fold generalized renormalizations of these two maps $g_l : \cup V_i \to U$ and $\widetilde{g}_l : \cup \widetilde{V}_i \to \widetilde{U}$. Then the holomorphic motion $\boldsymbol{h}$ transforms the domains of g_l to the corresponding domains of $\widetilde{g}_l$ respecting the boundary marking (coming from the external coordinate system, see §3.2). In this sense f_λ and $f_{\widetilde{\lambda}}$ have "the same combinatorics up to level" l.

Let us say that g_l and $\widetilde{g}_l$ K-qc pseudo-conjugate if there is a K-qc homeomorphism

$$h : (U, \cup V_i) \to (\widetilde{U}, \cup \widetilde{V}_i),$$

respecting the boundary marking. Thus it matches with the boundary holomorphic motion, and hence respects the boundary dynamics: $h(g_l z) = \widetilde{g}_l(hz)$ for $z \in \cup \partial V_i$.

***Theorem 4.7* (Uniformly qc pseudo-conjugacies [L3, §11]).** — *Assume that $\widetilde{\lambda} \in \Lambda^{l+1}(\lambda)$, where the tile $\Lambda^{l+1}(\lambda)$ is defined by (3.19). Then the corresponding generalized renormalizations g and $\widetilde{g}$ are K-qc pseudo-conjugate, with K depending only on the Misiurewicz wake $O(\lambda)$ and geometry of $(\boldsymbol{f}, \boldsymbol{h})$.*

Remarks

(1) Concerning the assumption $\widetilde{\lambda} \in \Lambda^{l+1}(\lambda)$, see **[L3]**, Remarks on pp. 272 and 277.

(2) In **[L3]** (see §4) the initial choice of straightening of two maps f and $\widetilde{f}$ is made independently and its dilatation depends only on the mod f and mod $\widetilde{f}$. In the above formulation, the choice should be consistent with the holomorphic motion $\boldsymbol{h}$, so that its dilatation depends on the geometry of $(\boldsymbol{f}, \boldsymbol{h})$.

4.5. Uniform bound of dilatation

Lemma 4.8. — *Let $*$ belong to a Misiurewicz wake O. For any principal parapuzzle piece $\Delta = \Delta^{l+1}(*)$, the corresponding holomorphic motion $\boldsymbol{h}_\Delta$ of $\mathbb{U}^{l+1} \smallsetminus \mathbb{V}_0^{l+1}$ over Δ has a uniformly bounded dilatation, depending only on the choice of the Misiurewicz wake O and the geometry of $(\boldsymbol{f}, \boldsymbol{h})$.*

Proof. — Let K be a dilatation bound given by Theorem 4.7. Find an $M = M(K)$ by Corollary 4.5. By Theorem 4.6 and (3.18), there exists an l_0 such that $\operatorname{mod}(W_*^l \smallsetminus L_*^l) \geq M$ for $l \geq l_0$.

For $l \leq l_0$, the desired dilatation bound is guaranteed by Lemmas 4.2 and 4.3.

Fix an $l \geq l_0$. Consider the generalized quadratic-like family $(\boldsymbol{g} : \cup \mathbb{V}_i \to \mathbb{U}, \boldsymbol{h})$ over $D \equiv \Delta^l(*)$. In what follows we will use the notations of §3.6. Let $* \in D^{(N-1)} \smallsetminus D^{(N)}$.

By Theorem 4.7, for $\lambda \in \Lambda^{l+1}(*)$, there is a K-qc pseudo conjugacy

$$\psi_\lambda : (U_*, \cup V_{i,*}) \to (U_\lambda, \cup V_{i,\lambda}),$$

with K depending only on the choice of wake O and geometry of $(\boldsymbol{f}, \boldsymbol{h})$. As we have $\operatorname{mod}(W_*^l \smallsetminus L_*^l) \geq M$, Corollary 4.5 can be applied. We conclude that

$$\operatorname{mod}(\Lambda^{l+1}(*) \smallsetminus \Delta^{l+1}(*)) \geq \log 2 \tag{4.2}$$

for l sufficiently big (depending on O and geometry of $(\boldsymbol{f}, \boldsymbol{h})$).

In §3.6 we have constructed a holomorphic motion $\boldsymbol{H}$ of $(\mathbb{U}, \mathbb{W}_{\bar{j}}, \mathbb{L}_{\bar{j}})$ over $D^{(N-1)}$. By the λ-lemma and (4.2), $\boldsymbol{H}$ is L-qc over $\Delta = \Delta^{l+1}(*)$, with an absolute L provided l is big enough. But the holomorphic motion $\boldsymbol{h}_\Delta$ on $\mathbb{U}^{l+1} \smallsetminus \mathbb{V}^{l+1}$ is the lift of $\boldsymbol{H}$ on $\mathbb{V}^{(N-1)} \smallsetminus \mathbb{L}_*$ over Δ by means of the fiberwise analytic double covering

$$\boldsymbol{g} : \mathbb{U}^{l+1} \smallsetminus \mathbb{V}^{l+1} \to \mathbb{V}^{(N-1)} \smallsetminus \mathbb{L}_*.$$

Hence $\boldsymbol{h}_\Delta$ on $\mathbb{U}^{l+1} \smallsetminus \mathbb{V}^{l+1}$ is also L-qc. □

4.6. Proof of Theorem A. — We are now prepared to complete the proof:

$$\operatorname{mod}(\Delta^l \smallsetminus \Delta^{l+1}) \geq K^{-1} \operatorname{mod}(W_{\bar{i}_*} \smallsetminus L_{\bar{i}_*}) \geq Bl.$$

The first estimate in the above row follows from Lemma 4.8 and Corollary 2.1. The last estimate is due to Theorem 4.6.

For the same reason,

$$\operatorname{mod}(\Delta^l \smallsetminus \Pi^l) \asymp \operatorname{mod}(U^l_* \smallsetminus V^l_{0,*}) \geq Bl. \qquad \square$$

Remark. — The author thanks the referee for the following note. Instead of exploiting holomorphic motions in the above proof, one can use (for high levels l) a refined version of Lemma 4.4. Indeed, the proof of this lemma shows that one can inscribe into $\mathbb{U} \smallsetminus \mathbb{V}$ a round condensator of modulus comparable with $\operatorname{mod}(U_* \smallsetminus V_*)$ (with a constant depending on the dilatation K only). This provides us with inscribed round condensators over principal parapuzzle pieces with linearly growing moduli, which yields linearly growing parameter moduli.

Thus, formally speaking, holomorphic motions are needed only on the initial levels. However, their actual role is more significant as their transversal quasi-conformality is a true mechanism behind commensurability of the dynamical and parameter geometries.

5. Application to the measure problem

In this section we will apply the previous results to the real quadratic family $P_c : z \mapsto z^2 + c$, $c \in \mathbb{R}$. Let $\mathcal{NR}$ stand for the set of non-renormalizable real parameter values $c \in [-2, -3/4)$. Note that all periodic points of the $P_c : z \mapsto z^2 + c$, $c \in \mathcal{NR}$, are repelling. Indeed, the interval $[-3/4, 1/4]$ where P_c has a non-repelling fixed point is excluded, while maps with non-repelling cycles of higher period are renormalizable.

Let $\mathcal{NC}$ stand for the set of parameter values $c \in \mathcal{NR}$ such that the principal nest of P_c contains only finitely many non-trivial (i.e., of length > 1) central cascades.

Theorem 5.1

- *The set $\mathcal{NR}$ has positive measure;*
- *The set $\mathcal{NC}$ has full Lebesgue measure in $\mathcal{NR}$.*

Remarks

(1) The former (positive measure) result is known (see **[BC]**, **[J]**). The latter (full measure) is new.
(2) The corresponding statements concerning at most finitely renormalizable parameter values are derived from the above statements by considering quadratic-like families associated with little copies of the Mandelbrot set.
(3) By the result of Martens & Nowicki **[MN]** together with **[L2]**, P_c has an absolutely continuous invariant measure for any $c \in \mathcal{NC}$. Altogether these yield Theorem B stated in the Introduction.

Proof of Theorem 5.1. — Let d stand for the real tip of the little Mandelbrot set attached to the main cardioid (i.e. $P^3_d(0) = \alpha$). As all parameter values $c \in [d, -3/4)$

are renormalizable, we can restrict ourselves to the interval $[-2, d) \supset \mathcal{NR}$. This interval belongs to the Misiurewicz wake O attached to d.

Given measurable sets $X, Y \subset \mathbb{R}$, with $\text{length}(Y) > 0$, let $\text{dens}(X|Y)$ stand for the $\text{length}(X \cap Y)/\text{length}(Y)$.

We will now restrict all tilings $\mathcal{D}^l$ constructed above to the real line, without change of notations. We will use the same notation, $\mathcal{D}^l$, for the union of all pieces of $\mathcal{D}^l$. For every $\Delta = \Delta^l(\lambda) \in \mathcal{D}^l$, let us consider the central piece $\Pi \subset \Delta$ corresponding to the central return of the critical point. By Theorem A, $\text{dens}(\Pi|\Delta) \le Cq^l$ for absolute $C > 0$ and $q < 1$. Let Γ^l be the union of these central pieces. Summing up over all $\Delta \in \mathcal{D}^l$, we conclude that

$$\text{length}(\Gamma^l) \le \text{dens}(\Gamma^l|\mathcal{D}^l) \le Cq^l \tag{5.1}$$

(the whole interval is normalized so that its length is equal to 1).

It follows that for l sufficiently big,

$$\text{dens}\Big(\bigcup_{k\ge 0} \Gamma^{l+k}|\mathcal{D}^l\Big) \le C_1 q^l < 1,$$

which means that with positive probability central returns will never occur again. This proves the first statement.

To prove the second one just notice that (5.1) together with the Borel-Cantelli Lemma yield that infinite number of central returns occurs with zero probability. □

6. Shapes of the Mandelbrot copies

In this section we will prove Theorem C stated in the Introduction. Let us fix a quadratic-like family $(\boldsymbol{f}, \boldsymbol{h})$ and a Misiurewicz wake O in it.

Lemma 6.1. — *All maximal Mandelbrot copies in O have a bounded shape depending only on the geometry of $(\boldsymbol{f}, \boldsymbol{h})$ and the choice of O. In particular, in the quadratic family the shape depends only on the wake O.*

Proof. — Take a maximal Mandelbrot copy $M' \subset O$ centered at $*$. Let $(\boldsymbol{f}_l : \mathbb{V}^l \to \mathbb{U}^l, \boldsymbol{h}_l)$ be the DH quadratic-like family in the principal nest of $*$ generating M'. By Lemma 4.8, the dilatation of $\boldsymbol{h}_l$ on $\mathbb{U}^l \smallsetminus \mathbb{V}^l$ is bounded by a constant K depending only on the geometry of $(\boldsymbol{f}, \boldsymbol{h})$ and the choice of O. By a weak form of Theorem 4.6, $\text{mod}(\boldsymbol{f}_l) \ge \varepsilon > 0$, where ε depends on the same data only. Hence by Lemma 3.2, M' has a bounded shape depending on the same data only. □

Corollary 6.2. — *All real maximal Mandelbrot copies in $(\boldsymbol{f}, \boldsymbol{h})$, except the doubling one, have a bounded shape depending only on the geometry of $(\boldsymbol{f}, \boldsymbol{h})$. In particular, all real maximal copies in the quadratic family, except the doubling one, have a bounded shape with an absolute bound.*

Let us consider a family $\mathcal{L}$ of maximal Mandelbrot copies supplied with a combinatorial parameter $\tau = \tau(M')$, $M' \in \mathcal{L}$, satisfying the following

Big Modulus Property. — Let $M' \in \mathcal{L}$ be a Mandelbrot copy generated by a quadratic-like family over a domain Δ'. Then for $c \in \Delta'$, $\mathrm{mod}(RP_c) \geq \mu(\tau) \to \infty$ as $\tau \to \infty$.

For the family $\mathcal{L} = \mathcal{L}_O$ of maximal Mandelbrot copies contained in a wake O, several examples of such parameters are given in **[L3]** (see Theorems IV and IV′): the height, the return time, etc. For the family $\mathcal{L} = \mathcal{L}_{\mathbb{R}}$ of maximal real Mandelbrot sets, all these parameters can be unified in a single one called the *essential period* p_e (see **[L3**, §8], **[LY]**).

We say that the shapes of Mandelbrot copies M' *approach the shape of* M as $\tau(M') \to \infty$ if the M' have a (K, ε)-bounded shape with $K \to 1$, $\varepsilon \to \infty$ as $\tau \to \infty$. Lemma 3.2 implies:

Lemma 6.3. — *Assume that we have a family $\mathcal{L}$ of maximal Mandelbrot copies and a combinatorial parameter $\tau : \mathcal{L} \to \mathbb{R}$ satisfying the Big Modulus Property. Then the shapes of the Mandelbrot copies $M' \in \mathcal{L}$ approach the shape of M as $\tau(M') \to \infty$. In particular, the shapes of the $M' \in \mathcal{L}_{\mathbb{R}}$ approach M as $p_e(M') \to \infty$.*

Putting together the above statements, we obtain Theorem C.

Remark. — If the essential period $p_e(M')$ stays bounded but the period $p(M')$ grows, we obtain Mandelbrot copies near a parabolic cusp. The shapes of such copies were analysed by Douady and Devaney **[DD]**.

6.1. Relation to MLC. — Let us consider a family $\mathcal{L}$ of Mandelbrot copies. Assume that any copy M' with combinatorial type $(M_0, \dots, M_n)$, $M_i \in \mathcal{L}$, has a bounded shape. Let us say that such an $\mathcal{L}$ is fine.

Given a family $\mathcal{L}$ of Mandelbrot copies, let $E_{\mathcal{L}}$ stand for the set of infinitely renormalizable parameter values with combinatorics $(M_0, M_1, \dots)$, where $M_n \in \mathcal{L}$.

Proposition 6.4. — *Let $\mathcal{L}$ be a fine family of Mandelbrot copies. Then*

- *The Mandelbrot set is locally connected at any point $c \in E_{\mathcal{L}}$;*
- *The set $E_{\mathcal{L}}$ has zero Lebesgue measure.*

Proof

• Take a string $\tau = (M_0, M_1, \dots)$ with $M_i \in \mathcal{L}$. It determines a nest $M^0 \supset M^1 \supset \cdots$ of little Mandelbrot copies shrinking to the combinatorial class $\mathcal{C}_\tau$ of infinitely renormalizable maps with combinatorics τ. To prove MLC at a point $c \in \mathcal{C}_\tau$ we need to show that c is a single point of $\mathcal{C}_\tau$.

Let c^n stand for the center of M^n, and $H^n \ni c^n$ stand for the corresponding hyperbolic component. Let r_n be the inner radius of H^n, i.e., the radius of maximal round disk centered at c^n and inscribed into H^n. As the domains H^n are pairwise disjoint, $r_n \to 0$. Since the sets M^n have a bounded shape, $r_n \asymp \operatorname{diam} M^n$. Hence $\operatorname{diam} M^n \to 0$, so that the combinatorial class $C_\tau = \cap M^n$ consists of a single point.

• To prove the second statement, note that the hyperbolic components H^n do not belong to $E_{\mathcal{L}}$ as they are not infinitely renormalizable. Hence near any point $c \in E_{\mathcal{L}}$ there are gaps of definite relative size in arbitrary small scales. By the Lebesgue density points theorem, $\operatorname{meas}(E_{\mathcal{L}}) = 0$. □

Examples of fine families

(a) Family of maximal Mandelbrot copies M' with big height: $\chi(M') \geq \overline{\chi}$.
(b) Family of real maximal Mandelbrot copies with big period: $p(M') \geq \overline{p}$ (see **[L7]**).

Remarks

(1) We conjecture that the whole family of real maximal Mandelbrot copies is fine (so that all real Mandelbrot copies have bounded shape). We are not sure whether this is still valid for the full family of complex maximal Mandelbrot copies. This would imply MLC but it may happen that MLC is still true, though there exist very distorted Mandelbrot copies.
(2) In **[L6]** we have constructed a fine family $\mathcal{L}$ of Mandelbrot copies such that $\mathrm{HD}(E_{\mathcal{L}}) \geq 1$. Thus the Hausdorff dimension of the set of infinitely renormalizable parameter values is at least 1. (The dimension of the set of infinitely renormalizable real parameter values is at least 1/2.)

References

[A] L. Ahlfors. Lectures on quasi-conformal maps. Van Nostrand Co, 1966.

[BC] M. Benedicks & L. Carleson. On iterations of $1-ax^2$ on $(-1,1)$. Annals Math., v. 122 (1985), 1–25.

[BH] B. Branner & J.H. Hubbard. The iteration of cubic polynomials, Part II. Acta Math. v. 169 (1992), 229–325.

[BR] L. Bers & H.L. Royden. Holomorphic families of injections. Acta Math., v. 157 (1986), 259–286.

[D] A. Douady. Chirurgie sur les applications holomorphes. In: Proc. ICM, Berkeley, 1986, p. 724–738.

[DD] A. Douady & R. Devaney. Homoclinic bifurcations and infinitely many Mandelbrot sets. Preprint.

[DH1] A. Douady & J.H. Hubbard. Étude dynamique des polynômes complexes. Publication Mathématiques d'Orsay, 84-02 and 85-04.

[DH2] A. Douady & J.H. Hubbard. On the dynamics of polynomial-like maps. Ann. Sci. Éc. Norm. Sup., v. 18 (1985), 287–343.

[GS] J. Graczyk & G. Swiatek. Induced expansion for quadratic polynomials. Ann. Sci. Éc. Norm. Sup., v. 29 (1996), pp. 399–482.

[H] J.H. Hubbard. Local connectivity of Julia sets and bifurcation loci: three theorems of J.-C. Yoccoz. In: "Topological Methods in Modern Mathematics, A Symposium in Honor of John Milnor's 60th Birthday", Publish or Perish, 1993.

[J] M. Jacobson. Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. Comm. Math. Phys., v. 81 (1981), 39–88.

[L1] M. Lyubich. Some typical properties of the dynamics of rational maps, Russian Math. Surveys, v. 38:5 (1983), 154–155.

[L2] M. Lyubich. Combinatorics, geometry and attractors of quasi-quadratic maps. Ann. Math, v. 140 (1994), 347–404.

[L3] M. Lyubich. Dynamics of quadratic polynomials, I-II. Acta Math., v. 178 (1997), 185–297.

[L4] M. Lyubich. Renormalization ideas in conformal dynamics. "Current Developments in Math.", International Press, 1995.

[L5] M. Lyubich. Feigenbaum-Coullet-Tresser Universality and Milnor's Hairiness Conjecture, Ann. Math., v. 149 (1999), 319–420.

[L6] M. Lyubich. How big is the set of infinitely renormalizable quadratics? AMS Transl. (2), v. 184 (1998), 131–143.

[L7] M. Lyubich. Almost every real quadratic map is either regular or stochastic. Preprint IMS at Stony Brook, # 1997/8. To appear in Ann. Math.

[LY] M. Lyubich & M. Yampolsky. Dynamics of quadratic polynomials: Complex bounds for real maps. Ann. Inst. Fourier, v. 47 (1997), 1219–1255.

[MN] M. Martens & T. Nowicki. Invariant measures for Lebesgue typical quadratic maps. Astérisque (this volume).

[M] J. Milnor. Self-similarity and hairiness in the Mandelbrot set, "Computers in geometry and topology", Lect. Notes in Pure Appl Math, **114** (1989), 211–257.

[MSS] R. Mañé, P. Sad & D. Sullivan. On the dynamics of rational maps, Ann. scient. Éc. Norm. Sup., v. 16 (1983), 193–217.

[MvS] W. de Melo & S. van Strien. One-dimensional dynamics. Springer, 1993.

[R] M. Rees. A partial description of parameter space of rational maps of degree two, Acta. Math., v. 168 (1992), 11–87.

[Sch] D. Schleicher. The structure of the Mandelbrot set. Preprint 1995.

[Sh] M. Shishikura. The Hausdorff dimension of the boundary of the Mandelbrot set and Julia sets. Ann. Math., v. 147 (1998), 225–267.

[ST] D. Sullivan & W. Thurston. Extending holomorphic motions. Acta Math., v. 157 (1986), 243–257.

[Sl] Z. Slodkowsky. Holomorphic motions and polynomial hulls. Proc. Amer. Math. Soc., v. 111 (1991), 347–355.

[TL] Tan Lei. Similarity between the Mandelbrot set and Julia sets. Comm. Math. Phys., v. 134 (1990), 587–617.

[W] L. Wenstrom. Parameter scaling for the Fibonacci point. Preprint IMS at Stony Brook, # 1996/4.

M. Lyubich, Institute for Math. Sciences, SUNY at Stony Brook, NY 11794, USA
E-mail : `mlyubich@math.sunysb.edu` • *Url :* `http://www.math.sunysb.edu/~mlyubich`

STEFANO LUZZATTO
MARCELO VIANA

Positive Lyapunov exponents for Lorenz-like families with criticalities

Astérisque, tome 261 (2000), p. 201-237
<http://www.numdam.org/item?id=AST_2000__261__201_0>

Astérisque
261, 2000, p. 201–237

POSITIVE LYAPUNOV EXPONENTS FOR LORENZ-LIKE FAMILIES WITH CRITICALITIES

by

Stefano Luzzatto & Marcelo Viana

Dedicated to Adrien Douady on the occasion of his 60^{th} birthday

Abstract. — We introduce a class of one-parameter families of real maps extending the classical geometric Lorenz models. These families combine singular dynamics (discontinuities with infinite derivative) with critical dynamics (critical points) and are based on the behaviour displayed by Lorenz flows over a fairly wide range of parameters. Our main result states that – nonuniform – expansion is the prevalent form of dynamics even after the formation of the criticalities.

1. Introduction and statement of results

Numerical analysis of the now famous system of differential equations

$$(1)\qquad \begin{cases} \dot{x} = -\sigma x + \sigma y \\ \dot{y} = rx - y - xz \\ \dot{z} = -bz + xy \end{cases}$$

for parameter values $r \approx 28$, $\sigma \approx 10$, $b \approx 8/3$, led Lorenz **[11]** to identify sensitive dependence of orbits with respect to the corresponding initial points as a main source of unpredictability in deterministic dynamical systems. His observations were then interpreted by **[1]**, **[6]**, who described expanding ("strange") attractors in certain geometric models for the behaviour of (1). Conjecturedly, such an attractor exists also for Lorenz' original equations, although this has not yet been proved.

1991 ***Mathematics Subject Classification.*** — 58F03, 58F13.
Key words and phrases. — Lorenz, positive Lyapunov exponents, Lorenz-like maps and flows.

S. Luzzatto is partially supported by EEC Contract No. ERBCHBGCT 920016 and by a CAPES/ITALIA fellowship. He is also thankful to IMPA-CNPq/Brazil for its kind hospitality and support during the preparation of this paper. M. Viana is partially supported by PRONEX Project "Dynamical Systems", Brazil.

Further study of (1) revealed that relatively small variations of these parameter values may lead to quite different, albeit even more complex, dynamical features. Indeed, already as r is increased past $r \approx 30$ Poincaré return maps cease to be described by the cusp-type pictures corresponding to the geometric models, instead they exhibit "folded cusps", or "hooks"; moreover, these hooks persist in a large window of values of r (extending beyond $r \approx 50$), see **[15]** for a thorough discussion. Trying to understand this folding process and its effect on the behaviour of the flow was, in fact, a main motivation behind Hénon's model of strange attractor for maps in two dimensions, **[7]**, **[8]**. In constructing this model he focused on the dynamics near the fold, in particular disregarding trajectories which pass close to equilibrium points.

Here we aim at a more global understanding of the dynamics of Lorenz flows, accounting for the interaction between *singular behaviour* (corresponding to trajectories near equilibria) and *critical behaviour* (near folding regions). Indeed, we introduce a one-dimensional prototype for this problem, largely inspired by the observations in **[15]**, which we call Lorenz-like families with criticalities. Apart from their present motivation, these families of maps are also of interest in their own right, as models of rich nonsmooth dynamics in dimension one. Moreover, in an ongoing work we are further pushing the present constructions and conclusions to the context of smooth flows in three-dimensional space, *cf.* comments below.

Let us begin by explaining what we mean by *Lorenz-like families with criticalities*. We consider one-parameter families $\{\varphi_a\}$ of real maps of the form

$$\varphi_a(x) = \begin{cases} \varphi(x) - a & \text{if } x > 0 \\ -\varphi(-x) + a & \text{if } x < 0 \end{cases}$$

where $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ is smooth and satisfies:

L1 : $\varphi(x) = \psi(x^\lambda)$ for all $x > 0$, where $0 < \lambda < 1/2$ and ψ is a smooth map defined on $\mathbb{R}$ with $\psi(0) = 0$ and $\psi'(0) \neq 0$;

L2 : there exists some $c > 0$ such that $\varphi'(c) = 0$;

L3 : $\varphi''(x) < 0$ for all $x > 0$.

As we already mentioned, this definition is motivated by a fair amount of numerical and analytical data concerning the behaviour of Lorenz flows. In particular, the condition $\lambda < 1/2$ corresponds to the fact that, for the parameter region we are interested in, the expanding eigenvalue λ_u of (1) at the origin is more than twice stronger than the weakest contracting eigenvalue λ_s (that is $\lambda_u + 2\lambda_s > 0$).

For small values of the parameter the maximal invariant set of φ_a in the interval $[-a, a]$ is a hyperbolic Cantor set. Under certain natural conditions, implied by L4 and L5 below, the entire interval $[-a, a]$ becomes forward invariant as a crosses some value $a_1 > 0$. This situation persists for a certain range of parameter values and corresponds to the class of maps usually associated to the "Lorenz attractor" (see **[5]**, **[6]**, **[1]**). The dynamics of such maps is relatively well understood: they admit an

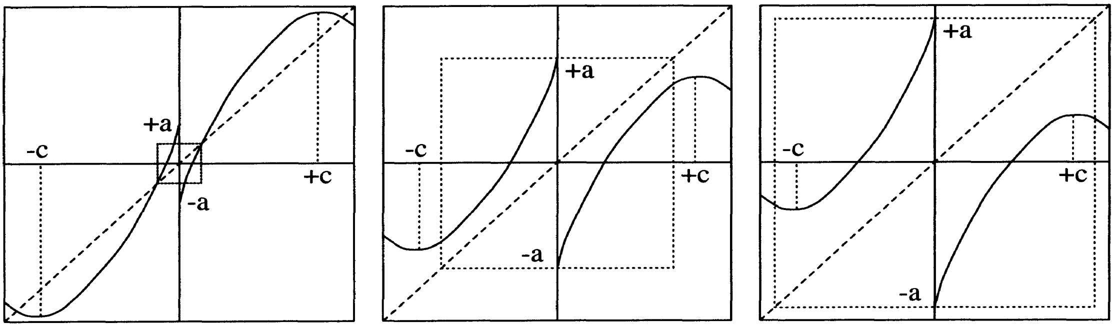


FIGURE 1. Lorenz-like families with criticalities

invariant measure which is absolutely continuous with respect to Lebesgue and has positive metric entropy; they are not structurally stable but are fully persistent in the sense that any small perturbation also admits an absolutely continuous invariant measure of positive entropy.

We are mainly interested in studying the bifurcation which occurs as the parameter crosses the value $a = c$. With this in mind, we add two natural assumptions on φ which ensure that a Lorenz attractor persists for all $a < c$.

Let $x_{\sqrt{2}}$ denote the unique point in $(0, c)$ such that $\varphi'(x_{\sqrt{2}}) = \sqrt{2}$; sometimes we also write $a_2 = x_{\sqrt{2}}$. Then we suppose

L4 : $0 < \varphi_a(x_{\sqrt{2}}) < \varphi_a(a) < x_{\sqrt{2}}$ for all $a \in (a_2, c]$.

The last inequality implies that given any y with $|y| \in [x_{\sqrt{2}}, a)$ there exists a unique $x \in [-a, a]$ such that $\varphi_a(x) = y$. Note that x and y have opposite signs. Moreover, the first inequality implies that $|x| < x_{\sqrt{2}}$. Our last assumption is

L5 : $|(\varphi_c^2)'(x)| > 2$ for all $x \in [-c, c] \setminus \{0\}$ such that $|\varphi_c(x)| \in [x_{\sqrt{2}}, c]$.

Observe that this is automatic if $\varphi_c(x) = x_{\sqrt{2}}$ (because $|x|$ is strictly smaller than $x_{\sqrt{2}}$, by the previous remarks) and also if $\varphi_c(x)$ is close to c (then x is close to zero and so $|(\varphi_c^2)'(x)| \approx |x|^{2\lambda - 1} \approx \infty$).

It is straightforward to check that L1-L5 are satisfied by a nonempty open set of one-parameter families, where openness is meant with respect to the C^2 topology in the space of real maps ψ. Moreover, we shall show that these hypotheses do imply that φ_a is essentially uniformly expanding for all parameters up to c:

Proposition 1.1. — *Given any $a \in [a_1, c]$,*

(1) *the interval $[-a, a]$ is forward invariant and $\varphi|[-a, a]$ is transitive*
(2) $|(\varphi_a^n)'(x)| \geq \min\{\sqrt{2}, |\varphi_a'(x)|\}(\sqrt{2})^{n-1}$ *for all $x \in [-a, a]$ such that $\varphi_a^j(x) \neq 0$ for every $j = 0, 1, \ldots, n - 1$.*

After the bifurcation $a = c$ such uniform expansivity is clearly impossible, due to the presence of the critical point in the domain of the map. However, our main result states that – nonuniform – expansivity persists, in a measure-theoretical sense, and is even the prevalent form of dynamics after the bifurcation. We denote $c_j(a) = \varphi_a^j(c)$ for each $j \geq 1$.

Theorem. — *Let $\{\varphi_a\}$ be a Lorenz-like family satisfying conditions L1-L5. Then there are $\sigma > 0$ and $\mathcal{A}^+ \subset \mathbb{R}$ such that $|(\varphi_a^j)'(c_1(a))| \geq e^{\sigma j}$ for all $a \in \mathcal{A}^+$ and $j \geq 1$ and*

$$\lim_{\varepsilon \to 0} \frac{m(\mathcal{A}^+ \cap [c, c+\varepsilon])}{\varepsilon} = 1 \qquad (m = \textit{Lebesgue measure on } \mathbb{R}).$$

Moreover, there is $\sigma_1 > 0$ such that if $a \in \mathcal{A}^+$ then for m-almost all $x \in [-a, a]$ we have $\limsup \frac{1}{n}|\log(\varphi_a^n)'(x)| \geq \sigma_1$ *as $n \to \infty$.*

Measure theoretic persistence of positive Lyapunov exponents (outside the class of uniformly expanding maps) was first proved by Jakobson **[9]**, for maps in the quadratic family $f_a(x) = 1 - ax^2$ close to parameter values $\bar{a}$ satisfying (**[12]**)

$$\inf_{j \in \mathbb{N}} \left| f_{\bar{a}}^j(c) - c \right| > 0 \qquad (c = \text{critical point} = 0). \tag{2}$$

There exist today many proofs of this theorem, *e.g.* **[4]**, **[2]**, as well as generalizations to families of smooth maps with finitely many critical points **[16]**, and to families of maps in which a single discontinuity coincides with the critical point **[14]**. A number of differences should be pointed out in this setting, between smooth maps and our Lorenz-like maps.

While all proofs of Jakobson's theorem in the smooth context rely in one way or the other on the nonrecurrence condition (2), here we need no assumption on the orbits of the critical points for $a = c$. Instead, we simply take advantage of the strong expansivity estimates given by Proposition 1.1 for that parameter value.

Various technical complications arise in the present situation from the existence of discontinuities and of regions where the derivative has arbitrarily large norm. Several estimates (including distortion bounds), which in the smooth case rely on the boundedness and Lipschitz continuity of the derivative, now require nontrivial reformulations together with a detailed study of the recurrence near the discontinuity (and not only near the critical points).

Lorenz-like families with criticalities undergo codimension-one bifurcations which mark a direct transition from uniformly expanding dynamics (for $a < c$), to nonuniformly expanding dynamics (for $a \in \mathcal{A}^+$), a kind of bifurcation which does not seem to be known in the smooth one-dimensional context. The fact that the bifurcation parameter $a = c$ is a Lebesgue density point for $\mathcal{A}^+$ is related to the strong form of expansivity exhibited by φ_c. An interesting question is whether some characterization of the density points of $\mathcal{A}^+$ can be given in terms of special hyperbolicity features of the corresponding maps (*e.g.* uniformly hyperbolic structure on periodic orbits ?).

We remark here that the symmetry inherent in our definition of Lorenz-like maps, though partly justified by the symmetry which exists in Lorenz' system of equations, is not strictly necessary for the proof of the theorem. We carry out the proof in the symmetric case in order to simplify the exposition (in particular we shall often discuss some construction or result with explicit reference to only one of the critical points with the implicit understanding that the same statements apply to the other one as well, by symmetry) but all the arguments hold, up to minor modifications, in a nonsymmetric setting.

Closing this section, we observe that the dynamics of the Lorenz flows in the parameter range we want to consider cannot be expected to fully reduce to that of one-dimensional maps (as happens for the geometric models of the Lorenz attractor mentioned previously). Indeed, the very phenomenon of "folding" which we want to encompass in our description, is also an obstruction to the existence of invariant foliations transverse to the flow. Nevertheless, drawing on the results obtained in this article we are developing, in a forthcoming paper by the same authors **[10]**, a natural extension of those geometric models to this wider range of parameters. Such Lorenz-like flows are amongst the simplest systems in which behaviour arising from the presence of equilibria interacts with dynamical features related to the presence of criticalities (homoclinic and heteroclinic tangencies). The understanding of the bifurcations taking place in this model is probably a necessary step towards a global description of the dynamics of flows, in the spirit of the program proposed a few years ago by Palis, see **[13]**.

The proof of our main result is organized as follows. In Section 2 we identify a pair of conditions on the parameter a which ensure that $a \in \mathcal{A}^+$. Sections 3 and 4 are then devoted to showing that the set of parameters for which such conditions are satisfied is large in the sense of the statement of the theorem. The whole global approach is inspired on **[3]**.

Acknowledgements. — This work was started during the Conference on Real and Complex Dynamics held in Hillerød, Denmark, in the Summer of 1993. It was concluded two years later in the pleasant surroundings of Orsay, during the Colloquium Adrien Douady. We are thankful to the organizers of both conferences for a fine ambiance. We are also grateful to Colin Sparrow for stimulating conversations, to Warwick Tucker for a thorough reading of a preliminary version and for pointing out several misprints and to Jacob Palis for his friendship and encouragement.

2. Positive Lyapunov Exponents

We begin by proving Proposition 1.1. In doing this we focus only on $a \in [a_2, c]$: the case $a < a_2$ corresponds to the situation in **[6]**, and it also follows from (simpler versions of) these same arguments.

2.1. Proof of Proposition 1.1. — The invariance of $[-a,a]$ is an immediate consequence of $\lim_{x\to 0}|\varphi_a(x)| = a$, $|\varphi_a(\pm a)| < x_{\sqrt{2}} \le a$ (recall L4), and the monotonicity of φ_a on $(-a,0)$ and $(0,a)$.

Next, let x and $1 \le j \le n-1$ be as in part (2). If $|\varphi_a^j(x)| \le x_{\sqrt{2}}$ then we have $|\varphi_a'(\varphi_a^j(x))| \ge \sqrt{2}$. If $|\varphi_a^j(x)| > x_{\sqrt{2}}$ then, by L4, there exists a unique $z \in [-c,c]$ such that $\varphi_c(z) = \varphi_a^j(x) = \varphi_a(\varphi_a^{j-1}(x))$. Moreover, z and $\varphi_a^{j-1}(x)$ have the same sign (opposite to that of $\varphi_a^j(x)$) and $|z| \ge |\varphi_a^{j-1}(x)|$, because $a \le c$. Using L5 we get

$$\begin{aligned}|(\varphi_a^2)'(\varphi_a^{j-1}(x))| &= |\varphi'(\varphi_a^{j-1}(x))\,\varphi'(\varphi_a^j(x))|\\ &\ge |\varphi'(z)\,\varphi'(\varphi_a^j(x))| = |(\varphi_c^2)'(z)| > 2.\end{aligned}$$

Part (2) follows directly from these remarks and it is easy to see that one even gets a somewhat better bound, with $\sqrt{2}$ replaced by some slightly larger constant θ.

To prove transitivity, we let $U_0, V_0 \subset [-a,a]$ be arbitrary open sets and show that $\varphi_a^n(U_0) \cap V_0 \ne \varnothing$ for some $n > 0$. Suppose, without loss of generality, that $0 \notin U_0$ and $U_0 \subset (-x_{\sqrt{2}}, x_{\sqrt{2}})$ (recall L4). As long as $0 \notin \varphi_a^j(U_0)$, write $U_j = \varphi_a^j(U_0)$ and notice that $|U_j| \ge \theta^j|U_0|$. Thus we must have $0 \in \varphi_a(U_{k_1-1})$ for some $k_1 \ge 1$. Let U_{k_1} denote the largest connected component of $\varphi_a(U_{k_1-1}) \setminus \{0\}$ and observe that $|U_{k_1}| \ge \frac{1}{2}|\varphi_a(U_{k_1-1})| \ge \frac{1}{2}\theta^{k_1}|U_0|$. Suppose first that $U_{k_1} \subset (z^-, z^+)$, where $z^- < 0 < z^+$ are the preimages of zero under φ_a; observe that $|z^\pm| < x_{\sqrt{2}}$ as a consequence of the first inequality in L4. Then we proceed as before, with U_0 replaced by U_{k_1}. More precisely, we define $U_{k_1+j} = \varphi_a^j(U_{k_1})$ until the first iterate $k_2 > k_1$ for which $0 \in \varphi_a(U_{k_2-1})$; at that point we take U_{k_2} to be the largest component of $\varphi_a(U_{k_2-1})$ and repeat the whole procedure again. As long as $U_{k_i} \subset (z^-, z^+)$ we have $k_{i+1} \ge k_i + 2$, hence

$$|U_{k_{i+1}}| \ge |\varphi_a(U_{k_i-1})|/2 \ge \theta^{k_{i+1}-k_i}|U_{k_i}|/2 \ge \theta^2|U_{k_i}|/2$$

grows exponentially with i. Thus, one eventually reaches some $k = k_j$ for which U_k contains either $(z^-,0)$ or $(0,z^+)$. In the first case $\varphi_a(U_k)$ contains $(0,a) \supset (0,z^+)$ and then $\varphi_a^2(U_k)$ contains $(-a,0)$, which ensures that either $\varphi_a(U_k)$ or $\varphi_a^2(U_k)$ intersect V_0. The second case is entirely analogous so the proof of the proposition is complete.

Now we fix a number of constants to be used in the sequel of our argument. Recall that $0 < \lambda < 1/2$. We take $\sigma_0 > 0$ and $\sigma > 0$ such that $0 < 2\sigma < \sigma_0 < \log\sqrt{2}$ and also choose

$$\gamma > 1 \text{ and } \delta, \iota > 0 \text{ such that } 1 < \gamma + \delta + \iota < 1/2\lambda.$$

We will be choosing δ small with respect to λ and ι small with respect to δ. We remark for future reference that this implies $\gamma + \delta + \iota < 1/\lambda - 1$. Then, we let $0 < \alpha < \beta$ be small, depending on the previous constants (the precise conditions are stated throughout the proof wherever they are required).

By conditions L1-L3 there exist $\eta_1, \eta_2 > 0$ such that

$$\lim_{x\to 0} \frac{|\varphi(x)|}{|x|^\lambda} = \eta_1 \qquad \text{and} \qquad \lim_{x\to c} \frac{|\varphi(x) - \varphi(c)|}{|x-c|^2} = \eta_2.$$

For each $i = 1, 2$, we fix constants $\eta_i^- = \eta_i - v$ and $\eta_i^+ = \eta_i + v$, where v is some small positive number (once more, precise conditions are to be stated along the way). Then we have

M1 : for all $x \neq 0$ close enough to the origin,

$$\begin{aligned} \eta_1^-|x|^\lambda \le \varphi_a(x) + a \le \eta_1^+|x|^\lambda \qquad & \text{if } x > 0, \\ -\eta_1^+|x|^\lambda \le \varphi_a(x) - a \le -\eta_1^-|x|^\lambda \qquad & \text{if } x < 0, \\ \text{and} \qquad \eta_1^-\lambda|x|^{\lambda-1} \le |\varphi_a'(x)| \le \eta_1^+\lambda|x|^{\lambda-1}; & \end{aligned}$$

M2 : for all x close enough to the critical point c

$$\begin{aligned} & \eta_2^-(x-c)^2 \le |\varphi_a(x) - \varphi_a(c)| \le \eta_2^+(x-c)^2 \\ \text{and} \qquad & 2\eta_2^-|x-c| \le |\varphi_a'(x)| \le 2\eta_2^+|x-c| \end{aligned}$$

and a similar fact holds for all x close enough to $-c$.

Now, for each small $\varepsilon > 0$ we let Δ_+^0, Δ_+^c, Δ_+^{-c} denote the ε^γ- neighbourhoods of the origin and of the critical points c and $-c$, respectively. We define partitions of Δ_+^0 and $\Delta_+^{\pm c}$ by writing $I_r = [\varepsilon^\gamma e^{-r}, \varepsilon^\gamma e^{-r+1})$ and

$$\Delta_+^0 = \{0\} \cup \bigcup_{|r|\ge 1} I_r^0 \qquad \text{and} \qquad \Delta_+^{\pm c} = \{\pm c\} \cup \bigcup_{|r|\ge 1} I_r^{\pm c},$$

where $I_r^0 = I_r$ and $I_{-r}^0 = -I_r$, for each $r \ge 1$, and the $I_r^{\pm c} = I_r^0 \pm c$ are simply the translates of the I_r^0. We shall always assume that $\varepsilon > 0$ is small enough so that Δ_+^0 and $\Delta_+^{\pm c}$ are contained in the regions for which M1 and M2 are valid. Moreover, we let $r_\varepsilon = [\delta \log \varepsilon^{-1}]$ (here $[x]$ is the integer part of x) and we consider restricted neighbourhoods

$$\Delta^0 = \{0\} \cup \bigcup_{|r|\ge r_\varepsilon+1} I_r^0 \qquad \text{and} \qquad \Delta^{\pm c} = \{\pm c\} \cup \bigcup_{|r|\ge r_\varepsilon+1} I_r^{\pm c}$$

(of radius $\approx \varepsilon^{\gamma+\delta}$) of the origin and the critical points. We shall also need an even smaller neighbourhood (of radius $\approx \varepsilon^{2(\gamma+\delta)+\iota}$) of the origin. So let

$$\Delta_{r_s}^0 = \{0\} \cup \bigcup_{|r|\ge r_s+1} I_r^0, \qquad r_s = [(\gamma + 2\delta + \iota)\log 1/\varepsilon].$$

We shall prove below that the preimages of the critical neighbourhoods $\Delta_+^{\pm c}$ are always contained in this smallest neighbourhood of the origin, *i.e.* $\varphi_a^{-1}(\Delta_+^{\pm c}) \subset \Delta_{r_s}^0$.

2.2. Breaking the hyperbolic structure. — The loss of expansivity occurring after the bifurcation $a = c$ and caused by the critical points entering the domain of the map is, in some sense, local: for $c \le a \le c + \varepsilon$, it occurs only in a neighbourhood of the critical points of size $\varepsilon^{\gamma+\delta} \ll \varepsilon$. More precisely, any piece of orbit that does not intersect $\Delta^{\pm c}$ has an exponentially growing derivative. Proving this fact requires two preliminary lemmas. First we determine the position and size of the preimage of $\Delta^{\pm c}$ for a convenient range of parameter values. Then we estimate the accumulated derivative of points which pass close to the discontinuity or to the critical points. In all that follows we write $\rho = 2^{-\lambda}$.

Lemma 2.1. — *If $\varepsilon > 0$ is sufficiently small then $(\varepsilon/\eta_1^+)^{1/\lambda} \le |\varphi_{c+\varepsilon}^{-1}(\pm c)| \le (\varepsilon/\eta_1^-)^{1/\lambda}$ and*

$$\frac{1}{e} \le \frac{|\varphi_a^{-1}(y)|}{|\varphi_{c+\varepsilon}^{-1}(\pm c)|} \le e$$

for every $y \in \Delta_+^{\pm c}$ and $a \in [c + \rho\varepsilon, c + \varepsilon]$.

Proof. — By symmetry it suffices to consider the preimages of points in Δ_+^c. Using the second inequality in M1,

$$-\eta_1^+ |\varphi_{c+\varepsilon}^{-1}(c)|^\lambda \le c - (c+\varepsilon) \le -\eta_1^- |\varphi_{c+\varepsilon}^{-1}(c)|^\lambda,$$

which immediately gives the first claim. To prove the second one notice that, for any y and a as in the statement, $y - a = (c-a) - (c-y) \in [-\varepsilon - \varepsilon^\gamma, -\rho\varepsilon + \varepsilon^\gamma]$ and so, using M1 in the same way as before,

$$\left(\rho\varepsilon - \frac{\varepsilon^\gamma}{\eta_1^+}\right)^{1/\lambda} \le |\varphi_a^{-1}(y)| \le \left(\frac{\varepsilon + \varepsilon^\gamma}{\eta_1^-}\right)^{1/\lambda}.$$

Combining with the first part of the present lemma, we get

$$\left(\frac{\eta_1^-}{\eta_1^+}\frac{\rho\varepsilon - \varepsilon^\gamma}{\varepsilon}\right)^{1/\lambda} \le |\frac{\varphi_a^{-1}(y)}{|\varphi_{c+\varepsilon}^{-1}(c)|}| \le \left(\frac{\eta_1^+}{\eta_1^-}\frac{\varepsilon + \varepsilon^\gamma}{\varepsilon}\right)^{1/\lambda}.$$

The left hand side is close to $1/2$, and hence larger than $1/e$, if ε is small (and υ has been fixed sufficiently small, recall the definition of $\eta_i^\pm$). Analogously, the right hand side is smaller than e if ε and υ are small enough. The proof is complete. □

Now we define $r_c = r_c(\varepsilon) \ge 1$ by the condition $\varphi_{c+\varepsilon}^{-1}(c) \in I_{-r_c}^0$. Observe that

$$fracle \left(\frac{\varepsilon}{\eta_1^+}\right)^{1/\lambda} \le \varepsilon^\gamma e^{-r_c} \le \left(\frac{\varepsilon}{\eta_1^-}\right)^{1/\lambda} \tag{3}$$

by the first part of the previous lemma. Moreover, the second part gives

$$\varphi_a^{-1}(\Delta_+^c) \subset I_{-r_c+1}^0 \cup I_{-r_c}^0 \cup I_{-r_c-1}^0, \quad \text{for every } a \in [c+\rho\varepsilon, c+\varepsilon]. \tag{4}$$

Notice that we have from (3) $e^{-r_c} \le \varepsilon^{1/\lambda-\gamma}(\eta_1^-)^{1/\lambda}$ which yields

$$r_c \ge (1/\lambda - \gamma)\log 1/\varepsilon + 1/\lambda \log \eta_1^- > (\gamma + 2\delta + \iota)\log 1/\varepsilon$$

if ε is small enough. This shows, as promised above, that $\varphi_a^{-1}(\Delta_+^{\pm c}) \subset \Delta_{r_s}^0$.

Lemma 2.2. — *For every $a \in [c+\rho\varepsilon, c+\varepsilon]$ and $x \in I_r^0$, $|r| \geq 1$,*

(1) *if $\varphi_a(x) \notin \Delta^{\pm c}$ then $|(\varphi_a^2)'(x)| \geq \mathrm{const}(\varepsilon^\gamma e^{-r})^{2\lambda-1} \geq e^{2\sigma_0}e^{\beta|r|}$;*

(2) *if $\varphi_a(x) \in \Delta_+^{\pm c}$, with $\varphi_a(x) \in I_{\tilde{r}}^{\pm c}$,*

then

$$|\varphi_a'(x)| \geq \mathrm{const}(\varepsilon^\gamma e^{-r})^{\lambda-1} > e^{\sigma_0}e^{\beta|r|}$$

and

$$|(\varphi_a^2)'(x)| \geq \varepsilon^{\gamma+1-1/\lambda}e^{-|\tilde{r}|} > e^{2\sigma_0}e^{-|\tilde{r}|}.$$

Proof. — We consider $r, \tilde{r} \geq 1$, the other cases being entirely analogous. For the sake of clearness let us split the proof of (1) into three different cases.

Suppose first that $r < r_c - 2$. Then, in view of (4), $|x - \varphi_a^{-1}(\pm c)| \geq |I_{r+1}| = (1-1/e)\varepsilon^\gamma e^{-r}$. Thus, by the mean value theorem,

$$\begin{aligned} |\varphi_a(x) \pm c| &\geq |x - \varphi_a^{-1}(\pm c)| \cdot \inf\{\varphi'(z) : z \in [x, \varphi_a^{-1}(c)]\} \\ &\geq (1-1/e)\varepsilon^\gamma e^{-r}\eta_1^-\lambda(\varepsilon^\gamma e^{-r+1})^{\lambda-1} \\ &\geq k_1\varepsilon^{\gamma\lambda}e^{-r\lambda}, \end{aligned}$$

with $k_1 = (1-1/e)\eta_1^-\lambda e^{\lambda-1}$. It follows, using M1, M2,

$$\begin{aligned} |(\varphi_a^2)'(x)| &\geq |\varphi_a'(x)||\varphi_a'(\varphi_a(x))| &&\geq \eta_1^-\lambda(\varepsilon^\gamma e^{-r+1})^{\lambda-1}2\eta_2^-k_1\varepsilon^{\gamma\lambda}e^{-r\lambda} \\ &\geq k_2\varepsilon^{-\gamma(1-2\lambda)}e^{r(1-2\lambda)} &&\geq e^{2\sigma_0}e^{\beta r}, \end{aligned}$$

where $k_2 = 2\eta_1^-\lambda e^{\lambda-1}\eta_2^-k_1$ and, for the last inequality, we suppose $\beta < 1-2\lambda$ and ε sufficiently small.

Clearly, exactly the same argument works for $|r| \geq r_c - 2$ where we have even greater expansion.

Finally, suppose that $\varphi_a(x) \in I_{\tilde{r}}^{\pm c} \subset \Delta_+^{\pm c}$. Clearly,

$$|\varphi_a'(x)| \geq \eta_1^-\lambda(\varepsilon^\gamma e^{-r+1})^{\lambda-1} \geq \eta_1^-\lambda e^{\lambda-1}\varepsilon^{-\gamma(1-\lambda)}e^{r(1-\lambda)} \geq e^{\sigma_0}e^{\beta r}$$

if β and ε are small. Moreover, by (4), $\varepsilon^\gamma e^{-r_c-1} \leq |x| \leq \varepsilon^\gamma e^{-r_c+2}$, which gives

$$\begin{aligned} |(\varphi_a^2)'(x)| &\geq |\varphi_a'(x)||\varphi_a'(\varphi_a(x))| &&\geq \eta_1^-\lambda(\varepsilon^\gamma e^{-r_c+2})^{\lambda-1}2\eta_2^-\varepsilon^\gamma e^{-\tilde{r}} \\ &\geq k_5\varepsilon^{\gamma\lambda}e^{(1-\lambda)r_c}e^{-\tilde{r}} &&\geq k_5\varepsilon^{\gamma\lambda+(1-\lambda)(\gamma-1/\lambda)}(\eta_1^-)^{(1/\lambda)-1}e^{-\tilde{r}} \\ &\geq k_6\varepsilon^{\gamma-(1/\lambda)+1}e^{-\tilde{r}} &&\geq e^{2\sigma_0}e^{-\tilde{r}}, \end{aligned}$$

where $k_5 = 2\eta_1^-\lambda e^{2(\lambda-1)}\eta_2^-$ and $k_6 = k_5(\eta_1^-)^{(1/\lambda)-1}$ and we use the relation (3) in the fourth inequality. □

Lemma 2.3. — *For any $a \in [c+\rho\varepsilon, c+\varepsilon]$ and $x \in [-a, a]$,*

(1) *if $\{\varphi_a^j(x)\}_{j=0}^{n-1} \cap \Delta^{\pm c} = \varnothing$ then $|(\varphi_a^n)'(x)| \geq \min\{e^{\sigma_0}, |\varphi_a'(x)|\}e^{\sigma_0(n-1)}$;*

(2) *if, in addition, $\varphi_a^n(x) \in \Delta_+^{\pm c}$ then $|(\varphi_a^n)'(x)| \geq e^{\sigma_0 n}$.*

Proof. — Denote $x_j = \varphi_a^j(x)$, for $0 \le j \le n-1$. We claim that given any $j \ge 1$ either $|\varphi_a'(x_j)| \ge e^{\sigma_0}$ or $|(\varphi_a^2)'(x_{j-1})| \ge e^{2\sigma_0}$. This is obvious if $|x_j| \le x_{\sqrt{2}}$, because we get $|\varphi_a'(x_j)| = |\varphi'(x_j)| > e^{\sigma_0}$. From now on we consider $x_j \ge x_{\sqrt{2}}$, the case $x_j \le -x_{\sqrt{2}}$ being entirely analogous. If $x_{j-1} \in \Delta_+^0$ then $|(\varphi_a^2)'(x_{j-1})| \ge e^{2\sigma_0}$ by part (1) of the previous lemma. Therefore, we may suppose $x_{j-1} \notin \Delta_+^0$, that is $|x_{j-1}| \ge \varepsilon^\gamma$. Then, recall M1, M2, $c - \varphi_c(x_{j-1}) \ge \eta_1^- \varepsilon^{\gamma\lambda}$ and so $|\varphi'(\varphi_c(x_{j-1}))| \ge 2\eta_1^-\eta_2^-\varepsilon^{\gamma\lambda}$. Hence, using also $\varphi_a(x_{j-1}) - \varphi_c(x_{j-1}) = a - c \le \varepsilon$, we get

$$\begin{aligned}\frac{|(\varphi_a^2)'(x_{j-1})|}{|(\varphi_c^2)'(x_{j-1})|} =& \frac{|\varphi'(\varphi_a(x_{j-1}))|}{|\varphi'(\varphi_c(x_{j-1}))|} \\ \ge& 1 - \frac{|\varphi'(\varphi_a(x_{j-1})) - \varphi'(\varphi_c(x_{j-1}))|}{|\varphi'(\varphi_c(x_{j-1}))|} \ge 1 - k_7\varepsilon^{1-\lambda\gamma},\end{aligned}$$

where $k_7 = k/(2\eta_1^-\eta_2^-)$, with k a Lipschitz constant for φ' on $\{x \ge x_{\sqrt{2}} - \varepsilon_0\}$ (ε_0 is some small constant, we take $\varepsilon \le \varepsilon_0$). Since $1 - \lambda\gamma > 0$, the left hand term is larger that $e^{2\sigma_0}/2$ if ε is small enough and then the claim follows from L5. Moreover, the first statement in the lemma is a direct consequence of our claim (*cf.* the proof of lemma 2.1).

In order to deduce the second part of the lemma we may suppose $|\varphi_a'(x)| \le e^{\sigma_0}$, for otherwise there is nothing to prove. Observe also that if $\varphi_a^n(x) \in \Delta_+^{\pm c}$ then, by (3), (4), we have $|\varphi_a'(\varphi_a^{n-1}(x))| \ge k_8\varepsilon^{1-1/\lambda}$, with $k_8 = \lambda e^{2(\lambda-1)}(\eta_1^-)^1/\lambda$.

Moreover, by hypothesis, $x \notin \Delta^{\pm c}$ and so $|\varphi_a'(x)| \ge \eta_1^-\lambda\varepsilon^{\gamma+\delta}$. Altogether, writing $k_9 = \eta_1^-\lambda k_8$,

$$\begin{aligned}|(\varphi_a^n)'(x)| &\ge |\varphi_a'(x)|e^{\sigma_0(n-2)}|\varphi'(\varphi_a^{n-1}(x))| &&\ge k_9\varepsilon^{\gamma+\delta}e^{\sigma_0(n-2)}\varepsilon^{1-1/\lambda} \\ &\ge k_9\varepsilon^{\gamma+\delta+1-1/\lambda}e^{\sigma_0(n-2)} &&\ge e^{\sigma_0 n},\end{aligned}$$

if ε is small enough. □

2.3. Recovering expansion. — Now we deal with the expansion losses occurring when trajectories pass close to some of the critical points $\pm c$. More precisely, we consider points $x \in \Delta^{\pm c}$. Assuming that the critical trajectories satisfy (exponential) expansivity and bounded recurrence conditions (during a convenient number of iterates, depending on $|x \pm c|$), we show that the small value of $\varphi_a'(x)$ is fully compensated in the subsequent iterates, during which the trajectory of x remains close to that of the critical point (and so exhibits rapidly increasing derivative).

For each $j \ge 0$ let $c_j = c_j(a) = \varphi_a^j(\pm c)$ and denote $d(c_j) = \min\{|c_j|, |c_j \pm c|\}$. In what follows $\varepsilon > 0$ is fixed and we suppose $a \in [c + \rho\varepsilon, c + \varepsilon]$.

Lemma 2.4. — *There exists $\theta = \theta(\beta - \alpha) > 0$ such that the following estimates hold. Let $x \in I_r^{\pm c}$ for some $|r| \ge r_\varepsilon$. Suppose that there is $n \ge |r|/\alpha$ such that*

(5) $\qquad d(c_j) \ge \varepsilon^\gamma e^{-\alpha j}$ *and* $|(\varphi_a^j)'(c_1)| \ge e^{\sigma j}$, *for all* $1 \le j \le n-1$.

Then there exists an integer $p = p(x) \ge 1$ such that

(1) *For all* $y_1, z_1 \in [\varphi_a(x), \varphi_a(\pm c)]$ *and for all* $1 \le k \le p$,

$$\frac{1}{\theta} \le \frac{|(\varphi_a^k)'(z_1)|}{|(\varphi_a^k)'(y_1)|} \le \theta$$

(2) $p \le (2|r| + \frac{3}{2}\gamma \log 1/\varepsilon)/\sigma \le n-1$;

(3) (a) $|(\varphi_a^{p+1})'(x)| \ge \varepsilon^{2\beta\gamma/\sigma} e^{(1-2\beta/\sigma)r}$;

(b) $|(\varphi_a^{p+1})'(x)| \ge \varepsilon^{-\beta/\sigma} e^{|r|-(\alpha n/2)}$;

(c) $|(\varphi_a^{p+1})'(x)| \ge e^{\sigma_0 + \beta p} \ge e^{\beta(p+1)}$.

Proof. — We suppose $x \in I_r^c$ with $r \ge 1$, the remaining cases being treated in precisely the same way. Define $p = p(x) \ge 1$ as the maximum integer such that

$$|x_i - c_i| \le \varepsilon^{\gamma} e^{-\beta i} \quad \text{for all } 1 \le i \le p \tag{6}$$

where $x_i = \varphi_a^i(x)$. Recall that we fix $\beta > \alpha$. Therefore, (6) and the first condition in (5) ensure that the intervals $[x_i, c_i]$, $1 \le i \le p$, do not contain the origin nor any of the critical points $\pm c$. Therefore, $\varphi_a^i : [x_1, c_1] \to [x_{i+1}, c_{i+1}]$ is a diffeomorphism for all $1 \le i \le p$. In particular, given any $y_1, z_1 \in [x_1, c_1]$ we have $y_i, z_i \in [x_i, c_i]$ for $1 \le i \le p$, where $y_i = \varphi_a^i(y)$ and $z_i = \varphi_a^i(z)$. By the chain rule,

$$\left|\frac{(\varphi_a^k)'(z_1)}{(\varphi_a^k)'(y_1)}\right| = \prod_{i=1}^{k} \left|\frac{\varphi_a'(z_i)}{\varphi_a'(y_i)}\right| = \prod_{i=1}^{k} \left|1 + \frac{\varphi_a'(z_i) - \varphi_a'(y_i)}{\varphi_a'(y_i)}\right|$$

and so part (1) will follow if we show that

$$\sum_{i=1}^{k} \left|\frac{\varphi_a'(z_i) - \varphi_a'(y_i)}{\varphi_a'(y_i)}\right| \tag{7}$$

is bounded by some constant depending only on $\beta - \alpha$. By the mean value theorem there exists, for each $1 \le i \le k$ some $\xi_i \in [z_i, y_i]$ s.t.

$$\left|\frac{\varphi_a'(z_i) - \varphi_a'(y_i)}{\varphi_a'(y_i)}\right| = \left|\frac{|z_i - y_i| \varphi_a''(\xi_i)}{\varphi_a'(y_i)}\right| \le \varepsilon^{\gamma} e^{-\beta i} \left|\frac{\varphi_a''(\xi_i)}{\varphi_a'(y_i)}\right|.$$

Thus it is sufficient to show that $|\varphi_a''(\xi_i)/\varphi_a'(y_i)| \le \mathrm{const}\, \varepsilon^{-\gamma} e^{\alpha i}$ to conclude that the terms of the sum (7) are decreasing exponentially and so the entire sum is bounded by a constant independent of k. We fix some small constant $\varepsilon' > 0$ independent of ε. The norm of $\varphi_a''(x)$ is bounded above and below outside $(-\varepsilon', \varepsilon')$ by some constant $C = \sup\{|\varphi_a''(x)| : x \notin (-\varepsilon', \varepsilon')\}$. For simplicity, and without loss of generality, we shall assume that this supremum is actually achieved at ε'. Inside $(-\varepsilon', \varepsilon')$ we have by the form of the map φ that $|\varphi_a''(x)| \le \eta^+ \lambda(\lambda - 1)|x|^{\lambda-2}$.

We distinguish two cases. If $[x_i, c_i] \cap (-\varepsilon', \varepsilon') = \varnothing$ then we have

$$|\varphi_a'(y_i)| \ge 2\eta_2^- |y_i - c| \ge 2\eta_2^- \varepsilon^{\gamma}(\varepsilon^{-\alpha i} - e^{-\beta i}) \ge 2\eta_2^-(1 - e^{\alpha - \beta})\varepsilon^{\gamma} e^{-\alpha i}$$

and so

$$\left|\frac{\varphi_a''(\xi_i)}{\varphi_a'(y_i)}\right| \le \frac{C}{2\eta_2^-(1 - e^{\alpha - \beta})} \varepsilon^{-\gamma} e^{\alpha i}$$

as desired. If $[x_i, c_i] \cap (-\varepsilon', \varepsilon') \neq \varnothing$ we have the following estimates. To simplify the notation we shall suppose that $[x_i, c_i] \subset (0, c)$. The other case $[x_i, c_i] \subset (-c, 0)$ is dealt with similarly. Taking ϵ' small and since $|x_i - c_i| \leq \varepsilon^\gamma$ we can suppose that $[x_i, c_i]$ is contained in the neighbourhood of 0 for which conditions M1 and M2 hold. We have

$$|\varphi_a'(y_i)| \geq |\varphi_a'(c_i + \varepsilon^\gamma e^{-\beta i})| \geq \eta_1^- \lambda (c_i + \varepsilon^\gamma e^{-\beta i})^{\lambda - 1}$$

and

$$|\varphi_a''(\xi_i)| \leq |\varphi_a''(c_i - \varepsilon^\gamma e^{-\beta i})| \leq \eta_1^+ \lambda(\lambda - 1)(c_i - \varepsilon^\gamma e^{-\beta i})^{\lambda - 2}.$$

This gives

$$\left|\frac{\varphi_a''(\xi_i)}{\varphi_a'(y_i)}\right| \leq \frac{\eta_1^+(\lambda-1)}{\eta_1^-}\left(\frac{c_i - \varepsilon^\gamma e^{-\beta i}}{c_i + \varepsilon^\gamma e^{-\beta i}}\right)^{\lambda-1} (c_i - \varepsilon^\gamma e^{-\beta i})^{-1} \leq \operatorname{const}(\varepsilon^\gamma e^{-\alpha i})^{-1}.$$

This follows from the fact that $|c_i| \geq \varepsilon^\gamma e^{-\alpha i}$ and therefore

$$(c_i - \varepsilon^\gamma e^{-\beta i})^{-1} \leq (\varepsilon^\gamma e^{-\alpha i})^{-1}(1 - e^{(\alpha-\beta)i})^{-1} \leq \operatorname{const}(\varepsilon^\gamma e^{-\alpha i})^{-1}$$

and that $(c_i - \varepsilon^\gamma e^{-\beta i})/(c_i + \varepsilon^\gamma e^{-\beta i}) \leq$ const. Indeed this last fact follows from observing that

$$c_i - \varepsilon^\gamma e^{-\beta i} \geq \varepsilon^\gamma e^{-\alpha i} - \varepsilon^\gamma e^{-\beta i} \geq \varepsilon^\gamma e^{-\alpha i}(1 - e^{(\alpha-\beta)i}) \geq (1 - e^{\alpha-\beta})\varepsilon^\gamma e^{-\alpha i}$$

and similarly

$$c_i + \varepsilon^\gamma e^{-\beta i} \leq (1 + e^{\alpha-\beta})\varepsilon^\gamma e^{-\alpha i}$$

which together give

$$\frac{c_i - \varepsilon^\gamma e^{-\beta i}}{c_i + \varepsilon^\gamma e^{-\beta i}} \leq \frac{1 - e^{\alpha-\beta}}{1 + e^{\alpha-\beta}} = \operatorname{const}.$$

This proves (1).

Starting the proof of (2), let $q = \min\{p, n-1\}$. As $x \in I_r^c$, we have $|x - c| \geq \varepsilon^\gamma e^{-r}$ and so $|x_1 - c_1| \geq \eta_2^- \varepsilon^{2\gamma} e^{-2r}$. Then, in view of the second condition in (5) and the distortion estimate we have just proved, the mean value theorem yields $\eta_2^- \varepsilon^{2\gamma} e^{-2r} \theta^{-1} e^{\sigma(q-1)} \leq |x_q - c_q| \leq \varepsilon^\gamma e^{-\beta q}$. Thus

$$q \leq \frac{2r + \gamma \log(1/\varepsilon) + \sigma - \log(\eta_2^-/\theta)}{\sigma + \beta} \leq \left(2r + \frac{3}{2}\gamma \log 1/\varepsilon\right)/\sigma$$

as long as ε is sufficiently small. Since we also take $\alpha n \geq r \geq [\delta \log 1/\varepsilon] \gg 1$, we find that $q \leq (2\alpha n + 3\gamma \alpha n / 2\delta)/\sigma < n$ (if α is small), so that it must be $q = p$. In this way we have proved that $p \leq (2r + \frac{3}{2}\gamma \log 1/\varepsilon)/\sigma < n$, as claimed in part (2) of the lemma.

Now, by the definition of p we have $|x_{p+1} - c_{p+1}| \geq \varepsilon^\gamma e^{-\beta(p+1)}$. Thus, using part (1) in conjunction with the mean value theorem,

$$|(\varphi_a^p)'(x_1)| \geq \frac{1}{\theta}\frac{|x_{p+1} - c_{p+1}|}{|x_1 - c_1|} \geq \frac{\varepsilon^\gamma e^{-\beta(p+1)}}{\theta \eta_2^- \varepsilon^{2\gamma} e^{-2r+2}} \geq \operatorname{const} \varepsilon^{-\gamma} e^{2r - \beta p}.$$

Since $|\varphi_a'(x)| \geq 2\eta_2^- \varepsilon^\gamma e^{-r}$, we find

$$|(\varphi_a^{p+1})'(x)| \geq \operatorname{const} e^{r-\beta p}. \tag{8}$$

Using part (2) we immediately get

$$e^{r-\beta p} \geq e^{r-\frac{\beta}{\sigma}(2r+\frac{3}{2}\gamma\log 1/\varepsilon)} \geq e^{(1-\frac{2\beta}{\sigma})r-\frac{3\beta\gamma}{2\sigma}\log 1/\varepsilon} \geq \varepsilon^{\frac{3\beta\gamma}{2\sigma}} e^{(1-\frac{2\beta}{\sigma})r}.$$

This proves the first statement in part (3) since

$$|(\varphi_a^{p+1})'(x) \geq \operatorname{const} \varepsilon^{3\beta\gamma/2\sigma} e^{(1-2\beta/\sigma)r} \geq \varepsilon^{2\beta\gamma/\sigma} e^{(1-2\beta/\sigma)r}.$$

Using part (2) again together with $\alpha n \geq r \geq [\delta \log 1/\varepsilon] \gg 1$, we get

$$\beta p \leq \frac{2\beta}{\sigma}\Big(\alpha n + \frac{\gamma}{\delta}2\alpha n + \frac{1}{\delta}2\alpha n - \log 1/\varepsilon\Big) \leq \frac{4\beta}{\sigma\delta\lambda}\alpha n - \frac{2\beta}{\sigma}\log 1/\varepsilon \leq \frac{1}{2}\alpha n - \frac{2\beta}{\sigma}\log 1/\varepsilon$$

as long as we take $\beta < \sigma\delta\lambda/8$ (we also used $\delta + \gamma + 1 < 1/\lambda$). Replacing in (8), and supposing ε sufficiently small, we get the second statement in part (3). Finally notice that $|r| \geq [\delta \log 1/\varepsilon]$ and therefore $p \leq 2/\sigma(r + \gamma \log 1/\varepsilon) \leq (1 + \gamma/\delta)|r|$ and $|r| \geq p/(1 + \gamma/\delta)$. therefore we have

$$|(\varphi_a^{p+1})'(x)| \geq \operatorname{const} e^{r-\beta p} \geq \operatorname{const} e^{(\frac{1}{1+\gamma/\delta}-\beta)p} \geq e^{\beta(p+1)}$$

if β is small. This completes the proof of part (3) and of the lemma. □

2.4. Proving positive Lyapunov exponents. — We can now state the main results of this section, asserting that, under two convenient assumptions on the parameter a to be stated below, the critical trajectories exhibit exponential growth of the derivative and, in fact, the same is true for most trajectories of φ_a.

As before, we write $c_j = c_j(a) = \varphi_a^j(c)$, for $j \geq 1$. For the time being we fix some $n \geq 1$ and assume that

CP1(n) : $d(c_j) = \min\{|c_j|, |c_j \pm c|\} \geq \varepsilon^\gamma e^{-\alpha j}$ for all $1 \leq j \leq n$.

and

EG($n-1$) : $|(\varphi_a^j)'(c_1)| \geq e^{\sigma j}$ for all $1 \leq j \leq n-1$.

Then we define sequences of integers ν_i, p_i, by $\nu_1 = \inf\{\nu \geq 1 : c_\nu \in \Delta^{\pm c}\}$ and

(i) $p_i = p(c_{\nu_i})$, as given by lemma 2.4;

(ii) $\nu_{i+1} = \inf\{\nu > \nu_i + p_i : c_\nu \in \Delta^{\pm c}\}$.

(CP1(n) ensures that $c_{\nu_i} \in I_r^{\pm c}$ for some $|r| \leq \alpha\nu_i \leq \alpha n$). We take $s \geq 0$ maximum such that $\nu_s \leq n$. Then either $\nu_s \leq n < \nu_s + p_s$ or $\nu_s < \nu_s + p_s \leq n$. Now we define $P_n = p_1 + \cdots + p_{s-1}$ in the first case and $P_n = p_1 + \cdots + p_{s-1} + p_s$ in the second one. Then we further assume that

CP2(n) : $P_j \leq j/2$ for all $1 \leq j \leq n$.

All iterates occurring during a binding period $[\nu_i + 1, \nu_i + p_i]$ are called *bound iterates.* All others (including returns ν_i) are called *free* iterates.

Lemma 2.5. — *Suppose that some parameter $a \in [c+\rho\varepsilon, c+\varepsilon]$ satisfies CP1(n), CP2(n), and EG(n − 1). Then it also satisfies*

EG(n) : $|(\varphi_a^j)'(c_1)| \geq e^{\sigma j}$ *for all* $1 \leq j \leq n$.

Proof. — We let ν_i, p_i, be as above and define $q_0 = \nu_1 - 1$ and $q_i = \nu_{i+1} - (\nu_i + p_i + 1)$ for $1 \leq i \leq s-1$. If $n \geq \nu_s + p_s$ we also write $q_s = n - (\nu_s + p_s)$. Then

$$(9)\quad |(\varphi_a^n)'(c_1)| = |(\varphi_a^{q_0})'(c_1)| \prod_{i=1}^{s-1} (|(\varphi_a^{p_i+1})'(c_{\nu_i})||(\varphi_a^{q_i})'(c_{\nu_i+p_i+1})|)\, |(\varphi_a^{n-\nu_s+1})'(c_{\nu_s})|.$$

The first factor on the right can be estimated as follows. Since $\varphi_a(c_{q_0}) = c_{\nu_1} \in \Delta^{\pm c}$, relations (3) and (4) yield $|\varphi_a'(c_{q_0})| \geq \text{const}(\varepsilon^\gamma e^{-r_c})^{\lambda-1} \geq \text{const}\, \varepsilon^{(1/\lambda)-1}$. Hence, using also the first part of lemma 2.3,

$$|(\varphi_a^{q_0})'(c_1)| = |(\varphi_a^{q_0-1})'(c_1)||(\varphi_a)'(c_{q_0})| \geq \text{const}\, e^{\sigma_0 q_0} \varepsilon^{1-1/\lambda}$$

(note that the last inequality in L4 implies $|c_1| < x_{\sqrt{2}}$ and so $|\varphi_a'(c_1)| > \sqrt{2} > e^{\sigma_0}$ for all a close to c). On the other hand, lemma 2.4(3) and lemma 2.3(2) give, for $1 \leq i \leq s-1$,

$$|(\varphi_a^{p_i+1})'(c_{\nu_i})| \geq e^{\sigma_0 + \beta p_i} \quad \text{and} \quad |(\varphi_a^{q_i})'(c_{\nu_i+p_i+1})| \geq e^{\sigma_0 q_i}.$$

For estimating the last factor in (9), we distinguish two cases. If $n \geq \nu_s + p_s$ then we use lemmas 2.4(3) and 2.3(1) once more and get

$$\begin{aligned} |(\varphi_a^{n-\nu_s+1})'(c_{\nu_s}) &= |(\varphi_a^{p_s+1})'(c_{\nu_s})| \cdot |(\varphi_a^{q_s})'(c_{\nu_s+p_s+1})| \\ &\geq e^{\sigma_0+\beta p_s} \min\{e^{\sigma_0}, |\varphi_a'(c_{\nu_s+p_s+1})|\} e^{\sigma_0(q_s-1)} \\ &\geq \text{const}\, e^{\sigma_0+\beta p_s} \varepsilon^{\gamma+\delta} e^{\sigma_0 q_s} \end{aligned}$$

(the final bound remains valid when $n = \nu_s + p_s$, *i.e.* $q_s = 0$). Replacing in (9),

$$(10)\qquad |(\varphi_a^n)'(c_1)| \geq \text{const}\, \varepsilon^{1-(1/\lambda)+\gamma+\delta} e^{(\sum_{i=0}^s \sigma_0 q_i + \sum_{i=1}^s (\sigma_0+\beta p_i))}.$$

Now, CP2(n) implies (recall that we take $\sigma_0 > 2\sigma$)

$$\sum_{i=0}^s \sigma_0 q_i + \sum_{i=1}^s (\sigma_0 + \beta p_i) \geq \sigma_0(n - P_n) + \beta P_n \geq \sigma_0 \frac{n}{2} > \sigma n$$

and the lemma follows by replacing this in (10) and assuming ε sufficiently small.

Suppose now that $\nu_s \leq n < \nu_s + p_s$. In this case we cannot take advantage of the estimates in lemma 2.4(3), as we did before. Instead, we use CP1(n), EG($n-1$), and the distortion estimate in lemma 2.4(1), to conclude that

$$|(\varphi_a^{n-\nu_s+1})'(c_{\nu_s})| = |\varphi'(c_{\nu_s})| \cdot |(\varphi_a^{n-\nu_s})'(c_{\nu_s+1})| \geq \text{const}\, \varepsilon^\gamma e^{-\alpha\nu_s}\ \text{const}\, e^{\sigma(n-\nu_s)}.$$

This gives

$$(11)\qquad |(\varphi_a^n)'(c_1)| \geq \text{const}\, \varepsilon^{1-(1/\lambda)+\gamma} e^{(\sum_{i=0}^{s-1} \sigma_0 q_i + \sum_{i=1}^{s-1}(\sigma_0+\beta p_i) - \alpha\nu_s + \sigma(n-\nu_s))}.$$

Now,

$$\sum_{i=0}^{s-1} \sigma_0 q_i + \sum_{i=1}^{s-1} (\sigma_0 + \beta p_i) - \alpha\nu_s + \sigma(n - \nu_s)$$

$$\geq \sigma_0(\nu_s - P_{\nu_s}) - \alpha\nu_s + \sigma(n - \nu_s) \geq \frac{\sigma_0}{2}\nu_s - \alpha\nu_s + \sigma(n - \nu_s) \geq \sigma n$$

as long as we take $2\alpha < \sigma_0 - 2\sigma$. Replacing in (11) (and assuming ε small) we get the conclusion of the lemma also in this case. Our argument is complete. □

Proposition 2.6. — *Suppose that some parameter $a \in [c + \rho\varepsilon, c + \varepsilon]$ satisfies CP1(n) and CP2(n) for all $n \geq 1$. Then $|(\varphi_a^n)'(c_1)| \geq e^{\sigma n}$ for all $n \geq 1$. Moreover, there is $\sigma_1 > 0$ such that for any $x \in [-a, a]$ satisfying $\varphi_a^j(x) \notin \{0, \pm c\}$ for all $j \geq 0$ we have $\limsup \frac{1}{n} \log |(\varphi_a^n)'(x)| \geq \sigma_1$.*

Proof. — The first claim follows directly from the previous lemma, by induction on n. Observe that the step $n = 1$ is an immediate consequence of L4 which, as we already remarked, implies $|c_1| < x_{\sqrt{2}}$ and so $|\varphi_a'(c_1)| > \sqrt{2} > e^{\sigma}$.

For the second statement we distinguish two cases according as to whether the orbit of x accumulates one of the critical points or not. If it does not, the result follows immediately from the previous lemmas which guarantee that $|(\varphi_a^j)'(x)| \geq Ce^{\beta j}$, $\forall j \geq 0$ which immediately implies the result taking $\sigma_1 < \beta$. If the orbit of x does accumulate one of the critical points then we claim that for every $N > 0$ there exists an $n \geq N$ such that $|(\varphi_a^n)'(x)| \geq e^{\beta n}$. This claim clearly implies the desired statement. Let $\mu_1 < \mu_2 < \cdots < \mu_k \leq N$ be all the returns of x to $\Delta^{\pm c}$ before time N and let $p_1, p_2, \ldots$ be the lengths of the corresponding binding periods. If $N \geq \mu_k + p_k$ then we have by the same arguments used in the proof of lemma 2.5 that $|(\varphi_a^N)'(x)| \geq e^{\beta N}$ which proves the claim in this case. If $N \in (\mu_k, p_k)$, just take $n = \mu_k + p_k + 1$ and repeat the argument above. This completes the proof of the claim and of the proposition. □

Remark 2.1. — A refinement of the previous arguments permits to show a stronger statement: $\liminf \frac{1}{n} \log |(\varphi_a^n)'(x)| \geq \sigma_2$ for some $\sigma_2 > 0$ and almost every point x. First one notes that this holds whenever x satisfies

$$d(\varphi^j(x)) = \min\{|\varphi^j(x)|, |\varphi^j(x) \pm c_j|\} \geq e^{-\alpha j}$$

for all $j \geq 0$, by using essentially the same argument as we did above for the critical orbits. Then, using the distortion bounds we have been deriving, one shows that for Lebesgue almost every point y there is some $k \geq 0$ such that $x = \varphi^k(y)$ is as above.

Finally, we make the simple, yet useful observation.

Lemma 2.7. — *Suppose that some parameter $a \in [c + \rho\varepsilon, c + \varepsilon]$ satisfies CP1(n) and CP2(n) for some $n \geq 1$. Let $1 \leq \mu_1 \leq \mu_2 \leq n$ be free iterates for the orbit of c. Then*

$$|(\varphi_a^{\mu_2 - \mu_1})'(c_{\mu_1}(a))| \geq \min\{|\varphi_a'(c_{\mu_1(a)})|, e^{\sigma_0}\} e^{\beta(\mu_2 - \mu_1 - 1)}.$$

In particular, if $|\varphi_a'(c_{\mu_1}(a))| \geq e^{\beta}$, *then* $|(\varphi_a^{\mu_2-\mu_1})'(c_{\mu_1}(a))| \geq e^{\beta(\mu_2-\mu_1)}$.

Proof. — The proof follows easily by arguments almost identical to those used in the proof of lemma 2.5. □

3. Partitions and distortion estimates

3.1. Preliminary distortion estimates. — In this section we set up the machinery which will enable us, in the next section, to estimate the size of the set of parameters satisfying conditions CP1(n) and CP2(n) for all $n \geq 1$. Most of this analysis deals with properties of the family of maps

$$c_j : \omega_0 \longrightarrow [-c-\varepsilon, c+\varepsilon], \qquad c_j(a) = \varphi_a^j(c),$$

where $\omega_0 = [c+\rho\varepsilon, c+\varepsilon]$ and $j \geq 1$. Our first result implies that the derivatives $c_j'(a)$ of such maps grow exponentially fast with j, as long as the phase-space derivatives $(\varphi_a^j)'(c_1(a)) = (\partial_x \varphi_a^j)(c_1(a))$ do.

Lemma 3.1. — *There is* $\eta > 1$ *such that if* $|(\varphi_a^j)'(c_1(a))| \geq e^{\sigma j}$ *for* $1 \leq j \leq n$, *then*

$$\frac{1}{\eta} \leq \frac{|c_{j+1}'(a)|}{|(\varphi_a^j)'(c_1(a))|} \leq \eta \qquad \textit{for all } 1 \leq j \leq n$$

Proof. — The arguments are fairly standard. Using the chain rule we can write

$$c_{j+1}' = \partial_a\varphi_a(c_j) + \partial_x\varphi_a(c_j)\partial_a\varphi_a(c_{j-1}) + \cdots + \partial_x\varphi_a^{j-1}(c_2)\partial_a\varphi_a(c_1) + \partial_x\varphi_a^j(c_1)c_1'$$

and so

$$\frac{c_{j+1}'}{(\varphi_a^j)'(c_1)} = \sum_{i=1}^{j} \frac{\partial_a\varphi_a(c_i)}{(\varphi_a^i)'(c_1)} + c_1'. \tag{12}$$

Note that $c_1' = 1$ and $|\partial_a\varphi_a(c_i)| = 1$ for each i. Hence, our hypothesis implies that (12) is bounded from above (in norm) by $\eta = \sum_{i=0}^{\infty} e^{-\sigma i}$.

The bound from below requires a more careful analysis. For the time being we restrict to $a = c$ and note that the first two terms in (12) are both positive (L4 gives $c_1 > 0$ and so $\partial_a\varphi_a(c_1) > 0$, $\varphi_a'(c_1) > 0$, and $\partial_a\varphi_a(c_1)\cdot(\varphi_a^2)'(c_1) > 0$). We distinguish three cases.

If $|\varphi_a'(c_2)| \geq \sqrt{2}$ and $|\varphi_a'(c_3)| \geq \sqrt{2}$ (*i.e.* $|c_i| \leq x_{\sqrt{2}}$ for $i = 2, 3$) then

$$|(\varphi_a^{2k+1})'(c_1)| \geq (\sqrt{2})^{2k}|\varphi_a'(c_1)| \quad \text{and} \quad |(\varphi_a^{2k+2})'(c_1)| \geq (\sqrt{2})^{2k}|(\varphi_a^2)'(c_1)|, \tag{13}$$

by Proposition 1.1. It follows that

$$\sum_{1\leq 2k+1\leq j} \frac{\partial_a\varphi_a(c_{2k+1})}{(\varphi_a^{2k+1})'(c_1)} \geq \frac{1}{\varphi_a'(c_1)}\left(1 - \sum_{k=1}^{\infty} 2^{-k}\right) \geq 0$$

and a similar estimate holds for the sum over even indices. Thus the quotient in (12) is bounded from below by $c_1'(a) = 1$, which proves the lemma in this case.

Suppose now that $|\varphi_a'(c_2)| \geq \sqrt{2}$ and $|\varphi_a'(c_3)| < \sqrt{2}$. Clearly, we still have the first estimate in (13) and so the sum of all the terms in (12) corresponding to odd values of i is again nonnegative. In order to bound the sum of the even terms we use Proposition 1.1 once more and get (recall that $|\varphi_a'(c_1)| \geq \sqrt{2}$ by L4) $|(\varphi_a^{2k})'(c_1)| \geq (\sqrt{2})^{2k}$ for each k. It follows that

$$\sum_{1\leq 2k\leq j} \frac{\partial_a\varphi_a(c_{2k})}{(\varphi_a^{2k})'(c_1)} \geq 0 - \sum_{k=2}^{\infty}(\sqrt{2})^{-2k} \geq -\frac{1}{2}.$$

Hence (12) is bounded from below by 1/2.

The case $|\varphi_c'(c_2)| < \sqrt{2}$ and $|\varphi_c'(c_3)| \geq \sqrt{2}$ is quite similar to the previous one. The second estimate in (13) is valid here and so the total contribution of the even terms in (12) is nonnegative. On the other hand, $|(\varphi_a^{2k-1})'(c_1)| \geq (\sqrt{2})^{2k-1}$ yields

$$\sum_{1\leq 2k-1\leq j} \frac{\partial_a\varphi_a(c_{2k-1})}{(\varphi_a^{2k-1})'(c_1)} \geq 0 - \sum_{k=2}^{\infty}(\sqrt{2})^{1-2k} \geq -\frac{1}{\sqrt{2}}$$

and so (12) is bounded from below by $1 - 1/\sqrt{2} > 1/4$.

Note also that we cannot have $|\varphi_a'(c_2)| < \sqrt{2}$ and $|\varphi_a'(c_3)| < \sqrt{2}$ simultaneously, by L5. This means that the lemma is proved for $a = c$. The general case now follows easily. Fix $l \geq 1$ large so that $\sum_{i>l} e^{-\sigma i} < 1/10$ and take ε to be small enough so that

$$\left|\frac{c_{j+1}'(a)}{(\varphi_a^j)'(c_1(a))} - \frac{c_{j+1}'(c)}{(\varphi_c^j)'(c_1(c))}\right| \leq \frac{1}{10} \quad \text{for all } j \leq l \text{ and all } a \in [c, c+\varepsilon].$$

It follows, immediately, that $c_{j+1}'(a)/(\varphi_a^j)'(c_1(a)) \geq (1/4) - (1/10) = (3/20)$ for all $j \leq l$. Moreover, for $j > l$ we have

$$\left|\frac{c_{j+1}'(a)}{(\varphi_a^j)'(c_1(a))} - \frac{c_{l+1}'(a)}{(\varphi_a^l)'(c_1(a))}\right| = \left|\sum_{i=l+1}^{j} \frac{\partial_a\varphi_a(c_i)}{(\varphi_a^i)'(c_1)}\right| \leq \sum_{l+1}^{\infty} e^{-\sigma i} < \frac{1}{10}$$

and so $c_{j+1}'(a)/(\varphi_a^j)'(c_1(a)) \geq (3/20) - (1/10) = (1/20)$. □

It follows from this lemma that as long as the space derivatives $(\varphi_a^j)'(c_1(a))$ are growing exponentially for all a belonging to some interval ω of parameter values, the maps $c_j : \omega \to c_j(\omega)$ are diffeomorphisms, since $|c_j'(a)| \geq 1/\eta|(\varphi_a^j)'(c_1(a))| \neq 0$. In particular the same is true for the maps $\Phi : c_i(\omega) \to c_j(\omega)$ defined by

$$x \to c_j \circ c_i^{-1}(x)$$

for $0 \leq i \leq j$, even though the space derivatives may not be growing exponentially between time i and time j. We have that

$$\Phi'(c_i(a)) = \frac{|c_j'(a)|}{|c_i'(a)|}, \qquad a \in \omega. \tag{14}$$

Thanks to this fact we can still estimate the average expansion of the "intermediate" images of ω.

Lemma 3.2. — *Suppose that we have $|(\varphi_a^j)'(c_1(a))| \geq e^{\sigma j}$ for all $1 \leq j \leq n$ and $a \in \omega$. Then, for all integers $1 \leq k \leq l \leq n$ we have for some $\xi \in \omega$*

$$\frac{1}{\eta^2}|(\varphi_\xi^{l-k})'(c_k(\xi))| \leq \frac{|c_l(\omega)|}{|c_k(\omega)|} \leq \eta^2|(\varphi_\xi^{l-k})'(c_k(\xi))|.$$

Proof. — Defining $\Phi : c_k(\omega) \to c_l(\omega)$ as above, we have by the Mean Value theorem $|c_l(\omega)| = |\Phi'(c_k(\xi))||c_k(\omega)|$ for some $\xi \in \omega$. Thus from Lemma 3.1 and the formula above for the derivative of Φ we have immediately the statement in the lemma. □

3.2. Partitions. — We shall now use the family of maps $\{c_j\}$, together with Lemma 3.1, to construct two nested families $\{F_n\}_{n\in\mathbb{N}}$ and $\{E_n\}_{n\in\mathbb{N}}$ of subsets of ω_0

$$\cdots \subseteq E_n \subseteq F_n \subseteq E_{n-1} \subseteq \cdots \subseteq \omega_0$$

and a monotone sequence $\{\mathcal{P}_n\}_{n\in\mathbb{N}}$ of families of subintervals of ω_0

$$\cdots \succ \mathcal{P}_n \succ \mathcal{P}_{n-1} \succ \cdots \succ \{\omega_0\} \quad (\text{given } \omega \in \mathcal{P}_n\ \exists\, \omega' \in \mathcal{P}_{n-1} \text{ with } \omega \subset \omega')$$

as follows. All the parameters belonging to F_n satisfy CP1(n): $d(c_j) \geq \varepsilon^\gamma e^{-\alpha j}$ for all $1 \leq j \leq n$. Each $\mathcal{P}_n$ is a partition of F_n into intervals. Moreover, E_n is a union of elements $\omega \in \mathcal{P}_n$ such that CP2(n): $P_j \leq j/2$ for all $1 \leq j \leq n$ and $a \in \omega$ holds. In view of Proposition 2.6, the parameters in $\mathcal{A}_\varepsilon^+ = \bigcap_{n\in\mathbb{N}} E_n$ will satisfy EG for all times (we shall take $\mathcal{A}^+ = \cup_\varepsilon \mathcal{A}_\varepsilon^+$, where the union is over all small values of $\varepsilon > 0$).

The construction of the objects described above is carried out inductively. As the first step of the induction we simply set $E_0 = F_0 = \omega_0$ and $\mathcal{P}_0 = \{\omega_0\}$. Now suppose that F_{n-1}, E_{n-1}, $\mathcal{P}_{n-1}$ have been defined and let us explain how parameters are excluded at the nth stage and the partition $\mathcal{P}_n$ and the sets F_n, E_n are constructed. For that we consider separately the cases $n < r_s/\alpha$ and $n \geq r_s/\alpha$, see Remark 3.5 below. In what follows we denote $\Delta_+ = \Delta_+^0 \cup \Delta_+^c \cup \Delta_+^{-c}$ and $\Delta = \Delta^0 \cup \Delta^c \cup \Delta^{-c}$. For $r > 0$ we let Δ_r denote the $\varepsilon^\gamma e^{-r}$-neighbourhood of the origin and of the critical points. Recall also that $r_\varepsilon = [\delta \log 1/\varepsilon]$ and $r_s = [(\gamma + 2\delta + \iota)\log 1/\varepsilon]$.

Suppose first that $n < r_s/\alpha$. Given any $\omega \in \mathcal{P}_{n-1}$ with $\omega \subset E_{n-1}$, there are two possibilities:

1 : If $c_n(\omega)$ does not intersect $\Delta_{[\alpha n]}$ then, by definition, ω is also an element of $\mathcal{P}_n$ and it is contained in F_n and in E_n.

2 : If $c_n(\omega) \cap \Delta_{[\alpha n]} \neq \varnothing$ then parameters have to be thrown out in order that CP1(n) hold. We write $\omega_e' = c_n^{-1}(\Delta_{[\alpha n]} \cap c_n(\omega))$ and we also let ω_e'' be the union of those connected components of $\omega \setminus \omega_e'$ whose image under c_n is completely contained in $\Delta_{[\alpha n]-1}$. Both these sets of parameters are excluded from the sequel of the argument: by definition

$$E_n \cap \omega = F_n \cap \omega = \omega \setminus (\omega_e' \cup \omega_e'')$$

and the elements of $\mathcal{P}_n$ contained in ω are precisely the connected components of $\omega \setminus (\omega_e' \cup \omega_e'')$. We observe, for future reference, that any such component $\tilde{\omega}$

contains an interval of the form I_r^0 with $|r| = [\alpha n] - 1$. We call this interval the *host interval* of $\widetilde{\omega}$ at the return n. Moreover this immediately implies that $|c_n(\widetilde{\omega})| \geq \varepsilon^\gamma e^{-[\alpha n]}$.

Thus we have for each $k < r_s/\alpha$ and $\omega \in \mathcal{P}_k$ a nested sequence of intervals

$$\omega = \omega_k \subseteq \omega_{k-1} \subseteq \cdots \subseteq \omega_1 \subseteq \omega_0 \text{ with } \omega_i \in \mathcal{P}_i, i = 0, \dots, k$$

and a sequence of *escape times*

$$1 \leq \nu_1 < \nu_2 < \cdots < \nu_s \leq k$$

defined by the fact that the situation described in case 2 occurs precisely at these times. The corresponding components $\omega_{\nu_i} \in \mathcal{P}_{\nu_i}$ are called *escaping components.* By definition we also call the original parameter interval $\omega_0 = (c + \varepsilon/\rho, c + \varepsilon)$ an escaping component. Notice that case 1 and 2 describe the only situations which can occur before time r_s/α. Since case 1 does not involve making any changes to existing parameter intervals (*i.e.* $\omega_i = \omega_{i-1}$ if $c_i(\omega_{i-1})$ satisfies the conditions described in 1) we have $\omega_{\nu_{i+1}} \subset \omega_{\nu_{i+1}-1} = \omega_{\nu_{i+1}-2} = \cdots = \omega_{\nu_i}$. As we shall see this is not true in general for iterates larger that r_s/α.

Now we treat the case $n \geq r_s/\alpha$. In order to define $\mathcal{P}_n$, F_n, E_n, we need a refinement of the partitions $\{I_r^*\}$, $* = 0, \pm c$, introduced in the previous section: for each $|r| \geq 1$ we let $\{I_{r,l}^* \colon 1 \leq l \leq r^2\}$ be the partition of I_r^* into r^2 intervals of equal length. We suppose that l is increasing in the same direction as r, *i.e.* as we get closer to the singularity or the critical points. As above we associate to every $\omega \in \mathcal{P}_{n-1}$ with $\omega \subset E_{n-1}$ a nested sequence of intervals $\omega = \omega_{n-1} \subseteq \cdots \subseteq \omega_0$ and a sequence of escape times

$$1 \leq \nu_1 < \cdots < \nu_l < r_s/\alpha \leq \nu_{l+1} < \cdots < \nu_s \leq n-1.$$

Each ω_i, $0 \leq i \leq n-1$, is just the element of $\mathcal{P}_i$ that contains ω. For $j \leq l$ the ν_j are exactly the escape times described above. However, for $l \leq j \leq s-1$ we also have between two consecutive escape times a (possibly empty) sequence of *return times*

$$\nu_j < \mu_{0,j} < \mu_{1,j} < \cdots < \mu_{q(j),j} < \nu_{j+1}$$

and similarly for $j = s$ we have

$$\nu_s < \mu_{0,s} < \cdots < \mu_{q(s),s} \leq n-1.$$

Moreover to each such sequence of return times is associated a sequence of integers $p_{0,j}, p_{1,j}, \dots, p_{q(j),j} \geq 0$.

As part of the inductive step of our construction we also explain when and how $\nu_{s+1}, \mu_{q(s)+1,s}$ and $p_{q(s)+1,s}$ are introduced, assuming that such sequences are defined for all iterates up to $n-1$. We consider four cases separately:

3 : If $q(s) > 0$ and $\mu_{q(s)+1,s} \leq n \leq \mu_{q(s),s} + p_{q(s),s}$ then $\omega \in \mathcal{P}_n$ and $\omega \subset E_n \subset F_n$. Moreover, we leave the sequences unchanged.

4 : If $c_n(\omega)$ does not intersect $\Delta^{\pm c} \cup \Delta^0_{r_s}$ or if this intersection is completely contained in one of the extreme subintervals of $\Delta^{\pm c} \cup \Delta^0_{r_s}$, *i.e.*

$$c_n(\omega) \cap (\Delta^{\pm c} \cup \Delta^0_{r_s}) \subseteq I^0_{\pm(r_s+1),(r_s+1)^2} \cup I^{\pm c}_{\pm(r_\varepsilon+1),(r_\varepsilon+1)^2},$$

then once more we let $\omega \in \mathcal{P}_n$ and $\omega \subset E_n \subset F_n$, and we keep the sequences unchanged.

From now on we assume that neither 3 nor 4 hold.

5 : If $c_n(\omega)$ intersects $\Delta^{\pm c} \cup \Delta^0_{r_s}$ but does not properly contain any subinterval $I^{\pm c}_{r,l}$ with $|r| > r_\varepsilon$ or $I^0_{r,l}$ with $|r| > r_s$, we set $\omega \in \mathcal{P}_n$ and $\omega \subset F_n$. Moreover, we let $\mu_{q(s)+1} = n$ and $p_{q(s)+1} = \min\{p(c_n(a))\colon a \in \omega\}$, recall Lemma 2.4, if $c_n(\omega) \cap \Delta^0_{r_s} = \varnothing$, and $p_{q(s)+1} = 0$ otherwise. We say that n is an *(inessential) return* for $\omega \in \mathcal{P}_n$ and that $p_{q(s)+1}$ is the length of the associated *binding period.* Note that $c_n(\omega)$ is contained in the union of at most two $I^*_{s,m}$. We shall prove in Lemma 3.4 that CP1(n) is automatically satisfied in this case. Then we take $\omega \subset E_n$ if all $a \in \omega$ satisfy CP2(n) and $E_n \cap \omega = \varnothing$ otherwise.

6 : If $c_n(\omega)$ intersects $\Delta^{\pm c} \cup \Delta^0_{r_s}$ and contains some $I^{\pm c}_{r,l}$ with $|r| > r_\varepsilon$ or $I^0_{r_s}$ with $|r| > r_s$, we carry out the following construction. We start by excluding the parameters which do not satisfy CP1(n). More precisely, we let

$$\omega'_e = c_n^{-1}(\Delta_{[\alpha n]} \cap c_n(\omega))$$

and ω''_e be the union of the connected components of $\omega \setminus \omega'_e$ whose image under c_n contains no subinterval $I^*_{s,m}$. By definition $F_n \cap \omega = \overline{\omega} = \omega \setminus (\omega'_e \cup \omega''_e)$. Then we partition $\overline{\omega}$ into subintervals $\overline{\omega} = (\cup_{r,l}\omega_{r,l}) \cup (\cup_i \widetilde{\omega}_i)$ where the first union runs over some subset of pairs (r,l) with $r_\varepsilon < |r| \le [\alpha n]$ or $r_s < |r| \le [\alpha n]$ (depending on whether n is a return to $\Delta^{\pm c}$ or Δ^0), the second one involves at most two $\widetilde{\omega}_i$, and

a : $c_n(\omega_{r,l}) \supset I^*_{r,l}$ but contains no other interval $I^*_{s,m}$ (thus it is contained in the union of $I^*_{r,l}$ with the two $I^*_{s,m}$ adjacent to it). We call $I^*_{r,l}$ the *host interval* associated to $\omega_{r,l}$ at time n.

b : $c_n(\widetilde{\omega}_i)$ is disjoint from $\Delta^{\pm c} \cup \Delta^0_{r_s}$ but contains some $I^{\pm c}_{r,1}$ with $|r| = r_\varepsilon$ or $I^0_{r,1}$ with $|r| = r_s$. Again we call this interval the *host interval* associated to $\widetilde{\omega}$ and time n.

The elements of $\mathcal{P}_n$ contained in ω are precisely these $\omega_{r,l}$ and $\widetilde{\omega}_i$. For $\omega_{r,l}$ we let $\mu_{q(s)+1} = n$ and $p_{q(s)+1} = \min\{p(c_n(a))\colon a \in \omega_{r,l}\}$ if $* = \pm c$ and $p_{q(s)+1} = 0$ if $* = 0$. In particular n is an *essential return* time for each $\omega_{r,l} \in \mathcal{P}_n$. For the intervals $\widetilde{\omega}_i$ described in 6b we let $\nu_{s+1} = n$. In particular n is an *escape time* for $\widetilde{\omega} \in \mathcal{P}_n$ and these intervals are *escaping components.* Finally, $E_n \cap \omega$ consists of the union of the intervals described above which satisfy CP2(n).

This completes the inductive definition of the sets E_n, F_n, the partitions $\mathcal{P}_n$, and the sequences ν_j, μ_j, and p_j.

Remark 3.1. — Host intervals can give some indication as to the type of situation we are dealing with. So, if $\omega \in \mathcal{P}_i$ has, at time i, an associated host interval of the form I_r^0, $|r| = [\alpha n] - 1$ then this implies that $i = \nu_j$ is an escape time and that ω is created at time ν_j as a consequence of a situation like the one described in case 2. Similarly if the host interval is of the form $I_{r,1}^0$, $|r| = r_s$ or $I_{r,1}^{\pm c}$, $|r| = r_\varepsilon$ then $i = \nu_j$ is also an escape time as described in 6b. On the other hand if the host interval is of the form $I_{r,l}^0$ with $|r| \geq r_s + 1$, $1 \leq l \leq r^2$ or $I_{r,l}^{\pm c}$ with $|r| \geq r_\varepsilon + 1$, $1 \leq l \leq r^2$ then $i = \mu_{k,j}$ for some k, j and $\mu_{k,j}$ is either an inessential return as described in case 5 or an essential return as described in case 6a.

A main difference between the latter two cases is that for essential returns we have upper and lower bounds for the length of $c_{\mu_{k,j}}(\omega)$ in terms of the associated host interval. More precisely, recall case 6a, $c_{\mu_{k,j}}(\omega)$ contains some $I_{r,l}^*$ and is contained in the union of $I_{r,l}^*$ and its two adjacent intervals of the form $I_{s,m}^*$. Therefore we have $\varepsilon^\gamma e^{-r}/r^2 \leq |c_{\mu_{k,j}}(\omega)| \leq 10\varepsilon^\gamma e^{-r}/r^2$. In the case of inessential returns we have the same upper bound but no a priori lower bound.

Remark 3.2. — We will sometimes talk about the sequence of escape times and returns associated to a single parameter value a or a subinterval $\omega' \subset \omega \in \mathcal{P}_n$ which does not itself necessarily belong to any partition (although we will always consider subsets of intervals which do belong to some $\mathcal{P}_n$). In these cases the sequences are just those associated to the the interval $\omega \in \mathcal{P}_n$ to which a or ω' belong.

Remark 3.3. — We will frequently talk about *returning situations* or *escape times* at time n for some interval $\omega_{n-1} \in \mathcal{P}_{n-1}$ (and *not* in $\mathcal{P}_n$). In these cases we will just be referring to the fact that $c_n(\omega_{n-1})$ intersects a neighbourhood of the origin or of the critical points in a way which is described in one of cases 2,5 or 6 above. Therefore, at time n some action may be required (parameter exclusions, subdivision of ω_{n-1} into smaller intervals) which yields the final classification of the surviving pieces of ω_{n-1} into pieces (now belonging to $\mathcal{P}_n$) for which n is either a return (essential or inessential) or an escape time.

Remark 3.4. — Notice that the definition of the binding period $p = p(\omega)$ given here does not completely coincide with the definition given in Lemma 2.4 for a fixed parameter value a. However all the estimates obtained in that lemma continue to hold for the slightly shorter binding period defined here.

To simplify the exposition we will often refer to a generic host interval of the form $I_{r,l}^*$. This will include the host intervals which occur in case 2 which, strictly speaking, are of the form I_r^*. Moreover we shall often suppose that $r > 0$ since most of the times we are only interested in the norm of r and not its sign.

Remark 3.5. — The condition $n < r_s/\alpha$ means that $\Delta_{[\alpha n]} \supset \Delta_{r_s}$ and so

$$c_n(\widetilde{\omega}) \cap \Delta^{\pm c} = \varnothing \quad \text{for all } \widetilde{\omega} \in \mathcal{P}_n.$$

Indeed, $c_n(a) \in \Delta_+^{\pm c}$ would imply $c_{n-1}(a) \in I_r^0$ with $|r| \approx r_c \approx ((1/\lambda) - \gamma)\log\varepsilon^{-1} \gg (2(\gamma+\delta)+\iota)\log\varepsilon^{-1} \approx r_s$, recall (3) and (4), and so $c_{n-1}(a) \in \Delta_{r_s}^0$. This is a contradiction, as such a parameter a would have been excluded already at time $n-1$. By construction (and the remark we have just made) all the elements of $\mathcal{P}_n$ obtained in that case contain some interval I_r^0 with $|r| = [\alpha n]$. Combined with Lemma 2.3 this gives

$$|(\varphi_a^j)'(c_1)| \geq e^{\sigma_0 j} \quad \text{for all } 1 \leq j \leq n \text{ and } a \in F_n.$$

Finally, r_s/α can be made arbitrarily large by fixing ε and α sufficiently small.

Remark 3.6. — We shall show below that for $\omega \in \mathcal{P}_n$ the distortion

$$\sup\{|(\varphi_a^n)'(c_1(a))/(\varphi_b^n)'(c_1(b))| : a, b \in \omega\}$$

is bounded above by some constant. It is crucial to the overall argument that this constant is independent of ω as well as of n. The proof of this fact relies on the exponential growth of the intervals $c_j(\omega)$ as well as on the fact that $c_j(\omega)$ is small compared to its distance from the critical points and the discontinuity, where the distortion explodes. This is why no action is required in case 5 above whereas we need to cut ω into smaller pieces in case 6.

Throughout the rest of the paper we will use C (resp. $\widetilde{C}$) to denote a generic small (resp. large) positive constant independent of ε, ω or the iterate under consideration.

Lemma 3.3. — *Let $\omega \in \mathcal{P}_{n-1}, \omega \subset E_{n-1}$ and suppose that n is a return for $\omega \in \mathcal{P}_{n-1}$. Let $\nu \leq n-1$ be the last essential return or escape time of ω before time n and let $I_{r,l}^*$ be the host interval associated to ν. Then we have the following estimates.*

(1) *If $* = 0$,* i.e. *if ν is a return or escape time associated to Δ^0, and*
   (a) *$n = \nu + 1$ is a return to $\Delta^{\pm c}$: $|c_n(\omega)| \geq \varepsilon^{1+\iota}$;*
   (b) *$n > \nu + 1$ is a return to Δ^0:*
   $$|c_n(\omega)| \geq C(\varepsilon^\gamma e^{-r})^{2\lambda}/r^2 \geq \varepsilon^{\gamma-\beta/\sigma}e^{-(1-\beta/\sigma)r};$$
   (c) *$n > \nu + 1$ is a return to $\Delta^{\pm c}$:*
   $$|c_n(\omega)| \geq C(\varepsilon^\gamma e^{-r})^{2\lambda}\varepsilon^{1-1/\lambda}/r^2 \geq \varepsilon^{1-1/\lambda}\varepsilon^{\gamma-\beta/\sigma}e^{-(1-\beta/\sigma)r}.$$

(2) *If $* = \pm c$,* i.e. *if ν is a return or an escape time associated to $\Delta^{\pm c}$, and $r_\varepsilon \leq r \leq (\gamma+\delta+\iota)\log 1/\varepsilon$ then n is* necessarily *a return to Δ^0 and*
   $$|c_n(\omega)| \geq \varepsilon^{-\iota/2}\varepsilon^\gamma e^{-r_s} \geq \varepsilon^{2(\gamma+\delta)+\iota/2}.$$

(3) *If $* = \pm c$ and $r \geq (\gamma+\delta+\iota)\log 1/\varepsilon$ and*
   (a) *n is a return to Δ^0:*
   $$|c_n(\omega)| \geq \varepsilon^{2\gamma+\delta+2\beta\gamma/\sigma}e^{-2\beta r/\sigma}/r^2 \geq \varepsilon^{\gamma-\beta/\sigma}e^{-(1-\beta/\sigma)r};$$
   (b) *n is a return to $\Delta^{\pm c}$:*
   $$|c_n(\omega)| \geq \varepsilon^{1-1/\lambda+2\gamma+\delta+2\beta\gamma/\sigma}e^{-2\beta r/\sigma}/r^2 \geq \varepsilon^{1-1/\lambda}\varepsilon^{\gamma-\beta/\sigma}e^{-(1-\beta/\sigma)r}.$$

Proof. — The basic strategy of the proof is to estimate $|(\varphi_a^{n-\nu})'(c_\nu(a))|$ for elements $a \in \omega$. Applying Lemma 3.2 and the Mean Value theorem we can then carry this expansion over to parameter space and get

$$|c_n(\omega)| \geq \inf\{|(\varphi_a^{n-\nu})'(c_\nu(a))| : a \in \omega\}|c_\nu(\omega)|/\eta^2. \tag{15}$$

Keeping in mind that the definition of host interval implies $|c_\nu(\omega)| \geq \varepsilon^\gamma e^r/r^2$ we shall get the desired result in each case.

Suppose first that ν is a return to Δ^0 and that $n = \nu + 1$ is a return to $\Delta^{\pm c}$. Then $c_\nu(\omega) \subset \varphi_a^{-1}(\Delta^{\pm c})$ for some $a \in \omega$ and Lemma 2.2 gives

$$|\varphi_a'((c_\nu(a))| \geq C(\varepsilon^\gamma e^{-r})^{\lambda-1} \geq \varepsilon^{1-1/\lambda}$$

since $r \approx (1/\lambda - \gamma)\log 1/\varepsilon$. Moreover $|c_\nu(\omega)| \geq C\varepsilon^{1/\lambda}/((1/\lambda+\gamma)\log 1/\varepsilon)^2$ and so

$$|c_\nu(\omega)| \geq C\varepsilon/((1/\lambda+\gamma)\log 1/\varepsilon)^2 \geq \varepsilon^{1+\iota}$$

for small $\iota > 0$. This proves (1a). Now let ν be a return to Δ^0 and consider first the situation in which $|r| \geq r_s + 1$. We have that for all $a \in \omega$, by Lemmas 2.2 and 2.7 we have

$$|(\varphi_a^2)'(c_\nu(a))| \geq (\varepsilon^\gamma e^r)^{2\lambda-1} \text{ and } |(\varphi_a^{n-\nu-2})'(c_{\nu+2}(a))| \geq Ce^{\beta(n-\nu-2)}.$$

For this last statement we have used the fact that $c_\nu \in \Delta^0$ implies $|c_{\nu+1}| > x_{\sqrt{2}}$ and this in turn implies $|c_{\nu+2}| < x_{\sqrt{2}}$ (see condition L4) which in particular means that $|\varphi'(c_{\nu+2})| > e^\beta$. Applying Lemma 3.2 we get $|c_n(\omega)| \geq C(\varepsilon^\gamma e^{-r})^{2\lambda}/\eta^2 r^2$ proving the first inequality in (1b). The second inequality just follows by taking $\beta/\sigma < 1 - 2\lambda$. If n is a return to $\Delta^{\pm c}$ then we have an additional factor of $\varepsilon^{1-1/\lambda}$ coming from the large derivative in $\varphi^{-1}(\Delta^{\pm c})$ and we get (1c). If $r \leq r_s$ we apply the same arguments to the subinterval $\overline{\omega} \subset \omega$ where $c_\nu(\overline{\omega}) = I^*_{r_s,1}$ and get $|c_n(\overline{\omega})| \geq C(\varepsilon^\gamma e^{-r})^{2\lambda}/\eta^2 r^2$ respectively $|c_n(\omega)| \geq C\varepsilon^{1-1/\lambda}(\varepsilon^\gamma e^{-r})^{2\lambda}/\eta^2 r^2$ which yields the desired result since $|c_n(\omega)| \geq |c_n(\overline{\omega})|$.

Now suppose that ν is a return to $\Delta^{\pm c}$. If $|r| = r_\varepsilon$ then the binding period has zero length by definition and we simply have, defining $\overline{\omega} = c_\nu^{-1}(I^{\pm c}_{r_\varepsilon,1}) \subset \omega$,

$$|c_n(\omega)| \geq |c_n(\overline{\omega})| \geq C\varepsilon^{\gamma+\delta}(\varepsilon^\gamma e^{-r_\varepsilon})/\eta^2 r_\varepsilon^2 \geq C\varepsilon^{2(\gamma+\delta)}/\eta^2 r_\varepsilon^2 \geq \varepsilon^{-\iota/2}\varepsilon^\gamma e^{r_s}.$$

If $|r| \geq r_\varepsilon + 1$ we have by by Lemmas 2.4 and 2.7 that, for all $a \in \omega$,

$$|(\varphi_a^{n-\nu})'(c_\nu(a))| \geq \varepsilon^{2\beta\gamma/\sigma} e^{(1-2\beta/\sigma)r} \cdot \inf\{\varphi_a'(c_{\nu+p+1}(a)), e^{\sigma_0}\} e^{\beta(n-\nu-p-1)} \tag{16}$$

$$\geq \varepsilon^{\gamma+\delta+2\beta\gamma/\sigma} e^{(1-2\beta/\sigma)r} \tag{17}$$

if n is a return to Δ^0 and $|(\varphi_a^{n-\nu})'(c_\nu(a))| \geq \varepsilon^{1-1/\lambda+\gamma+\delta+2\beta\gamma/\sigma} e^{(1-2\beta/\sigma)r}$ if it is a return to $\Delta^{\pm c}$. Notice that in both cases we have used $|\varphi'(c_{\nu+p+1}(a))| \geq \varepsilon^{\gamma+\delta}$. Now applying Lemma 3.2 and equations (15)(16) we get

$$|c_n(\omega)| \geq \varepsilon^{2\gamma+\delta+2\beta\gamma/\sigma} e^{-2\beta r/\sigma}/\eta^2 r^2. \tag{18}$$

Now if $r \le (\gamma + \delta + \iota)\log 1/\varepsilon$ then we have from (18)

$$|c_n(\omega)| \ge \varepsilon^{2\gamma+\delta+2\beta\gamma/\sigma+2\beta(\gamma+\delta+\iota)/\sigma}/r^2 \ge \varepsilon^{(2\gamma+\delta)+\iota/2} \ge \varepsilon^{-\iota/2}\varepsilon^{\gamma}e^{r_s}.$$

If $r > (\gamma+\delta+\iota)\log 1/\varepsilon$ then we write $e^{-2\beta r/\sigma} = e^{-(1-\beta/\sigma)r}e^{(1-\beta/\sigma)r}$ and $e^{(1-\beta/\sigma)r} \ge \varepsilon^{-(1-\beta/\sigma)(\gamma+\delta+\iota)}$ and therefore, using (18),

$$|c_n(\omega)| \ge \varepsilon^{2\gamma+\delta+2\beta\gamma/\sigma-(\gamma+\delta+\iota)+\beta/\sigma(\gamma+\delta+\iota)}e^{-(1-\beta/\sigma)r} \ge \varepsilon^{\gamma-\beta/\sigma}e^{-(1-\beta/\sigma)r}.$$

This concludes the proof of the lemma. □

We now formulate some easy consequences of these estimates.

Lemma 3.4. — *If n is an inessential return for some $\omega \in \mathcal{P}_{n-1}$ with $\omega \subset E_{n-1}$, then CP1(n) holds for every $a \in \omega$.*

Proof. — We claim that

$$|c_n(\omega)| \ge 6\varepsilon^{\gamma}e^{-[\alpha n]}. \tag{19}$$

This implies the statement in the lemma for the following reason. Suppose by contradiction that some $a \in \omega$ did not satisfy CP1(n), *i.e.* $d(c_n(a)) \le \varepsilon^{\gamma}e^{-\alpha n}$ and in particular $c_n(\omega) \cap \Delta^*_{[\alpha n]} \ne \varnothing, * = \pm c, 0$. Then, either $c_n(\omega) \subset \Delta^*_{[\alpha n]-1}$ or $c_n(\omega) \supset I^*_{[\alpha n]}$. The second alternative is not possible since it would contradict the fact that n is an inessential return. The first alternative cannot happen either since the claim implies $|c_n(\omega)| > 2e\varepsilon^{\gamma}e^{[\alpha n]-1} = |\Delta^*_{[\alpha n]-1}|$. Thus we have reduced the proof of the lemma to that of (19). However this is an easy consequence of Lemma 3.3. Indeed recall that $r \le \alpha n$ and therefore if the last essential return $\nu \le n-1$ occurred in Δ^0 we have

$$|c_n(\omega)| \ge C\varepsilon^{2\lambda\gamma}e^{-2\lambda r}/r^2 \gg 6\varepsilon^{\gamma}e^{-[\alpha n]}.$$

If ν is return to $\Delta^{\pm c}$ and $|r| = r_\varepsilon$ then the result follows immediately from part (2) in Lemma 3.3 keeping in mind that returns to $\Delta^{\pm c}$ can occur only for iterates $n \ge \alpha/r_s$. If $|r| \ge r_\varepsilon + 1$ we distinguish two cases. Suppose first that $|r| \ge (\gamma+\delta+\iota)\log 1/\varepsilon$. Then

$$\begin{aligned}|c_n(\omega)| &\ge \varepsilon^{2\gamma+\delta+2\beta\gamma/\sigma}e^{-r}e^{1-2\beta/\sigma}r/r^2 \ge \varepsilon^{2\gamma+\delta+2\beta\gamma/\sigma-(1-2\beta/\sigma)(\gamma+\delta+\iota)}e^{-r}\\ &\ge \varepsilon^{\gamma+2\beta\gamma/\sigma-\iota+2\beta(\gamma+\delta+\iota)/\sigma}e^{-r} \ge \varepsilon^{-\iota/2}\varepsilon^{\gamma}e^{-r} \gg 6\varepsilon^{\gamma}e^{-[\alpha n]}.\end{aligned}$$

Now suppose that $r < (\gamma+\delta+\iota)\log 1/\varepsilon$. Since $\nu \le n-1$ is a return to $\Delta^{\pm c}$ we also have $n \ge r_s/\alpha = ([\gamma+2\delta+\iota]/\alpha)\log 1/\varepsilon$. Therefore it is sufficient to show that $|c_n(\omega)| \ge 6\varepsilon^{\gamma}e^{-r_s} \ge \varepsilon^{2(\gamma+\delta)+\iota}$. This follows from part (2) of Lemma 3.3 which gives

$$|c_n(\omega)| \ge \varepsilon^{2\gamma+\delta+2\beta\gamma/\sigma}e^{-2\beta r/\sigma} \ge \varepsilon^{2\gamma+\delta+2\beta\gamma/\sigma+2\beta(\gamma+\delta+\iota)/\sigma} \gg 6\varepsilon^{2(\gamma+\delta)+\iota}.$$

This concludes the proof of the lemma. □

Lemma 3.5. — *Suppose that $\omega \in \mathcal{P}_{n-1}, \omega \subset E_{n-1}$ is an escaping component created at some time $\nu \le n-1$. Then, if n is a return for ω we have*

$$|c_n(\omega)| \ge \varepsilon^{-\iota/2}\varepsilon^{\gamma}e^{-r_s}.$$

In particular n is a return to Δ^0 for ω and there is a component $\widetilde{\omega} \subset \omega$ for which n is an escape time.

Proof. — The statement in the lemma follows immediately from Lemma 3.3. If ν is a return to $\Delta^{\pm c}$ the result is part (2). If ν is a return to Δ^0 then we have $|c_n(\omega)| \geq C(\varepsilon^\gamma e^{-r_s})^{2\lambda}/r_s^2 \gg \varepsilon^{-\iota/2}\varepsilon^\gamma e^{-r_s}$. □

3.3. More distortion estimates. — For the following lemma we fix ε' independent of ε. Let $\Delta_{\varepsilon'}$ and $\Delta_{2\varepsilon'}$ denote respectively ε' and $2\varepsilon'$ neighbourhoods of the origin and of the critical points. We suppose that ε' is chosen sufficiently small so that conditions M1 and M2 hold in $\Delta_{2\varepsilon'}$. Let $I \subset [-a, a]$ be an interval. For each $x \in I$ define $d(x) = \min\{|x|, |x \pm c|\}$ and $d(I) = \inf\{d(x) : x \in I\}$. Finally let $\widetilde{D}(I) = \sup\{|\varphi''(x)/\varphi'(y)| : x, y \in I\}$. We call an interval *admissible* if $I \cap \Delta_{\varepsilon'} \neq \varnothing$ implies $|I| \leq \varepsilon'$.

Lemma 3.6. — *For any constant $C_1 > 0$ there exists a constant $C_2 > 0$ such that if I is an admissible interval then*

$$|I| \leq C_1 d(I) \implies \widetilde{D}(I) \leq C_2/d(I).$$

Proof. — If $I \cap \Delta_{\varepsilon'} = \varnothing$ then both $\varphi'(x)$ and $\varphi''(x)$ are bounded above and below by constants which depend only on the map and on ε', and the statement in the lemma follows immediately. So suppose that $I \cap \Delta_{\varepsilon'} \neq \varnothing$. Then, since I is an admissible interval we have $|I| \leq \varepsilon'$ and in particular $I \subset \Delta_{2\varepsilon'}$. Therefore either $I \subset \Delta^0_{2\varepsilon'}$ or $I \subset \Delta^{\pm c}_{2\varepsilon'}$. If $I \subset \Delta^0_{2\varepsilon'}$ then, by condition M1,

$$|\varphi''(x)| \leq \widetilde{C} d(I)^{\lambda-2}$$

and

$$|\varphi'(x)| \geq C(d(I) - |I|)^{\lambda-1} \geq C((1 - C_1)d(I))^{\lambda-1} \geq Cd(I)^{\lambda-1}.$$

and so $\widetilde{D}(I) \leq C_2/d(I)$ for some constant $C_2 > 0$. If $I \subset \Delta^{\pm c}_{2\varepsilon'}$ then

$$|\varphi''(x)| \leq \widetilde{C} \text{ and } |\varphi'(x)| \geq Cd(I)$$

which immediately gives $\widetilde{D}(I) \leq C_2/d(I)$. □

Lemma 3.7. — *There exists a constant $A > 1$ (independent of ε or n) such that if $\omega \in \mathcal{P}_{n-1}$ with $\omega \subset E_{n-1}$ and n is a return to $\Delta^{\pm c}$ then*

$$\left| \frac{(\varphi^k_{\overline{a}})'(c_1(\overline{a}))}{(\varphi^k_a)'(c_1(a))} \right| \leq A \qquad \textit{for all } \overline{a}, a \in \omega \textit{ and all } 0 \leq k \leq n-1.$$

If n is a return to Δ^0 we have the same result for any $\overline{a}, a$ belonging to a subinterval $\overline{\omega} \subset \omega$ with $|c_n(\overline{\omega})| \leq \max\{(\varepsilon^\gamma e^{-\alpha n})^{2\lambda}, \varepsilon^{2(\gamma+\delta)}\}$.

Proof. — We shall prove the result for the case $k = n - 1$. It will be apparent from the proof that the result holds for all other values of k as well. Write $c_i = \varphi_a^i(c)$ and $\overline{c}_i = \varphi_{\overline{a}}^i(c)$. By the chain rule

$$\left|\frac{(\varphi_{\overline{a}}^k)'(c_1(\overline{a}))}{(\varphi_a^k)'(c_1(a))}\right| = \prod_{i=1}^k \left|1 + \frac{\varphi_{\overline{a}}'(\overline{c}_i) - \varphi_a'(c_i)}{\varphi_a'(c_i)}\right| \leq \prod_{i=1}^k 1 + |A_i|,$$

letting $A_i = (\varphi_{\overline{a}}'(\overline{c}_i) - \varphi_a'(c_i))/\varphi_a'(c_i)$, and thus the proof reduces to showing that $\sum_{i=1}^k |A_i|$ is bounded above by some constant independent of ε, n, or ω. By the Mean Value theorem we also have $|\varphi_{\overline{a}}'(\overline{c}_i) - \varphi_a'(c_i)| \leq |\overline{c}_i - c_i|\varphi''(\xi)$ for some $\xi \in c_i(\omega)$. Therefore we have

$$|A_i| \leq |c_i(\omega)| \cdot \widetilde{D}(c_i\omega). \tag{20}$$

Let $1 < \mu_1 < \mu_2 < \cdots < \mu_s < \mu_{s+1} = n$ be the essential and inessential returns (cases 5 and 6a) of ω in the time interval $[1, n]$. Notice that we do not include escape times in this list. Let $p_1, \ldots, p_s$ be the corresponding binding periods as defined in the previous subsection (recall that $p_j = 0$ if μ_j is a return close to Δ^0) and $r_1, \ldots, r_s$ be the values associated to the corresponding time intervals.

We start by considering the case in which the sequence of essential and inessential returns is empty, *e.g.* if $n < r_s/\alpha$. Then n is necessarily a return to Δ^0 for otherwise $n - 1$ would have been such a return. We suppose without loss of generality that $|c_n(\omega)| \leq (\varepsilon^\gamma e^{-\alpha n})^{2\lambda}$, for otherwise we could restrict ourselves to some subinterval $\overline{\omega}$ for which this condition is satisfied. Let ε' be the constant fixed in Lemma 3.6 and suppose first that $c_i(\omega) \cap \Delta_{\varepsilon'} = \varnothing$. Then $d(c_i(\omega)) \geq \varepsilon'$ and $|\varphi_a'(c_i(\omega))| \geq C\varepsilon'$ for all $a \in \omega$, and therefore by the standard arguments which we have used repeatedly above we get $|c_i(\omega)| \leq \widetilde{C}e^{-\sigma_0(n-1)}(\varepsilon^\gamma e^{-\alpha n})^{2\lambda} \leq \widetilde{C}d(c_i(\omega))$. Then by Lemma 3.6 we get $|A_i| \leq \widetilde{C}e^{-\sigma_0(n-1)}$. Now suppose that $c_i(\omega) \cap \Delta_{\varepsilon'}^0 \neq \varnothing$. A preliminary estimate for the length of $c_i(\omega)$ is given by

$$|c_i(\omega)| \leq e^{-\sigma_0(n-1)}|c_n(\omega)| \leq e^{-\sigma_0(n-1)}(\varepsilon^\gamma e^{-\alpha n})^{2\lambda} \ll \varepsilon'. \tag{21}$$

This shows that $c_i(\omega)$ is an admissible interval. Now we need to obtain a stronger estimate to show that it actually satisfies the hypothesis of Lemma 3.6. We distinguish two cases according as to whether $d(c_i(\omega)) > (\varepsilon^\gamma e^{-\alpha n})^{2\lambda}$ or $d(c_i(\omega)) \leq (\varepsilon^\gamma e^{-\alpha n})^{2\lambda}$. In the first case we have from (21) that the hypothesis of the lemma are satisfied and $|A_i| \leq \widetilde{C}|c_i(\omega)|/d(c_i(\omega)) \leq \widetilde{C}e^{-\sigma_0(n-1)}$. In the second case we have that the *maximum* distance between $c_i(\omega)$ and the origin, for $a \in \omega$ is

$$|c_i(\omega)| + d(c_i(\omega)) \leq 2(\varepsilon^\gamma e^{-\alpha n})^{2\lambda} \ll \varepsilon'.$$

Therefore $c_i(\omega)$ is entirely contained in the region in which condition M1 applies and it is easy to see by a simple variation of the argument in the proof of Lemma 2.2 that we have, for any $a \in \omega$,

$$|(\varphi_a^2)'(c_i(a))| \geq (\varepsilon^\gamma e^{-\alpha n})^{2\lambda(2\lambda-1)}.$$

From this we obtain an improved estimate for the size of $c_i(\omega)$, namely

$$|c_i(\omega)| \le e^{-\sigma_0(n-1)}(\varepsilon^\gamma e^{-\alpha n})^{2\lambda(2\lambda-1)}|c_n(\omega)| \le e^{-\sigma_0(n-1)}(\varepsilon^\gamma e^{-\alpha n})^{(2\lambda)^2}.$$

Then, from this and Lemma 3.6, $|A_i| \le |c_i(\omega)|/d(c_i(\omega)) \le Ce^{-\sigma_0(n-1)}$. Now suppose that $c_i(\omega) \cap \Delta^{\pm c}_{\varepsilon'} \ne \varnothing$. Since $|\varphi'(c_i(a))| \approx d(c_i(\omega)$ we have

$$|c_i(\omega)| \le e^{-\sigma_0(n-1)}(d(c_i(\omega)))^{-1}|c_n(\omega)| \le e^{-\sigma_0(n-1)}(d(c_i(\omega)))^{-1}(\varepsilon^\gamma e^{-\alpha n})^{2\lambda}.$$

Moreover, in this case we necessarily have $d(c_i(\omega)) \ge (\varepsilon^\gamma e^{-\alpha n})^{2\lambda}$ for the following reason. Since $\omega \subset E_{n-1}$ and $i \le n-1$ we have that, by definition, all $a \in \omega$ satisfy condition CP1 up to time $n-1$ and, in particular, at time $i-1$. Therefore $d(c_{i-1}(\omega)) \ge \varepsilon^\gamma e^{-\alpha(i-1)} \ge \varepsilon^\gamma e^{-\alpha n}$. Then because of the form of the map near the origin (condition M1) this implies $d(c_i(\omega)) \ge (\varepsilon^\gamma e^{-\alpha(i-1)})^\lambda \gg (\varepsilon^\gamma e^{-\alpha n})^{2\lambda}$. Therefore $c_i(\omega)$ is an admissible interval and we have

$$|c_i(\omega)| \le e^{-\sigma_0(n-i)}(\varepsilon^\gamma e^{-\alpha n})^{-\lambda}(\varepsilon^\gamma e^{-\alpha n})^{2\lambda}$$

which give $|A_i| \le |c_i(\omega)|/d(c_i(\omega)) \le Ce^{\sigma_0(n-i)}$. Therefore we can sum over all iterates to get

$$\sum_{i=0}^{n-1} |A_i| \le C.$$

This proves the lemma if there are no essential or inessential returns for ω before time n.

Now we consider the cases in which there is a non empty sequence of returns. We start by estimating the values $|A_{\mu_j}|$ for $j = 1, \dots, s$. Since c_{μ_j} is contained in the union of three intervals of the form $I^*_{r,l}$ we have $|c_{\mu_j}(\omega)| \le \widetilde{C}\varepsilon^\gamma e^{-r_j}/r_j^2 \le \widetilde{C}d(c_{\mu_j}(\omega))$ and therefore by Lemma 3.6,

$$\widetilde{D}(c_{\mu_j}(\omega)) \le \widetilde{C}/d(c_{\mu_j}(\omega)) \le \widetilde{C}\varepsilon^\gamma e^{r_j}.$$

Substituting in (20) we have

$$|A_{\mu_j}| \le \widetilde{C}/r_j^2.$$

We now consider A_i where i is not a return iterate. Notice first of all that any return μ_{j+1} to $\Delta^{\pm c}$ is *immediately preceded* by a return $\mu_j = \mu_{j+1} - 1$ to Δ^0. Therefore if $i \in (\mu_j, \mu_{j+1})$ we necessarily have that μ_{j+1} is a return to $\Delta^0_{r_s}$. Therefore we only need to distinguish two cases according as to whether μ_j is a return to Δ^0 or to $\Delta^{\pm c}$. For the moment we also restrict our attention to values of $j \le s-1$.

Suppose first that μ_j is a return to Δ^0. We distinguish two subcases: either $|\varphi'(c_i)| \ge e^\beta$ for all $a \in \omega$ or there is some $a \in \omega$ for which $|\varphi'(c_i(a))| < e^\beta$. Then Lemma 3.2 and Lemma 2.3 give $|(\varphi^{\mu_{j+1}-i})'(c_i(a))| \ge e^{\beta(\mu_{j+1}-i)}$ in the first subcase and $|(\varphi^{\mu_{j+1}-i})'(c_i(a))| \ge \varepsilon^{\gamma+\delta}e^{\beta(\mu_{j+1}-i-1)}$ in the second subcase. Moreover applying Lemma 3.2 this gives

$$|c_i(\omega)| \le \eta^2 e^{-\beta(\mu_{j+1}-i)}|c_{\mu_{j+1}}(\omega)| \le \eta^2 e^{-\beta(\mu_{j+1}-i)}\varepsilon^{2(\gamma+\delta)+\iota}/r_{j+1}^2$$

and

$$|c_i(\omega)| \leq \eta^2 \varepsilon^{-\gamma+\delta} e^{-\beta(\mu_{j+1}-i-1)} |c_{\mu_{j+1}}(\omega)| \leq \eta^2 e^{-\beta(\mu_{j+1}-i-1)} \varepsilon^{\gamma+\delta+\iota}/r_{j+1}^2$$

respectively, using the fact that $|c_{\mu_{j+1}(\omega)}| \leq \varepsilon^\gamma e^{-r_{j+1}}/r_{j+1}^2$ and that $r_{j+1} > r_s$ since, as we mentioned above, μ_{j+1} is necessarily a return to $\Delta^0_{r_s}$. Moreover we have that, in the first case, since i is not a return iterate, $d(c_i(\omega)) \geq \varepsilon^{2(\gamma+\delta)+\iota}$. In the second case we know (from the fact that $c_i(\omega)$ is small and $|\varphi'(c_i(a))| < e^\beta$ for some $a \in \omega$) that $c_i(\omega)$ is relatively far from Δ^0 and relatively close to $\pm c$ and therefore since i is not a return iterate, $d(c_i(\omega)) \geq \varepsilon^{\gamma+\delta}$. In both cases we have that $c_i(\omega)$ is an admissible interval and applying Lemma 3.6 and substituting in (20) we get $|A_i| \leq Ce^{-\beta(\mu_{j+1}-i)}/r_{j+1}^2$ in each case. Moreover we can sum over all $i \in (\mu_j, \mu_{j+1})$ to get

$$\sum_{i=\mu_j+1}^{\mu_{j+1}-1} |A_i| \leq C/r_{j+1}^2.$$

Now suppose that μ_j is a return to $\Delta^{\pm c}$. Then there follows a binding period $(\mu_j, \mu_j + p_j]$ and a (possibly empty) free period $(\mu_j + p_j + 1, \mu_{j+1})$. For iterates $i \in (\mu_j + p_j + 1, \mu_{j+1})$ the situation is exactly as in the case considered above and we have $|A_i| \leq Ce^{-\beta(\mu_{j+1}-i)}/r_{j+1}^2$ and $\sum_{i=\mu_j+p_j+1}^{\mu_{j+1}-1} |A_i| \leq C/r_{j+1}^2$. So it just remains to consider *bound* iterates $i \in (\mu_j, \mu_j + p_j]$.

First of all recall that $d(c_{\mu_j}(\omega)) \approx \varepsilon^\gamma e^{-r_j}$ and in particular, for all $a \in \omega$,

$$|c_1(a) - c_{\mu_j+1}(a)| \leq \widetilde{C}(\varepsilon^\gamma e^{-r_j})^2.$$

By the mean value theorem we have $|c_i(\omega)| = |c_{\mu_{j+1}}(\omega)| \cdot |(\varphi^{i-\mu_j-1})'(\zeta)|$ for some $\zeta \in c_{\mu_j+1}(\omega)$. Then using the bounded distortion estimate in Lemma 2.4(1) and the fact that $|c_i(\omega)| \leq \varepsilon^\gamma e^{\beta(\mu_j-i)}$ by the definition of binding period we get

$$|(\varphi^{i-(\mu_j-1)})'(\zeta)| \leq \frac{\varepsilon^\gamma e^{-\beta(\mu_j-1)}}{|c_{\mu_{j+1}}(\omega)|} \leq \frac{\varepsilon^\gamma e^{-\beta(\mu_j-1)}}{(\varepsilon^\gamma e^{-r_j})^2}.$$

Since $|\varphi'(x)| \approx \varepsilon^\gamma e^{-r_j}$ for all $x \in c_{\mu_j}(\omega)$ we then have

$$|(\varphi^{i-\mu_j})'(x)| \leq \varepsilon^\gamma e^{-\beta(\mu_j-1)}/\varepsilon^\gamma e^{-r_j}.$$

By Lemma 3.2 this gives

$$|c_i(\omega)| \leq \frac{\varepsilon^\gamma e^{-\beta(\mu_j-1)}}{\varepsilon^\gamma e^{-r_j}} |c_{\mu_j}(\omega)| \leq \varepsilon^\gamma e^{-\beta(\mu_j-1)}/r_j^2.$$

Moreover $d(c_i(\omega)) \geq \varepsilon^\gamma e^{-\alpha(i-\mu_j)} - e^{-\beta(i-\mu_j)} \geq C\varepsilon^\gamma e^{-\alpha(i-\mu_j)}$ and therefore, substituting in (20) we get $|A_i| \leq Ce^{(\alpha-\beta)(i-\mu_j)}/r_j^2$ which also yields

$$\sum_{\mu_j+1}^{\mu_j+p_j} |A_i| \leq C/r_j^2.$$

Finally we consider the last piece of orbit $(\mu_s + p_s, \mu_{s+1} - 1]$. This interval can be empty if the return μ_{s+1} occurs immediately after the end of the binding period, *i.e.* if $\mu_{s+1} = \mu_s + p_s + 1$, or if μ_{s+1} is a return to $\Delta^{\pm c}$. Indeed, in the latter case we also have $\mu_{s+1} = \mu_s + p_s + 1$ keeping in mind that μ_s is necessarily a return to Δ^0 and that therefore $p_s = 0$ by definition. So suppose that $i \in (\mu_s + p_s, \mu_{s+1} - 1]$. By the comments above this implies that μ_{s+1} is a return to Δ^0. Moreover we are assuming, by the hypotheses in the lemma, that $c_{\mu_{s+1}} \subset \Delta^0_{2(\gamma+\delta)}$. Suppose first that $|\varphi'(c_i(a))| < e^\beta$ for some $a \in \omega$. Then, repeating the exact same arguments used above we have

$$|c_i(\omega)| \leq C\varepsilon^{-(\gamma+\delta)} e^{-\beta(n-i-1)} |\Delta^0_{2(\gamma+\delta)}| \leq Ce^{-\beta(n-i-1)} \varepsilon^{\gamma+\delta}.$$

Moreover, from Lemma 3.6 we have $\widetilde{D}(c_i(\omega)) \leq C\varepsilon^{-(\gamma+\delta)}$ and so

$$|A_i| \leq |c_i(\omega)| \cdot \widetilde{D}(c_i(\omega)) \leq Ce^{-\beta(n-i-1)}.$$

Now suppose that $|\varphi'(c_i)| \geq e^\beta$. We distinguish two further subcases according as to whether $d(c_i(\omega)) \geq \varepsilon^{2(\gamma+\delta)}$ or not. Suppose first that $d(c_i(\omega)) \geq \varepsilon^{2(\gamma+\delta)}$, *i.e.* $c_i(\omega) \cap \Delta^0_{2(\gamma+\delta)} = \varnothing$. Then we have $|c_i(\omega)| \leq e^{-\beta(n-i)} \varepsilon^{2(\gamma+\delta)}$ and applying Lemma 3.6,

$$|A_i| \leq |c_i(\omega)| \cdot \widetilde{D}(c_i(\omega)) \leq Ce^{-\beta(n-i)}.$$

If $d(c_i(\omega)) < \varepsilon^{2(\gamma+\delta)}$ then we still have $|c_i(\omega)| \leq e^{-\beta(n-i)} \varepsilon^{2(\gamma+\delta)}$ and, applying Lemma 3.6

$$\widetilde{D}(c_i(\omega)) \leq C/d(c_i(\omega)) \leq C\varepsilon^{-2(\gamma+\delta)-\iota}.$$

Notice however that since $c_i(\omega)$ is small ($\leq \varepsilon^{2(\gamma+\delta)}$) and $d(c_i(\omega)) < \varepsilon^{(\gamma+\delta)}$, $c_i(\omega)$ is completely contained in a small neighbourhood of the origin, say $c_i(\omega) \subset \Delta^0$ where the derivative is very large. In particular we have from Lemma 2.2 that $|(\varphi^2)'(c_i(a))| \geq \varepsilon^{(\gamma+\delta)(1-2\lambda)}$ for any $a \in \omega$. Therefore arguing as above we can obtain a much stronger bound on the size of $c_i(\omega)$, more precisely,

$$|c_i(\omega)| \leq e^{-\beta(n-i+2)} \varepsilon^{(\gamma+\delta)(1-2\lambda)} |\Delta_{2(\gamma+\delta)}| \leq Ce^{-\beta(n-i+2)} \varepsilon^{(\gamma+\delta)(1-2\lambda)} \varepsilon^{2(\gamma+\delta)}.$$

and therefore substituting in (20) we get

$$|A_i| \leq C\varepsilon^\iota e^{-\beta(n-i)}.$$

Finally, let $R(q)$ be the set of indices j for which $|r_j| = q$ and when $R(q)$ is nonempty we denote by j(q) the largest of its elements. Notice that for all $j \in R(q)$ we have $c_{\mu_j}(\omega) \leq Ce^{-\beta(\mu_{j(q)} - \mu_j)} |c_{\mu_{j(q)}}|$ and therefore we have

$$\sum_1^{n-1} |A_i| = \sum_1^{\mu_s+p_s} |A_i| + \sum_{\mu_s+p_s+1}^{n-1} |A_i| \leq C \sum_{q:R(q)\neq\varnothing} q^{-2} + C \leq A.$$

This completes the proof of the lemma. $\square$

Lemma 3.8. — *There exists a constant $B > 1$ (independent of ε or n) such that if $\omega \in \mathcal{P}_{n-1}$ with $\omega \subset E_{n-1}$ and n is a return to $\Delta^{\pm c}$ then*

$$\left|\frac{c_k'(\overline{a})}{c_k'(a)}\right| \leq B \quad \textit{for all } \overline{a}, a \in \omega \textit{ and all } 0 \leq k \leq n-1.$$

If n is a return to Δ^0 we have the same result for any $\overline{a}, a$ belonging to a subinterval $\overline{\omega} \subset \omega$ with $|c_n(\overline{\omega})| \leq \max\{(\varepsilon^\gamma e^{-\alpha n})^{2\lambda}, \varepsilon^{2(\gamma+\delta)}\}$.

Proof. — This is a direct consequence of Lemmas 3.1 and 3.7, just take $B = A\eta$. □

4. Parameter exclusions

We now wish to estimate the total measure of the parameters excluded during every step of the induction. We shall treat separately the exclusions due to each one of the two conditions on the parameter. We shall start by showing that for some positive constants δ_1, α_1 we have

$$|E_{n-1} \setminus F_n| \leq \varepsilon^{\delta_1} e^{-\alpha_1 n} |E_{n-1}| \leq \varepsilon^{\delta_1} e^{-\alpha_1 n} |E_0|$$

i.e. the proportion of parameters excluded by CP1 at each iteration is exponentially small as $n \to \infty$ and as $\varepsilon \to 0$. Recall that there are no binding periods and therefore no exclusions due to CP2 for iterates $n \leq N = [r_s/\alpha]$ and that N can be made arbitrarily large by taking ε small. Thus we have $E_n = F_n$ for all $n \leq N$ and so

$$|E_{n-1} \setminus E_n| \leq \varepsilon^{\delta_1} e^{-\alpha_1 n} |E_0| \quad \forall n \leq N \tag{22}$$

and, inductively,

$$|E_n| \geq |E_{n-1}| - \varepsilon^{\delta_1} e^{-\alpha_1 n} |E_0| \geq |E_0| (1 - \sum_{i=0}^{n} \varepsilon^{\delta_1} e^{-\alpha_1 i}) \quad \forall n \leq N.$$

For general n we shall show below that $|F_n \setminus E_n| \leq e^{-\alpha_1 n} |E_0|$ and therefore we get from (22), $|E_{n-1} \setminus E_n| \leq 2e^{-\alpha_1 n} |E_0|$. This then gives, for $n > N$,

$$|E_n| \geq |E_N| - |E_0| \sum_{i=N+1}^{n} 2e^{-\alpha_1 i} \geq |E_0| \left(1 - \sum_{i=0}^{N} \varepsilon^{\delta_1} e^{-\alpha_1 i} - \sum_{i=N+1}^{n} 2e^{-\alpha_1 i}\right)$$

and

$$|E_\infty| \geq |E_0| \left(1 - \sum_{i=0}^{N} \varepsilon^{\delta_1} e^{-\alpha_1 i} - \sum_{i=N+1}^{\infty} 2e^{-\alpha_1 i}\right) \geq |E_0|(1 - C(\varepsilon))$$

where $C(\varepsilon) \to 0$ as $\varepsilon \to 0$.

4.1. Exclusions due to CP1. — Recall that for each n we need to throw away parameters a for which $c_n(a)$ falls into the $\varepsilon^\gamma e^{-\alpha n}$-neighbourhood of the origin and of the critical points, which we denoted $\Delta_{\alpha n}$. Moreover, if exclusion of these parameters leads to the formation of small connected components in parameter space (smallness being expressed in terms of their image under c_n) then such components are also excluded.

Given an interval $\omega \in \mathcal{P}_{n-1}$ with $\omega \subset E_{n-1}$, the subset $\widehat{\omega} \subset \omega$ of parameters which get thrown out at the n^{th} iteration satisfies $c_n(\widehat{\omega}) \subset \Delta_{[\alpha n]-1}$. The aim of this section is to show that the ratio $|\widehat{\omega}|/|\omega|$ is exponentially small for small ε and large n. In principle this is achieved by estimating the ratio $|c_n(\widehat{\omega})|/|c_n(\omega)|$ and then invoking the bounded distortion estimates in Lemma 3.8 to show that this ratio is essentially preserved when pulled back by c_n^{-1}. This works well for returns to $\Delta^{\pm c}$ as the bounded distortion estimates (*cf.* Lemma 3.8) can then be applied to the entire interval ω. For returns to Δ^0 we face the problem that the bounded distortion estimates only apply to intervals $\overline{\omega}$ which satisfy $|c_n(\overline{\omega})| \leq \max\{(\varepsilon^\gamma e^{-\alpha n})^{2\lambda}, \varepsilon^{2(\gamma+\delta)}\}$. Nevertheless the next lemma show that this is sufficient for our purposes.

Lemma 4.1. — *There exist constants $\delta_1 > 0$ and $\alpha_1 > 0$ (independent of n or ε) such that*

$$\frac{|E_{n-1}|}{|F_n|} \geq 1 - \varepsilon^{\delta_1} e^{-\alpha_1 n} \quad \textit{for any } \omega \in \mathcal{P}_{n-1} \textit{ with } \omega \subset E_{n-1}.$$

Proof. — Consider an element $\omega \in \mathcal{P}_{n-1}, \omega \subset E_{n-1}$. Let $\widehat{\omega} \subset \omega$ be those parameters which get excluded at time n for failing to satisfy CP1. Clearly it is enough to estimate $|\widehat{\omega}/|\omega|$ and we can suppose that n is a returning situation for ω otherwise the statement would be trivially true.

Suppose first that if n is not a return to $\Delta^{\pm c}$ then $|c_n(\omega)| < (\varepsilon^\gamma e^{-\alpha n})^{2\lambda}$, in particular the hypotheses of Lemma 3.8 are satisfied. Then we have

$$|\widehat{\omega}|/|\omega| \leq B|\Delta_{[\alpha n]-1}|/|c_n(\omega)| \leq 2e^2 B\varepsilon^\gamma e^{-\alpha n}/|c_n(\omega)|.$$

The estimates for $c_n(\omega)$ have all been obtained in Lemma 3.3 and we just need to consider the various cases. If n is a return to $\Delta^{\pm c}$ we have either $|c_n(\omega)| \geq \varepsilon^{1+\iota}$ from (1a) or $|c_n(\omega)| \geq \varepsilon^{1-1/\lambda}\varepsilon^{\gamma-\beta/\sigma}e^{-(1-\beta/\sigma)r}$ from (1c) and (3b). Using the fact that $r \leq \alpha n$ we clearly get the desired estimate in this case. If n is a return to Δ^0 then (1b) and (3a) give $|c_n(\omega)| \geq \varepsilon^{\gamma-\beta/\sigma}e^{-(1-\beta/\sigma)r}$ which again yields the statement in the lemma since $r \leq \alpha n$. Finally case (2) gives $|c_n(\omega)| \geq \varepsilon^{-\iota/2}\varepsilon^\gamma e^{-r_s}$. Notice that this case can only occur after a return to $\Delta^{\pm c}$ and such returns can only occur for large values of n, more precisely for $n \geq r_c/\alpha \gg r_s/\alpha$. Therefore we have

$$\varepsilon^\gamma e^{-\alpha n}/\varepsilon^{-\iota/2}\varepsilon^\gamma e^{-r_s} \leq \varepsilon^{\iota/2}e^{-\alpha n+r_s} \leq \varepsilon^{\iota/2}e^{-n(\alpha-r_s/n)} \leq \varepsilon^{\iota/2}e^{-\alpha' n}$$

which proves the result in this case also.

Finally suppose that $|c_n(\omega)| \geq (\varepsilon^\gamma e^{-\alpha n})^{2\lambda}$. Let $\overline{\omega} \subset \omega$ be such that $|c_n(\overline{\omega})| = (\varepsilon^\gamma e^{-\alpha n})^{2\lambda}$, in particular, the hypotheses of Lemma 3.8 are satisfied by $\overline{\omega}$ and we have

$$\frac{|\widehat{\omega}|}{|\omega|} \leq \frac{|\widehat{\omega}|}{|\overline{\omega}|} \leq B\frac{|\Delta_{[\alpha n]-1}|}{|c_n(\overline{\omega})|} \leq 2e\beta\frac{\varepsilon^\gamma e^{-\alpha n}}{(\varepsilon^\gamma e^{-\alpha n})^{2\lambda}} \leq 2eB(\varepsilon^\gamma e^{-\alpha n})^{1-2\lambda}.$$

This completes the proof of the lemma. □

4.2. Exclusions due to CP2. — In this section we consider elements $\omega \subset F_n, \omega \in \mathcal{P}_n$ which satisfy CP1(n). We will set up a statistical argument to show that most of these elements (in measure theoretical terms) have spent a small proportion of their time in binding periods and therefore also satisfy CP2(n).

Recall from Section 3.2 that a sequence of escape times $0 = \nu_0 < \cdots < \nu_{s+1} = n$ is associated to each element $\omega \in \mathcal{P}_n$. Here we set $\nu_0 = 0$ and $\nu_{s+1} = n$ for notational convenience and we call these escape times as well. Between any two escape times we have a (possibly empty) sequence of essential and inessential returns $\mu_0 < \cdots < \mu_q$ with $\mu_i = \mu_{i,j}$ and $q = q(j)$. To each such return μ_i is associated a positive integer $p_i \geq 0$, the length of the associated binding period, and an integer $r_i > 0$ determined by the associated host interval. We let $\widetilde{P}_j = p_{0,j} + \cdots + p_{q,j}$, $R_j = r_{0,j} + \cdots + r_{q,j}$, $\widetilde{P}_+ = \widetilde{P}_1 + \cdots + \widetilde{P}_s$ and $R_+ = R_1 + \ldots R_s$. In particular $\widetilde{P}_+ = P_n$, the total number of iterates before time n belonging to binding periods. From Lemma 2.4 we immediately get the following

Lemma 4.2. — *For $\omega \in \mathcal{P}_n, \omega \subset F_n$ we have*

$$P_n \leq \frac{2(R_+ + 3/2\log 1/\varepsilon)}{\sigma} \leq \frac{2(1+\gamma/\delta)R_+}{\sigma}.$$

In particular we can formulate an alternative condition

CP2' : $R_+ \leq \sigma n/4(1+\gamma/\delta)$

which immediately implies CP2: $P_n \leq n/2$.

Lemma 4.3. — *Let $\omega_{\nu_j} \in \mathcal{P}_{\nu_j}$ and $\omega_{\nu_{j+1}} \in \mathcal{P}_{\nu_j}$ be escaping components with $\omega_{\nu_j} \subset \omega_{\nu_{j+1}}$ and suppose that there exists a non empty sequence $\mu_0 < \cdots < \mu_q$ of* essential *returns between time ν_j and time ν_{j+1}. Then*

$$|\omega_{\nu_{j+1}}| \leq \varepsilon^\iota e^{-\iota R}|\omega_{\nu_j}|.$$

Proof. — Let $\omega_{\mu_i} \in \mathcal{P}_{\mu_i}$ be the subintervals of ω_{μ_j} corresponding to the returns μ_i, $i = 0, \ldots, q$. We write

$$\frac{|\omega_{\nu_{j+1}}|}{|\omega_{\nu_j}|} = \frac{|\omega_{\mu_0}|}{|\omega_{\nu_j}|}\frac{|\omega_{\mu_1}|}{|\omega_{\mu_0}|}\cdots\frac{|\omega_{\mu_q}|}{|\omega_{\mu_{q-1}}|}\frac{|\omega_{\nu_{j+1}}|}{|\omega_{\mu_q}|} \tag{23}$$

and begin by estimating $|\omega_{\mu_0}|/|\omega_i|$. We have $|c_{\mu_0}(\omega_i)| \geq \varepsilon^{-\iota/2}\varepsilon^\gamma e^{-r_s}$, by Lemma 3.5. By the definition of the components ω_{μ_i}, notice that the first return after an escape time is always a return to Δ^0, we have $|c_{\mu_o}(\omega_{\mu_0})| \leq 10\varepsilon^\gamma e^{r_0}/r_0^2$ for some $r_0 \geq r_s$. We distinguish two cases. Suppose first that $c_{\mu_0}(\omega_i) \subset \Delta_{2(\gamma+\delta)}$. Then we can apply the bounded distortion estimates and we have

$$\frac{|\omega_{\mu_0}|}{\omega_{\nu_j}} \leq \frac{B|c_{\mu_0}(\omega_{\mu_0})|}{|c_{\mu_0}(\omega_{\nu_j})|} \leq 10B\varepsilon^{-\iota/2}e^{r_s - r_0}/r_0^2. \tag{24}$$

If $c_{\mu_0}(\omega_i)$ is not completely contained in $\Delta_{2(\gamma+\delta)}$ then, using the fact that $\varepsilon^{\gamma}e^{-r_s} \approx \varepsilon^{2(\gamma+\delta)}$ we have $|c_{\mu_0}(\overline{\omega}_{\nu_j})| \geq \varepsilon^{2(\gamma+\delta)} - \varepsilon^{\gamma}e^{-r_s} \geq \varepsilon^{2(\gamma+\delta)}(1-\varepsilon^{\iota})$ where $\overline{\omega} \subset \omega_i$ is that subinterval of ω_{ν_j} whose image is completely contained in $\Delta_{2(\gamma+\delta)}$. Then

$$\begin{aligned}\frac{|\omega_{\mu_0}|}{|\omega_{\nu_j}|} \leq \frac{|\omega_{\mu_0}|}{|\overline{\omega}_{\nu_j}|} \leq \frac{B|c_{\mu_0}(\omega_{\mu_0})|}{|c_{\mu_0}(\overline{\omega}_i)|} &\leq \frac{10B\varepsilon^{\gamma}e^{-r_0}}{r_0^2\varepsilon^{2(\gamma+\delta)}(1-\varepsilon^{\iota})} \\ &\leq 10B\varepsilon^{-(\gamma+2\delta)}e^{-r_0}/r_0^2 \ll 10B\varepsilon^{-\iota/2}e^{r_s-r_0}.\end{aligned}$$

We now turn to estimating the ratios $|\omega_{\mu_j}|/|\omega_{\mu_{j-1}}|$ for $j = 1, \ldots, q$. Each time we need to distinguish as above the cases in which $c_{\mu_{j-1}}(\omega_{\mu_{j-1}})$ is completely contained in $\Delta_{2(\gamma+\delta)}$ so that we can apply the bounded distortion estimates, and the cases in which this is not true. However, repeating the argument above we see that the estimates which we obtain in the former cases are always satisfied in the latter. Thus we shall consider in detail only the situation in which $c_{\mu_{j-1}}(\omega_{\mu_{j-1}}) \subset \Delta_{2(\gamma+\delta)}$. Suppose that μ_{j-1} is a return to Δ^0. Then we have

$$\frac{|\omega_{\mu_j}|}{|\omega_{\mu_{j-1}}|} \leq \frac{10Br_{j-1}^2\varepsilon^{\gamma}e^{-r_j}}{r_j^2\varepsilon^{2\lambda\gamma}e^{-2\lambda r_{j-1}}} \leq 10Br_{j-1}^2\varepsilon^{(1-2\lambda)\gamma}e^{2\lambda r_{j-1}-r_j} \tag{25}$$

If $\mu_{j-1}$is a return to $\Delta^{\pm c}$ then we have

$$\frac{|\omega_{\mu_j}|}{|\omega_{\mu_{j-1}}|} \leq \frac{10Br_{j-1}^2\varepsilon^{\gamma}e^{-r_j}}{r_j^2\varepsilon^{2\gamma+\delta+2\beta\gamma/\sigma}e^{-2\beta\gamma/\sigma r_{j-1}}} \leq 10Br_{j-1}^2\varepsilon^{-(\gamma+\delta+2\beta\gamma/\sigma)}e^{2\beta r_{j-1}/\sigma - r_j}. \tag{26}$$

Recall moreover that if μ_j is a return to $\Delta^{\pm c}$ we gain an extra factor of $\varepsilon^{1-1/\lambda}$ on the right hand side of (25) and (26). Finally we have $|\omega_{i+1}|/|\omega_{\mu_q}| \leq 1$.

Now let $\widetilde{q}$ denote the number of returns between μ_0 and μ_{q-1} (inclusive) which occur in Δ^0 and $\widehat{q}$ denote the number of those which occur in $\Delta^{\pm c}$. We do not include μ_q in this count and therefore we have $q = \widetilde{q} + \widehat{q} + 1$. Let $\widetilde{R} = \sum \widetilde{r}_i$ and $\widehat{R} = \sum \widehat{r}_i$ and $R = \sum_{i=0}^{q} r_i = \widetilde{R} + \widehat{R} + r_q$. Then we have from (23)(24)(25)(26)

$$\frac{|\omega_{i+1}|}{|\omega_i|} \leq (10B)^{q+1}\varepsilon^{\iota/2}e^{r_s}\varepsilon^{(1-2\lambda)\gamma\widetilde{q}}e^{2\lambda\widetilde{R}}\varepsilon^{(1/\lambda-1)\widehat{q}}\varepsilon^{-(\gamma+\delta+2\beta\gamma/\sigma)\widehat{q}}e^{2\beta\widehat{R}/\sigma}e^{-R}. \tag{27}$$

To simplify this expression we make the following three observations:

(i) $e^{2\lambda\widetilde{R}} = e^{(\lambda+1/2)\widetilde{R}}e^{-(1/2-\lambda)\widetilde{R}}$;

(ii) $(1/\lambda - 1) > \gamma + \delta + 2\beta\gamma/\sigma$ and therefore

$$\varepsilon^{(1/\lambda-1-\gamma-\delta-2\beta\gamma/\sigma)\widehat{q}}e^{2\beta\widehat{R}/\sigma} \leq e^{2\beta\widehat{R}/\sigma} \leq e^{(\lambda+1/2)\widehat{R}};$$

(iii) μ_0 is a return to Δ^0 and therefore $\widetilde{q} \geq 1$ and $\widetilde{R} \geq r_s$.
From (27) and these observations we get

$$\frac{|\omega_{i+1}|}{|\omega_i|} \leq (10B)^{q+1}\varepsilon^{\iota/2}e^{r_s}\varepsilon^{(1-2\lambda)\gamma}e^{-(1/2-\lambda)r_s-(\lambda+1/2)r_q}e^{(\lambda+1/2)(\widetilde{R}+\widehat{R}+r_q)-R}.$$

Now using the fact that $r_s = [(\gamma + 2\delta + \iota)\log 1/\varepsilon$ and $r_q \geq r_\varepsilon = [(\gamma+\delta)\log 1/\varepsilon]$ we see that

$$\begin{aligned} e^{(\lambda-1/2)r_s-(\lambda+1/2)r_q} &\leq \varepsilon^{(\lambda-1/2)(\gamma+2\delta+\iota)+(\lambda+1/2)(\gamma+\delta)} \\ &\leq \varepsilon^{\lambda(2\gamma+3\delta+\iota)-(\delta+\iota)/2} \leq \varepsilon^{\iota/2}. \end{aligned}$$

Thus we have

$$\frac{|\omega_{i+1}|}{|\omega_i|} \leq (10B)^{q+1}\varepsilon^{\iota}e^{(\lambda+1/2-1)R}.$$

Finally, since $q \leq R/r_\varepsilon$ we have $(10B)^q \leq ((10B)^{1/r_\varepsilon})^R$ and so, taking $\varepsilon > 0$ small we get

$$\frac{|\omega_{i+1}|}{|\omega_i|} \leq \varepsilon^{\iota}e^{-\iota R}.$$

This concludes the proof. $\square$

Now let $\eta_q(R)$ denote the number of possible sequences $r_1, \ldots, r_q$ with $r_i \geq \delta \log \varepsilon^{-1}$ and $r_i + \cdots + r_q = R$ and let $\eta(R) = \sum_{q\geq 0} \eta_q(R)$.

Lemma 4.4. — *For $\varepsilon > 0$ sufficiently small we have that, for all $R \in \mathbb{N}$,*

$$\eta(R) \leq e^{\iota R/2}.$$

Proof. — The result is purely combinatorial. We want to estimate, for each fixed q,

$$\eta_q(R) \leq \binom{R}{q} \leq \frac{R!}{q!(R-q)!}.$$

Using Stirling's approximation formula for factorials:

$$\sqrt{2\pi k}k^k e^{-k} \leq k! \leq \sqrt{2\pi k}k^k e^{-k}(1 + 1/4k),$$

we get

$$\begin{aligned} \eta_q(R) &\leq \frac{\sqrt{2\pi R}R^R e^{-R}(1+1/4R)}{\sqrt{2\pi q}q^q e^{-q}\sqrt{2\pi(R-q)}(R-q)^{R-q}e^{-(R-q)}} \\ &\leq 2\frac{R^R}{q^q(R-q)^{R-q}} \qquad \text{for small } \varepsilon > 0 \text{ (and therefore } R \text{ large)} \\ &\leq 2\left(\frac{R}{q}\right)^q \left(\frac{R}{R-q}\right)^{R-q} \leq 2\left[\left(\frac{1}{q/R}\right)^{q/R}\left(\frac{1}{1-q/R}\right)^{1-q/R}\right]^R. \end{aligned}$$

Now since $q/R \leq 1/\delta \log 1/\varepsilon \to 0$ we have that $(1/(q/R))^{q/R}$ and $(1/(1-q/R))^{1-q/R}$ both tend to 1 as $\log\varepsilon \to 0$. Thus we get $\eta_q(R) \leq e^{(\iota R/2)/2}$. Notice that the value of

q is bounded by $q \leq R/(\delta \log 1/\varepsilon) \leq R$, for each R, and so, summing over all possible values of q we get $\eta(R) \leq \sum_{q \leq R} \eta_q(R) \leq e^{\iota R}$. □

Now let $\omega_{\nu_j} \in \mathcal{P}_{\nu_j}$ be an escaping component and let $\widetilde{\eta}(R)$ denote the total number of subintervals $\omega \subset \omega_{\nu_j}$ which are escaping components of the form $\omega_{\nu_{j+1}}$, $\nu_{j+1} = \nu_{j+i}(\omega)$ which undergo a sequence of returns $\mu_0, \ldots, \mu_q$ between time ν_j and time μ_{j+1} with $R_j = R$. Then we have

Lemma 4.5. — *For all $R \in \mathbb{N}$,*

$$\widetilde{\eta}(R) \leq e^{2\iota R}.$$

Proof. — Notice that the subintervals ω can be indexed in a unique way by the sequence of host intervals corresponding to the returns $\mu_0, \ldots, \mu_q$, *i.e.* by a sequence $(*_0, r_0, l_0) \ldots (*_q, r_q, l_q)$ where $* = 0, \pm c*$ and $1 \leq l_i \leq r_i^2$. It follows that for each r there exist at most $6r^2$ intervals with $r_i = r$ and therefore the previous lemma immediately gives $\widetilde{\eta}(R) \leq 6R^2 e^{\iota R} \leq e^{2\iota R}$. □

For the final step of our argument we introduce the following notation. For $\omega \in \mathcal{P}_n$ let ω_{ν_j}, $j = 0, \ldots, s$ be the escaping components containing ω as defined above. Then define for $j = s+1, \ldots, n$, $\omega_{\nu_j} = \omega$ and call these escaping components as well. Thus we have a formally defined sequence of nested intervals

$$\omega = \omega_n = \cdots = \omega_{\nu_{s+1}} \subset \omega_{\nu_s} \subset \cdots \subset \omega_{\nu_0} = \omega_0$$

associated to each $\omega \in \mathcal{P}_n$. For each such ω_{ν_j} let $\langle \omega_{\nu_j} \rangle = \omega_{\nu_j} \cap F_n$ and let Q_j denote the union of all the escaping components of the form ω_{ν_j}. Notice that $Q_0 = \omega_0$ and $Q_n = F_n$.

Lemma 4.6. — *For every $n \geq 1$, we have*

$$\int_{F_n} e^{\iota R/4} da \leq (1 + \varepsilon^{\iota})^n |\omega_0|.$$

Proof. — For a given ω_{ν_j} we have

$$\begin{aligned}\int_{<\omega_{\nu_j}>} e^{\iota R/4} da &= |\Omega_0| + \sum_{R \geq r_s} |\Omega(R)| e^{\iota R/4} \leq |\omega_{\nu_j}| + \sum_{R \geq r_s} \varepsilon^{\iota} e^{-\iota R/4} |\omega_i| \\ &\leq (1 + \varepsilon^{\iota} \sum_{R \geq r_s} e^{-\iota R/4}) |\omega_{\nu_j}| \leq (1 + \varepsilon^{\iota}) |\omega_{\nu_j}|.\end{aligned}$$

Clearly this implies

$$\int_{Q_i} e^{\iota R/4} da \leq (1 + \varepsilon^{\iota/2}) Q_i$$

and, inductively, the statement in the lemma. □

We are now in a position to estimate the proportion of parameters satisfying $CP2'$: $R_+ \leq \sigma n/4(1+\gamma/\delta)$. This condition is equivalent to $e^{\iota R/4} \leq e^{\iota\sigma n/16(1+\gamma/\delta)} \leq e^{\xi n}$ where $\xi = \iota\sigma/16(1+\gamma/\delta)$. Thus we have

$$m\{a \in F_n : e^{\iota R/4} \geq e^{\xi n}\} e^{\xi n} \leq \int_{F_n} e^{\iota R/4} da \leq (1+\varepsilon^{\iota/2})^n |\omega_0|$$

which gives

$$m\{E_n \setminus F_n\} \leq (1+\varepsilon^{\iota})^n e^{-\xi n}|\omega_0| \leq e^{-\xi n/2}|\omega_0|$$

taking ε small.

References

[1] V.S. Afraimovich, V.V. Bykov, and L.P. Šil'nikov, *On attracting structurally unstable limit sets of Lorenz type attractors*, Trans. Moscow Math. Soc. **44** (1982), 153–216.

[2] M. Benedicks and L. Carleson, *On iterations of* $1-ax^2$ *on* $(-1,1)$, Annals of Math. **122** (1985), 1–25.

[3] M. Benedicks and L. Carleson, *The dynamics of the Hénon map*, Annals of Math. **133** (1991), 73–169.

[4] P. Collet and J.-P. Eckmann, *On the abundance of aperiodic behaviour*, Comm. Math. Phys. **73** (1980), 115–160.

[5] J. Guckenheimer, *A strange, strange attractor*, The Hoproof bifurcation theorem and its applications (M. J.E. Marsden, McCracken, ed.), Springer-Verlag, 1976, 368–381.

[6] J. Guckenheimer and R.F. Williams, *Structural stability of Lorenz attractors*, Publ. Math. IHES **50** (1979), 307–320.

[7] M. Hénon, *A two dimensional mapping with a strange attractor*, Comm. Math. Phys. **50** (1976), 69–77.

[8] M. Hénon and Y. Pomeau, *Two strange attractors with a simple structure*, Lect. Notes in Math. **565** (1976), 29–68.

[9] M.V. Jakobson, *Absolutely continuous invariant measures for one-parameter families of one-dimensional maps*, Comm. in Math. Phys. (1981), 39–88.

[10] S. Luzzatto and M. Viana, *Lorenz-like attractors without invariant foliations*, in preparation.

[11] E. N. Lorenz, *Deterministic nonperiodic flow*, J. Atmosph. Sci. **20** (1963), 130–141.

[12] M. Misiurewicz, *Absolutely continuous invariant measures for certain interval maps*, Publ. Math. IHES **53** (1981), 17–51.

[13] J. Palis and F. Takens, *Hyperbolicity and sensitive-chaotic dynamics at homoclinic bifurcations*, Cambridge University Press, 1993.

[14] A. Rovella, *The dynamics of perturbations of the contracting Lorenz attractor*, Bull. Braz. Math. Soc. **24** (1993), 233–259.

[15] C. Sparrow, *The Lorenz equations bifurcations, chaos and strange attractors*, Applied Mathematical Sciences, vol. 41, Springer-Verlag, 1982.

[16] P. Thieullen, C. Tresser, and L.-S. Young, *Exposant de Lyapunov positif dans des familles à un paramètre d'applications unimodales*, C. R. Acad. Sci. Paris Sér. I Math. **315** (1992), 69–72.

S. Luzzatto, SISSA, Via Beirut 4, 34014 Trieste, Italy • *E-mail :* luzzatto@maths.warwick.ac.uk
Url : http://www.maths.warwick.ac.uk/~luzzatto

• *Current Address:* Mathematics Institute, Warwick University, Coventry CV4 7AL, UK

M. Viana, IMPA, Est. D. Castorina 110, 22460-320 Rio de Janeiro, Brazil • *E-mail :* viana@impa.br
Url : http://www.impa.br/~viana

Astérisque

MARCO MARTENS
TOMASZ NOWICKI

Invariant measures for typical quadratic maps

Astérisque, tome 261 (2000), p. 239-252
<http://www.numdam.org/item?id=AST_2000__261__239_0>

Astérisque
261, 2000, p. 239–252

INVARIANT MEASURES FOR TYPICAL QUADRATIC MAPS

by

Marco Martens & Tomasz Nowicki

Abstract. — A sufficient geometrical condition for the existence of absolutely continuous invariant probability measures for $S-$unimodal maps will be discussed. The Lebesgue typical existence of Sinai-Bowen-Ruelle-measures in the quadratic family will be a consequence.

1. Introduction

A general belief, or hope, in the theory of dynamical systems is that typical dynamical systems have well-understood behavior. This belief has two forms, depending on the meaning of the word "typical". It could refer to the topological generic situation or to the Lebesgue typical situation in parameter space. In this work *typical* will refer to Lebesgue typical and the behavior of a Lebesgue typical quadratic map on the interval will be discussed.

The quadratic family is formed by the maps $q_t : [-1,1] \to [-1,1]$ with $t \in [0,1]$ and

$$q_t(x) = -2tx^\alpha + 2t - 1,$$

with the critical exponent $\alpha = 2$. The maps in this family can be classified as follows. The maps in

$$\mathcal{P} = \{t \in [0,1] \mid q_t \text{ has a periodic attractor}\}$$

have a unique periodic orbit whose basin of attraction is an open and dense set. Moreover this basin has full Lebesgue measure. In particular the invariant measure on the periodic attractor is the SBR-measure for the map. Recall that a measure μ on $[-1,1]$ is called an S(inai)-B(owen)-R(uelle)-measure for q_t if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \delta_{q_t^k(x)} = \mu,$$

1991 ***Mathematics Subject Classification.*** — 58F03, 58F11.
Key words and phrases. — Unimodal maps, SRB-measures.

T. N. Partially supported by the Polish KBN Grant 2 P03A 02208.

for typical $x \in [-1,1]$.

The maps in

$$\mathcal{R} = \{t \in [0,1] \mid q_t \text{ is infinitely renormalizable}\}$$

have a unique invariant minimal Cantor set which attracts both generic and typical orbits. This Cantor set is uniquely ergodic and has zero Lebesgue measure, **[BL2]**, **[G]**, **[M1]**. The unique invariant measure on this Cantor set is the SBR-measure for the system. The maps in

$$\mathcal{I} = [0,1] \setminus \{\mathcal{P} \cup \mathcal{R}\}$$

have a periodic interval whose orbit is the limit set of generic orbits. The orbit of this periodic interval absorbs also the orbit of typical points. These maps are ergodic with respect to the Lebesgue measure, **[BL1]**, **[GJ]**, **[K]**, **[M1]**. In the quadratic family, $\alpha = 2$, the limit set of typical points is actually also the orbit of this periodic interval, **[L1]**. However, in families with α big enough there are maps in $\mathcal{I}$ whose typical limit set is not this periodic interval, **[BKNS]**.

Before discussing the behavior of typical quadratic maps let us include the behavior of generic quadratic maps .

***Theorem 1.1* ([GS], [L3]).** — *Hyperbolicity is dense in the quadratic family, e.g.* $\overline{\mathcal{P}} = [0,1]$.

We will continue to specify the behavior of a typical map in $\mathcal{I}$. The dynamics of maps in

$$\mathcal{M} = \{t \in \mathcal{I} \mid q_t \text{ has an absolutely continuous invariant probability measure}\}$$

is well-understood. The measure is unique and its support is the orbit of the above periodic interval. Moreover it has positive Lyapunov exponent, **[K]**, **[Ld]**. Starting in **[NU]**, where it was shown that $q_1 \in \mathcal{M}$, more and more maps q_t were shown to have such a measure (**[B]**, **[R]**, **[Mi]**). Finally it was shown in **[Ja]** that $\mathcal{M}$ has positive measure.

***Main Theorem 1.2* (joint with Lyubich).** — *A typical quadratic map has a unique SBR-measure. More specifically*

(1) *for $t \in \mathcal{P}$ the support of the SBR-measure is the periodic attractor,*
(2) *for $t \in \mathcal{R}$ the SBR-measure is supported on a Cantor set,*
(3) *for $t \in \mathcal{M}$ the SBR-measure is an absolutely continuous measure supported on the orbit of a periodic interval,*
(4) *the set $\mathcal{P} \cup \mathcal{R} \cup \mathcal{M} \subset [0,1]$ has full Lebesgue measure.*

The general belief that typical dyamical systems have well understood behavior has been precisely formulated in the Palis-Conjecture **[P]**. Now, by Theorem 1.2, this Conjecture has been proved for the quadratic family.

Johnson constructed unimodal maps in $\mathcal{I}$ (with arbitrary critical exponent) which do not have an absolutely continuous invariant measure, **[Jo]**. More careful combinatorial Johnson-Examples were made without SBR-measure **[HK]**. The same work shows the existence of maps in $\mathcal{I} \setminus \mathcal{M}$ which have an SBR-measure but this measure is not absolutely continuous. The complications which occur in $\mathcal{I} \setminus \mathcal{M}$ are thoroughly studied in **[Br]**.

In this work we will formulate a geometrical condition on maps in $\mathcal{I}$ sufficient for the existence of absolutely continuous invariant probability measures. The geometric condition is formulated in terms of a decreasing sequence of *central intervals* $U_n = (-u_n, u_n)$, $n \geq 1$, which are defined for all unimodal maps with recurrent critical orbit. The domain $D_n \subset U_n$ of the first return map $R_n : D_n \to U_n$, $n \geq 1$ is a countable collection of intervals. The central component of D_n is U_{n+1}. The first return map R_n is said to have a *central return* when

$$R_n(0) \in U_{n+1}.$$

The sufficient geometrical condition for the existence of absolutely continuous measures is stated in terms of *scaling factors*

$$\sigma_n = \frac{u_{n+1}}{u_n}, \quad n \geq 1.$$

These scaling factors describe the small scale geometrical properties of the system but they are also strongly related to distortion questions. The main consequence of the distortion Theory developed in **[M1]** are the a priori bounds on the distortion of each R_n. The renormalization Theory developed in **[L1]** and **[LM]** achieved much stronger results: if a quadratic unimodal map has only finitely many central returns then the scaling factors tend exponentially to zero.

The scaling factors are related to small scale geometry, distortion but also expansion. The technical step in this work is to show that small scaling factors imply strong expansion along the critical orbit. In **[NS]** it was shown that enough expansion along the critical orbit causes the existence of an absolutely continuous invariant probability measure. In particular, if

$$\sum_{n \geq 1} \left| Dq_t^n(q_t(0)) \right|^{-1/\alpha} < \infty$$

then q_t has an absolutely continuous invariant probability measure.

Main Theorem 1.3. — *Let f be an $S-$unimodal map with critical exponent $\alpha > 1$. If f has summable scaling factors, that is*

$$\sum_{n \geq 1} \sigma_n^{1/\alpha} < \infty,$$

then it has an absolutely continuous invariant probability measure.

Corollary 1.4. — *If a quadratic map has only finitely many central returns then it has an absolutely continuous invariant probability measure.*

The (Johnson-)Examples in **[Jo]** have infinitely many (cascades of) central returns. The corollary states that the only quadratic unimodal maps in $\mathcal{I}$ which do not have an absolutely continuous invariant measure are Johnson-Examples. The families $\{q_t\}$ with α big enough have maps in $\mathcal{I}$ which do not have an absolutely continuous invariant probability measure and which are also not Johnson-Examples, **[BKNS]**.

In **[L2]**, Lyubich studies the parameter space of the (holomorphic) quadratic family. A new proof showing that $\mathcal{I}$ has positive Lebesgue measure is given (compare with the Jacobson-Theorem **[Ja]**). Moreover it is shown that for almost every parameter in $\mathcal{I}$ the corresponding quadratic map has only finitely many central returns. This, together with Theorem 1.3, implies Theorem 1.2.

Conjecture 1.5. — *A typical map in the family $\{q_t\}$, with critical exponent $\alpha > 1$, has a unique SBR-measure. More specifically*

(1) *for $t \in \mathcal{P}$ the support of the SBR-measure is the periodic attractor,*
(2) *for $t \in \mathcal{R}$ the SBR-measure is supported on a Cantor set,*
(3) *for $t \in \mathcal{M}$ the SBR-measure is an absolutely continuous measure supported on the orbit of a periodic interval,*
(4) *the set $\mathcal{P} \cup \mathcal{R} \cup \mathcal{M} \subset [0,1]$ has full Lebesgue measure.*

An appendix is added to collect the standard notions and Lemmas in interval dynamics.

Acknowledgements. — This work has been done during the second authors visit to SUNY at Stony Brook. The second author would like to thank SUNY at Stony Brook for its kind hospitality.

2. Central Intervals

Throughout the following sections we will fix an S−unimodal map $f : [-1,1] \to [-1,1]$ with critical exponent $\alpha > 1$ and without periodic attractors. Furthermore assume that the critical orbit is recurrent.

The set of nice points is

$$\mathcal{N} = \{x \in [-1,1] \mid \forall i \geq 0 \; f^i(x) \notin (-|x|, |x|)\}$$

This set is closed and not empty. For example the fixed point of f in $(0,1)$ is in $\mathcal{N}$.

For $x \in \mathcal{N}$ let $D_x \subset U_x = (-|x|, |x|)$ be the set of points whose orbit returns to U_x. The first return map to U_x is denoted by

$$R_x : D_x \longrightarrow U_x.$$

The next Lemma is a straightforward consequence of the fact that the boundary of each U_x is formed by nice points.

Lemma 2.1 ([**M1**]). — *For every $x \in \mathcal{N}$ there exists a collection of pairwise disjoint intervals $\mathcal{U}_x$ with*

(1) $I \subset U_x$ *for all* $I \in \mathcal{U}_x$,
(2) $\bigcup_{I \in \mathcal{U}_x} I = D_x$,
(3) *if* $I \in \mathcal{U}_x$ *and* $0 \notin I$ *then* $R_x : I \to U_x$ *is monotone and onto,*
(4) *if* $I \in \mathcal{U}_x$ *and* $0 \in I$ *then* $R_x : I \to U_x$ *is 2 to 1 onto the image. Moreover* $R_x(\partial I) = \{x\}$ *or* $\{-x\}$.

Define the function $\psi : \mathcal{N} \to \mathcal{N}$ by

$$\psi(x) = \partial V_x \cap (0, 1),$$

where $V_x \in \mathcal{U}_x$ is the *central* interval: $0 \in V_x$. Say $R_x|_{V_x} = f^{q_x}$ and observe that

$$\{f(\psi(x)), f^2(\psi(x)), \ldots, f^{q_x}(\psi(x)) = x\} \cap U_x = \varnothing$$

which follows from the fact that $R_x : D_x \to U_x$ is the first return map. In particular $\psi(x) \in \mathcal{N}$ and we can consider the first return map to V_x. It will also satisfy Lemma 2.1.

Choose $u_1 \in \mathcal{N}$ and consider the sequence $u_n = \psi(u_{n-1})$ with $n \geq 1$. Use the simplified notation $\mathcal{U}_n$ for $\mathcal{U}_{u_n}$ and denote the first return maps by

$$R_n : D_n \longrightarrow U_n$$

instead of $R_{u_n} : D_{u_n} \to U_{u_n}$. All these first return maps have the properties stated in Lemma 2.1. Observe that $|U_n| = 2u_n$.

Let $\sigma_n = u_{n+1}/u_n$, $n \geq 1$. We call σ_n the *scaling factor* of level n. We will assume that

$$\sigma_n \longrightarrow 0.$$

However, the main Proposition 3.1, can also be proved by using only an a priori bound on the scaling factors. The assumption $\sigma_n \to 0$ will make the exposition less cumbersome.

Lemma 2.2. — *If* $I \in \mathcal{U}_n$ *and* $R_n|_I = f^t$ *then there exists an interval* $J \supset f(I)$ *such that*

$$f^{t-1} : J \longrightarrow U_{n-1}$$

is monotone onto. In particular all the maps $f^{t-1} : f(I) \to U_n$, $I \in \mathcal{U}_n$ *have uniformly bounded distortion. Moreover these maps will be essentially linear when* $n \to \infty$.

Proof. — Let $I \in \mathcal{U}_n$ with $R_n|I = f^t$ and let $J \supset f(I)$ be the maximal interval on which f^{t-1} is monotone. The maximality implies the existence of $i < t-1$ such that $0 \in \partial f^i(J)$. Observe that $f^i(f(I)) \cap U_n = \varnothing$, the first return happens after $t - 1 > i$ steps. So u_n (or $-u_n$) $\in f^i(J)$. We observed before that the orbit of $f(u_n)$ never enters U_{n-1}, $f^{t-1}(J) \supset U_{n-1}$. □

Lemma 2.3. — *For $\varepsilon > 0$ there exists $n_0 \geq 1$ such that the hyperbolic length of any $I \in \mathcal{U}_n$ is small,*

$$\mathrm{hyp}(I, U_n) \leq \varepsilon \textit{ and also } \frac{|f(I)|}{|f(U_n)|} \leq \varepsilon,$$

whenever $n \geq n_0$.

Proof. — Let $I \in \mathcal{U}_n$, say $R_n|I = f^t|I$. The previous Lemma states the existence of an interval $J \supset f(I)$ such that $f^{t-1} : (J, f(I)) \to (U_{n-1}, U_n)$ is monotone onto. For n large we see that U_n is a very small middle interval in U_{n-1}, it has a very small hyperbolic length. Because the map f has negative Schwarzian derivative we get that $f(I) \subset J$ has a very small hyperbolic length. Observe that $f^{-1}(J) \subset U_n$. Otherwise the orbit of u_n would pass through U_{n-1}. This is impossible: we saw before that the orbit of $\psi(x) = u_n$ does not cross $U_x = U_{n-1}$. The Lemma will be proved by pulling back the pair $(J, f(I))$ one step more. $\square$

3. Derivatives along Recurrent Orbits

Let $\rho_n = \min\{1/\sigma_{n-1}, 1/\sigma_n\}$. In this section we will prove

Proposition 3.1. — *There exist $n_0 \geq 1$, $\theta < 1$ and $C > 0$ with the following property. If $n \geq n_0$, $x \in U_{n+1}$ and $R_n^i(x) \notin U_{n+1}$ for $i \leq s$ then*

$$|Df^T(f(x))| \geq C \cdot \rho_n \cdot \theta^{-(s-1)},$$

where $R_n^s(x) = f^T(x)$.

In **[VT]** a similar estimate in the case $s = 1$ was obtained for circle homeomorphisms with irrational rotation number of bounded type. The proof of Proposition 3.1 will use the following Lemmas and notation. Fix $x \in U_{n+1}$ according to the Proposition, say $R_n^i(x) = f^{t_i}(x)$ with $i \leq s$.

Lemma 3.2. — *For each $i \leq s$ there exists an interval $S_i \ni f(x)$ such that*

$$f^{t_i - 1} : S_i \longrightarrow U_n$$

is monotone and onto.

Proof. — Lemma 2.2 applied to $U_{n+1} \in \mathcal{U}_n$ states the existence of S_1. The proof will proceed by induction. Assume that $S_i \ni f(x)$ exists. Then $f^{t_i-1} : S_i \to U_n$ monotone and onto. Moreover $f^{t_i-1}(f(x)) \in I_{i+1} \in \mathcal{U}_n$. Because $f^{t_i-1}(f(x)) \notin U_{n+1}$ we have that $I_{i+1} \neq U_{n+1}$ and $R_n : I_{i+1} \to U_n$ is monotone and onto. Now let $S_{i+1} = f^{-(t_i-1)}(I_{i+1}) \cap S_i$. $\square$

Observe that $f^{t_{i-1}-1}(S_i) = I_i \in \mathcal{U}_n$.

Lemma 3.3. — *There exist $n_0 \geq 1$ and $K < \infty$ with the following property. If the distortion of*

$$f^{t_i - 1} : S_i \longrightarrow U_n \text{ with } n \geq n_0$$

is bigger than K then

$$I_i \subset \Big(-\frac{3}{4} \cdot u_n, \frac{3}{4} \cdot u_n \Big).$$

Proof. — Lemma 2.2 states that $f^{t_1 - 1} : S_1 \to U_n$ has a monotone extension up to U_{n-1}, the map is essentially linear for big enough n. Hence $i \geq 2$. Consider the following decomposition

$$f^{t_i - 1}|S_i = f^{t_i - t_{i-1} - 1}|f(I_i) \circ f|I_i \circ f^{t_{i-1}-1}|S_i.$$

The factor $f^{t_{i-1}-1}|S_i$ has a monotone extension up to U_n. In particular, for big enough n, it is essentially linear. This is because the image I_i has a small hyperbolic length within U_n (Lemma 2.3). The factor $f^{t_i - t_{i-1} - 1}|f(I_i)$ has a monotone extension up to U_{n-1} (Lemma 2.2), it is also essentially linear. The distortion of $f^{t_i - 1}|S_i$ can only be caused by the factor $f|I_i$ and only if I_i is very close to 0. There is some n_0 such that $I_i \subset (-\frac{3}{4} \cdot u_n, \frac{3}{4} \cdot u_n)$, whenever $n \geq n_0$. Here we used Lemma 2.3 which states that I_i has also very small hyperbolic length within U_n. □

Lemma 3.4. — *For any $\theta < 1$ there exist $n_0 \geq 1$ and $C < \infty$ such that*

$$|S_i| \leq C \cdot u_{n+1}^{\alpha} \cdot \theta^{i-1},$$

whenever $n \geq n_0$ and $i \geq 2$.

Proof. — Observe that $f(0) \in S_1 \supset S_2 \supset \cdots \supset S_i$ and $f(U_{n+1}) \subset S_1$, $i \geq 2$. Let $L_j \subset S_1$ be the connected component of $S_1 - S_j$ with $L_j \subset f(U_{n+1})$, $2 \leq j \leq i$. For n big enough we get from the proof of Lemma 3.2 and from Lemma 2.3 that the hyperbolic length of S_j within S_{j-1} is very small, $2 \leq j \leq i$. It is easily seen that this implies

$$\begin{aligned} |S_i| &\leq C \cdot \theta^{i-1} \cdot |L_i| \\ &\leq C \cdot \theta^{i-1} \cdot |f(U_{n+1})| \\ &\leq C \cdot \theta^{i-1} \cdot u_{n+1}^{\alpha}. \end{aligned}$$

□

Lemma 3.5. — *There exist $n_0 \geq 1$ and $C > 0$ such that the following holds for $n \geq n_0$. If $|R_n(U_{n+1})| = |f^{t_1}(U_{n+1})| \geq \frac{1}{10} u_n$ then*

$$|Df^{t_1 - 1}_{|S_1}| \geq C \cdot \frac{u_n}{u_{n+1}^{\alpha}}.$$

If $|R_n(U_{n+1})| < \frac{1}{10} u_n$ then

$$|Df^{t_1 - 1}_{|S_1}| \geq C \cdot \frac{u_{n-1}}{u_n^{\alpha}}.$$

Proof. — Consider the map $f^{t_1-1}: S_1 \to U_n$. From Lemma 2.2 we know that this map has a monotone extension up to U_{n-1}. The map is essentially linear because $u_{n-1} \gg u_n$ whenever n is big enough. The derivative $|Df^{t_1-1}_{|S_1}|$ is essentially constant and can be estimated by

$$|Df^{t_1-1}_{|S_1}| \geq C\frac{|R_n(U_{n+1})|}{|f(U_{n+1})|} \geq C\frac{u_n}{u^{\alpha}_{n+1}}.$$

Here we used that $|R_n(U_{n+1})| \geq \frac{1}{10}u_n$. Now consider the other case: $|R_n(U_{n+1})| < \frac{1}{10}u_n$. Let $K \supset f(U_{n+1})$ be the interval which is mapped monotonically onto U_{n-1}: $f^{t_1-1}: K \to U_{n-1}$. Observe that $f^{-1}(K) \subset U_n$. This follows from the fact that the orbit of $f(u_n)$ never hits U_{n-1}, which was observed in section 2. Let $K' \subset K$ be such that $f^{t_1-1}: K' \to (-\frac{3}{4}\cdot u_{n-1}, \frac{3}{4}\cdot u_{n-1})$ is monotone and onto. This map has uniform bounded distortion because it has a monotone extension up to U_{n-1}. Let $K'' = f^{-1}(K') \subset U_n$. The derivative $|Df^{t_1-1}_{|S_1}|$ can be estimated by

$$|Df^{t_1-1}_{|S_1}| \geq C\frac{|f^{t_1}(K'')|}{|f(K'')|} \geq C\frac{u_{n-1}}{u^{\alpha}_{n}}.$$

□

Proof of Proposition 3.1. — Assume first that $s = 1$. This is an application of the previous Lemma 3.5. If $|R_n(U_{n+1})| \geq \frac{1}{10}u_n$ then

$$|Df^{t_1}(f(x))| \geq C\cdot\frac{u_n}{u^{\alpha}_{n+1}}\cdot u^{\alpha-1}_{n+1} = C\cdot\frac{u_n}{u_{n+1}} \geq C\cdot\rho_n,$$

where we used that $R_n(x) \notin U_{n+1}$.

In the other case when $|R_n(U_{n+1})| < \frac{1}{10}u_n$, we have

$$|Df^{t_1}(f(x))| \geq C\cdot\frac{u_{n-1}}{u^{\alpha}_{n}}\cdot u^{\alpha-1}_{n} = C\cdot\frac{u_{n-1}}{u_n} \geq C\cdot\rho_n,$$

where we used that in this case $f^{t_1-1}(x) \in R_n(U_{n+1})$ which is close to the boundary of U_n. The case with $s = 1$ is finished.

Now assume that $s \geq 2$. The proof will be split in two cases. Let $n \geq n_0 \geq 1$ be big enough such that Lemma 3.3 and 3.4 can be applied. Let $K < \infty$ be the constant from Lemma 3.3 and $\theta < 1$ the constant from Lemma 3.4.

Case I ($f^T(x) = R^s_n(x) \in (-\frac{3}{4}\cdot u_n, \frac{3}{4}\cdot u_n)$). — Let $H \subset S_s$ be such that $f^{T-1} = f^{t_s-1}: H \to (-\frac{3}{4}\cdot u_n, \frac{3}{4}\cdot u_n)$ is onto. This map has a monotone extension up to U_n.

Hence it has a uniformly bounded distortion,

$$
\begin{aligned}
|Df^T(f(x))| &\geq C \cdot \frac{|f^{T-1}(H)|}{|H|} \cdot u_{n+1}^{\alpha-1} \\
&\geq C \cdot \frac{u_n}{|S_s|} \cdot u_{n+1}^{\alpha-1} \\
&\geq C \cdot \frac{u_n}{\theta^{s-1} \cdot u_{n+1}^{\alpha}} \cdot u_{n+1}^{\alpha-1} \\
&\geq C \cdot \rho_n \cdot \theta^{-(s-1)}.
\end{aligned}
$$

Case II ($f^T(x) = R_n^s(x) \notin (-\frac{3}{4} \cdot u_n, \frac{3}{4} \cdot u_n)$). — If the distortion of $f^{T-1} : S_s \to U_n$ is bounded by K then we can give the same argument as in case I:

$$
\begin{aligned}
|Df^T(f(x))| &\geq C \cdot \frac{|U_n|}{|S_s|} \cdot u_{n+1}^{\alpha-1} \\
&\geq C \cdot \frac{u_n}{\theta^{s-1} \cdot u_{n+1}^{\alpha}} \cdot u_{n+1}^{\alpha-1} \\
&\geq C \cdot \rho_n \cdot \theta^{-(s-1)}.
\end{aligned}
$$

Now let us assume that this distortion is bigger than K. Apply Lemma 3.3, which states $I_s \subset (-\frac{3}{4} \cdot u_n, \frac{3}{4} \cdot u_n)$. Then

$$
\begin{aligned}
|Df^T(f(0))| &= |Df^{t_{s-1}}(f(0))| \cdot |Df^{T-t_{s-1}}(f^{t_{s-1}}(f(0)))| \\
&\geq C \cdot \rho_n \cdot \theta^{-(s-2)} \cdot |Df^{T-t_{s-1}}(f^{t_{s-1}}(f(0)))|.
\end{aligned}
$$

For $s-1 \geq 1$ we get this estimate from case I: $R_n^{s-1}(0) \in I_s \subset (-\frac{3}{4} \cdot u_n, \frac{3}{4} \cdot u_n)$. When $s-1 = 1$ it follows from the Proof of Proposition 3.1 for $s = 1$.

The last factor can be estimated by using the fact that $f^{T-t_{s-1}-1} : f(I_s) \to U_n$ has a monotone extension up to U_{n-1}, see Lemma 2.2. It is essentially linear and its derivative can be estimated

$$
|Df^{T-t_{s-1}-1}|_{f(I_s)}| \geq C \cdot \frac{|U_n|}{|f(I_s)|} \geq C \frac{|U_n|}{\varepsilon \cdot |f(U_n)|},
$$

where $\varepsilon > 0$ is given by Lemma 2.3. By taking $n_0 \geq 1$ big enough we can assume that $\varepsilon > 0$ is arbitrarily small.

Observe that $|Df(f^{T-1}(f(0)))| \geq C \cdot u_n^{\alpha-1}$. This is because $|f^{T-1}(f(0))| \geq \frac{3}{4} \cdot u_n$. We can finish the estimate for $|Df^T(f(0))|$ by observing that

$$
\begin{aligned}
|Df^{T-t_{s-1}}(f^{t_{s-1}}(f(0)))| &= |Df^{T-t_{s-1}-1}(f^{t_{s-1}}(f(0)))| \cdot |Df(f^{T-1}(f(0)))| \\
&\geq C \cdot \frac{|U_n|}{\varepsilon \cdot |f(U_n)|} \cdot u_n^{\alpha-1} \\
&\geq C \cdot \frac{u_n}{\varepsilon \cdot u_n^{\alpha}} \cdot u_n^{\alpha-1} \geq C \cdot \frac{1}{\varepsilon} \geq \frac{1}{\theta}.
\end{aligned}
$$

□

4. Telemann Decomposition of the Critical Orbit

In this section we will prove Theorem 1.3. Let f be a unimodal map such that

$$\sum_{n\geq 1} \sigma_n^{1/\alpha} < \infty.$$

The existence of an absolutely continuous invariant probability measure follows from [NS] in where it was shown that the Summability Condition on Derivatives

$$\sum_{k\geq 1} |Df^k(f(0))|^{-1/\alpha} < \infty$$

is sufficient for the existence of absolutely continuous invariant probability measures.

In the sequel we will prove that the summability of scaling factors implies the Summability Condition of derivatives. Choose $n_0 \geq 1$ big enough such that Proposition 3.1 can be applied and moreover

$$a = \sum_{\substack{n\geq n_0 \\ s\geq 0}} \frac{\theta^{s/\alpha}}{(C\cdot\rho_n)^{1/\alpha}} = \frac{1}{1-\theta^{1/\alpha}} \sum_{n\geq n_0} \frac{1}{(C\cdot\rho_n)^{1/\alpha}} < 1,$$

where C and θ are the constants from Proposition 3.1.

Fix $k \geq 1$. The estimate for $|Df^k(f(0))|$ is based on the *Telemann decomposition* of the critical orbit up to time k. Consider the orbit of $f(0)$ up to time $k-1$. Let $m \geq 0$ be such that U_{n_0+m} is the smallest central interval which is crossed by this piece of the orbit:

$$n_0 + m = \max\{j \geq 0 \mid \exists\, 0 < i \leq k,\ f^i(0) \in U_j\}$$

and the last moment of crossing is denoted by

$$k_m = \max\{1 \leq j \leq k \mid f^j(0) \in U_{n_0+m}\}.$$

The moments $k_m \leq k_{m-1} \leq \ldots k_1 \leq k_0$ are such that k_i is the last moment that the orbit hits U_{n_0+i}: if $\{k_i < j \leq k \mid f^j(0) \in U_{n_0+i-1}\} = \varnothing$ then $k_{i-1} = k_i$ otherwise

$$k_{i-1} = \max\{k_i < j \leq k \mid f^j(0) \in U_{n_0+i-1}\} \text{ with } 1 \leq i \leq m.$$

Let $r(k) = k - k_0$ and if $k_{i-1} \neq k_i$ then

$$s_{i-1}(k) = \#\{k_i < j \leq k_{i-1} \mid f^j(0) \in U_{n_0+i-1}\}, \quad 1 \leq i \leq m,$$

the number of returns trough U_{n_0+i-1}.

The chain-rule applied to $Df^k(f(0))$ gives

$$Df^k(f(0)) = Df^{r(k)}(f^{k_0}(f(0))) \cdot Df^{k_m}(f(0)) \cdot \prod_{i=0}^{m-1} Df^{k_i - k_{i+1}}(f^{k_{i+1}}(f(0))).$$

The first factor can be estimated by using

Proposition 4.1 **([G], [Ma]).** — *Given $n_0 \geq 1$ there exist constants $C > 0$ and $\lambda > 1$ such that*

$$|Df^r(x)| \geq C\lambda^r,$$

whenever $f^i(x) \notin U_{n_0}$ with $i \leq r$.

The other factors can be estimated by Proposition 3.1. The decomposition was set up to make Proposition 3.1 applicable to the factors:

$$\begin{aligned}|Df^k(f(0))|^{-1/\alpha} &\leq \Big(C\lambda^{r(k)} \cdot \prod_{\substack{i\leq m-1\\ k_i\neq k_{i+1}}} C\cdot \rho_{n_0+i}\cdot\theta^{-(s_i(k)-1)}\Big)^{-1/\alpha}\\ &\leq C\Big(\frac{1}{\lambda^{1/\alpha}}\Big)^{r(k)} \cdot \prod_{\substack{i\leq m-1\\ k_i\neq k_{i+1}}} \frac{(\theta^{1/\alpha})^{s_i(k)-1}}{(C\cdot\rho_{n_0+i})^{1/\alpha}}\end{aligned}$$

Lemma 4.2. — *Let s_i, r and s'_i, r' correspond to the Teleman decomposition of respectively k and k'. If $k \neq k'$ then $r \neq r'$ or $s_i \neq s'_i$ for some $i \geq 0$.*

Proof. — Assume that $r = r'$ and $s_i = s'_i$ for all $i \geq 1$. We have to show that $k = k'$. Observe that $f^{k_m}(0) = R^{s_m}_{n_0+m}(0)$ but also $f^{k'_m}(0) = R^{s'_m}_{n_0+m}(0) = R^{s_m}_{n_0+m}(0)$. So $k_m = k'_m$. Now repeat this argument to show that $k_i = k'_i$ for $0 \leq i \leq m$. In particular we get $k_0 = k'_0$. So

$$k' = k'_0 + r' = k_0 + r = k. \qquad \square$$

Proof of the Summability Condition for Derivatives. — The number $a < 1$ was defined in the beginning of this section. Let $\beta = 1/\alpha$.

$$\begin{aligned}\sum_{k\geq 0}|Df^k(f(0))|^{-1/\alpha} &= \sum_{r\geq 0}\sum_{\substack{k\geq 0\\ r(k)=r}} |Df^k(f(0))|^{-1/\alpha}\\ &\leq \sum_{r\geq 0} C\Big(\frac{1}{\lambda^\beta}\Big)^r \cdot \sum_{\substack{k\geq 0\\ r(k)=r}} \prod_{\substack{0\leq i\leq m-1\\ k_i\neq k_{i+1}}} \frac{(\theta^\beta)^{s_i(k)-1}}{(C\cdot\rho_{n_0+i})^\beta}.\end{aligned}$$

Now observe that for each r there are no two products appearing in the second sum which are formed by the same factors, according to Lemma 4.2. If we expand a^n we see that the sum of all possible different products can be estimated by $1+a+a^2+a^3+\ldots$. Hence

$$\begin{aligned}\sum_{k\geq 0}|Df^k(f(0))|^{-1/\alpha} &\leq \sum_{r\geq 0} C\Big(\frac{1}{\lambda^\beta}\Big)^r\cdot(1+a+a^2+a^3+\ldots)\\ &\leq \frac{1}{1-a}\sum_{r\geq 0} C\Big(\frac{1}{\lambda^\beta}\Big)^r\\ &\leq C\cdot\frac{1}{1-a}\cdot\frac{1}{1-1/\lambda^\beta} < \infty.\end{aligned}$$

5. Appendix

In this appendix some basic notions of interval dynamics are collected. The details can be found in [**MS**].

The hyperbolic length of an interval $I \subset T \subset [-1,1]$ within T is defined to be

$$\mathrm{hyp}(I,T) = \ln \frac{|L \cup I| \cdot |R \cup I|}{|L| \cdot |R|},$$

where $L, R \subset T$ are the connected components of $T \setminus I$ and $|J|$ stands for the length of the interval $J \subset [-1,1]$.

The Schwarzian derivative of a C^3 map $f : [-1,1] \to [-1,1]$ is

$$Sf(x) = \frac{D^3 f(x)}{Df(x)} - \frac{3}{2} \cdot \frac{D^2 f(x)}{Df(x)}.$$

where $D^i f(x)$ stands for the i^{th} derivative of f in $x \in [-1,1]$.

Expansion-Lemma 5.1. — *If $f : [-1,1] \to [-1,1]$ has $Sf(x) \leq 0$ for all $x \in [-1,1]$ and $f^n|T$ is monotone then*

$$\mathrm{hyp}(f^n(I), f^n(T)) \geq \mathrm{hyp}(I,T),$$

where $I \subset T$.

Koebe-Lemma 5.2. — *For each $\tau > 0$ there exists $1 \leq K(\tau) < \infty$ with the following property. Let $f^n : T \to f^n(T)$ be monotone and $Sf(x) \leq 0$ for all $x \in [-1,1]$. If $I \subset T$ is an interval such that both component $L, R \subset T \setminus I$ satisfy*

$$\frac{|f^n(L)|}{|f^n(T)|}, \frac{|f^n(R)|}{|f^n(T)|} \geq \tau$$

then $f^n|I$ has bounded distortion

$$\frac{|Df^n(x)|}{|Df^n(y)|} \leq K(\tau),$$

for all $x, y \in I$. Moreover $K(\tau) \to 1$ when $\tau \to \infty$.

An S-unimodal map is a C^3 endomorphism $f : [-1,1] \to [-1,1]$ such that

(1) $f(\pm 1) = -1$,
(2) $Df(x) > 0$ for $x < 0$,
(3) $Df(x) < 0$ for $x > 0$,
(4) $Df(0) = 0$ and up to a C^1 change of coordinates f equals locally $x \to |x|^\alpha$ with $\alpha > 1$. The point $x = 0$ is called the critical point and $\alpha > 1$ is called the critical exponent of f.
(5) $Sf(x) < 0$, $x \neq 0$.

The orbit of the critical point $x = 0$ is called the critical orbit. The critical orbit is said to be recurrent if it accumulates at the critical point.

Usage of constants. — Every uniform constant, that is a constant which is independent of the actual map, appearing in estimates will be denoted by C. Constants which play a specific role in the statement of Lemmas will usually have a specific name.

References

[BKNS] H. Bruin, G. Keller, T. Nowicki, S. van Strien, *Wild Attractors Exist*, Ann. of Math. **143** (1996), 97-130.

[B] L.A. Bunimovich, *A Transformation of the Circle*, Math. Notes **8** no. 2 (1970), 587-592.

[Br] H. Bruin, *Invariant Measures of Interval Maps*, Thesis (1994), Technical University Delft, Delft.

[BL1] A.M. Blokh, M. Lyubich, *Attractors of Maps on the interval*, Func. Anal. and Appl. **21** (1987), 32-46.

[BL2] A.M. Blokh, M. Lyubich, *Measure of Solenoidal Attractors of Unimodal Maps of the Segment*, Math. Notes **48** no. 5 (1990), 1085-1090.

[G] J. Guckenheimer, *Limit Sets of S−unimodal Maps with Entropy Zero*, Comm. Math. Phys. **110** (1987), 655-659.

[GJ] J. Guckenheimer, S. Johnson, *Distortion of S−unimodal Maps*, Ann. of Math. **132** (1990), 71-130.

[GS] J. Graczyk, G. Swiatek, *Hyperbolicity in the Real Quadratic Family*, Report no. PM 192 PennState (1995).

[HK] F. Hofbauer, G. Keller, *Quadratic Maps without Asymptotic Measure*, Comm. Math. Phys. **127** (1990), 319-337.

[Ja] M.V. Jakobson, *Absolutely Continuous Invariant Measures for one-parameter Families of One-dimensional Maps*, Comm. Math. Phys. **81** (1981), 39-88.

[Jo] S. Johnson, *Singular Measures without restrictive interval*, Comm. Math. Phys. **110** (1987), 185-190.

[K] G. Keller, *Exponents, Attractors and Hopf Decomposition for Interval Maps*, Erg. Th. & Dyn. Sys. **10** (1990), 717-744.

[L1] M. Lyubich, *Combinatorics, Geometry and Attractors of Quasi-Quadratic Maps*, Ann. of Math **140** (1994), 347-404.

[L2] M. Lyubich, *Geometry of Quadratic Polynomials: Moduli, Rigidity, and Local Connectivity*, IMS at Stony Brook preprint 1993/9.

[L3] M. Lyubich, *Dynamics of Quadratic Polynomials, II Rigidity*, IMS at Stony Brook preprint 1995/14.

[L4] M. Lyubich, *Dynamics of Quadratic Polynomials, III Parapuzzles*, IMS at Stony Brook preprint 1996/5 and this volume.

[LM] M. Lyubich, J.W. Milnor, *The Fibonacci Unimodal Map*, J.A.M.S. **6** (1993), 425-457.

[Ld] F. Ledrappier, *Some Properties of Absolutely Continuous Invariant Measures of an Interval*, Erg. Th. & Dyn. Sys. **1** (1981), 77-93.

[Ma] R. Mañé, *Hyperbolicity, Sinks, and Measure in One-dimensional Dynamics*, Comm. Math. Phys. **100** (1985), 495-524 and Erratum, Comm. Math. Phys. **112** (1987), 721-724.

[M1] M. Martens, *Interval Dynamics*, Thesis (1990), Technical University Delft, Delft, or **[M2]**.

[M2] M. Martens, *Distortion Results and Invariant Cantor Sets of Unimodal Maps*, Erg. Th. & Dyn. Sys. **14** (1994), 331-349.

[Mi] M. Misiurewicz, *Absolutely Continuous Measures for certain maps of an Interval*, Publ. Math. I.H.E.S. **53** (1981), 17-51.

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

[NS] T. Nowicki, S. van Strien, *Absolutely Continuous Measures under a Summability Condition*, Invent. Math. **93** (1988), 619-635.

[NU] J. von Neumann, S.M. Ulam, *On Combinatorics of Stochastic and Deterministic Processes*, Bull. A.M.S. **53** (1947), 1120.

[P] J. Palis, *A Global View on Dynamics: Conjectures and Results*, this volume.

[R] D. Ruelle, *Applications conservant une mesure absolument continue par rapport à dx sur* $[0,1]$, Comm. Math. Phys. **55** (1977), 47-51.

[VT] J.J.P. Veerman, F.M. Tangerman, *Scalings in Circle Maps (I)*, Comm. Math. Phys. **134** (1990), 89-107.

M. MARTENS, Institute for Mathematical Sciences, SUNY at Stony Brook, Stony Brook, NY 11794-3651 • *E-mail :* martens@watson.ibm.com

T. NOWICKI, University of Warsaw, Warsaw, Poland

Astérisque

JEAN-FRANÇOIS MATTEI

Quasi-homogénéité et équiréductibilité de feuilletages holomorphes en dimension deux

Astérisque, tome 261 (2000), p. 253-276

<http://www.numdam.org/item?id=AST_2000__261__253_0>

Astérisque
261, 2000, p. 253–276

QUASI-HOMOGÉNÉITÉ ET ÉQUIRÉDUCTIBILITÉ DE FEUILLETAGES HOLOMORPHES EN DIMENSION DEUX

par

Jean-François Mattei

Résumé. — Après avoir étudié la dépendance analytique des séparatrices d'une famille « équisingulière » de germes de feuilletages holomorphes à l'origine de $\mathbb{C}^2$, nous définissons la quasihomogénéité comme une propriété de rigidité. Nous obtenons un théorème de type K. Saito pour les germes de feuilletages quasi-homogènes et un théorème de type Briançon-Skoda dans le cas général.

Vocabulaire et notations

On désignera par $\mathcal{O}_M$ et par Λ^1_M respectivement les faisceaux des germes de fonctions holomorphes et de 1-formes différentielles holomorphes sur une variété holomorphe M. Pour $m \in Z \subset M$ le germe de Z en m sera noté Z, m.

Le *lieu singulier* $\mathrm{Sing}(\eta)$ d'une 1-forme $\eta \in \Lambda^1_M(M)$ est l'ensemble des points m de M où $\eta(m) \in T^*_m M$ est nulle. Lorsque $\mathrm{Sing}(\eta), m$ est de codimension 1 on peut décomposer le germe η_m de η en ce point,

$$\eta_m = h\eta'_m \qquad h \in \mathcal{O}_{M,m}, \quad \eta' \in \Lambda^1_{M,m},$$

de manière que $\mathrm{codim}\,(\mathrm{Sing}(\eta'_m)) \geq 2$. Cette décomposition est unique, à unité multiplicative près et η'_m s'appellera le *saturé local de η au point m.* On notera

$$\Sigma(\eta) := \{m \in M \mid \eta'_m(m) = 0\}$$

que l'on appellera *lieu singulier strict de η.*

Lorsque la forme η vérifie la *condition d'intégrabilité* $\eta \wedge d\eta \equiv 0$, elle définit un feuilletage $\mathcal{F}_\eta$, éventuellement singulier, de *lieu singulier* $\mathrm{Sing}(\mathcal{F}_\eta) := \mathrm{Sing}(\eta)$. La collection des saturés locaux de η définit le *feuilletage saturé* $\mathcal{F}^{\mathrm{sat}}_\eta$ *associé à* $\mathcal{F}_\eta$; son *lieu singulier* $\mathrm{Sing}(\mathcal{F}^{\mathrm{sat}}_\eta)$ est égal à $\Sigma(\eta)$.

***Classification mathématique par sujets* (1991).** — 32A10, 32A20, 32B10,34C20, 34C35, 58F23, 32G34, 32S15, 32S30, 32S45, 32S65.
Mots clefs. — Singularités, champs de vecteurs, feuilletage, equisingularité, courbes, quasi-homogénéité, déformation, modules, formes normales.

Un sous-ensemble analytique Z de M est dit *ensemble invariant* (resp. *strictement invariant*) de η si la restricion de η (resp. des saturés locaux de η) à la partie régulière de Z est identiquement nulle. Lorsque η est intégrable, un *ensemble invariant de* $\mathcal{F}_{\eta}^{\text{sat}}$ est un ensemble strictement invariant de η.

Si au voisinage d'un point m d'une sous-variété lisse V de M le saturé local η'_m de η vérifie l'une des conditions suivantes :

(1) $m \in \Sigma(\eta)$,

(2) $m \notin \Sigma(\eta)$, $\eta'_m \mid_V \not\equiv 0$ et $\eta'_m \mid_V (m) = 0$,

on dira que m est un *point singulier strict du couple* $(\eta; V)$. Pour une hypersurface H à croisements normaux $\Sigma(\eta; H)$ désignera l'ensemble des points $m \in H$ qui sont points singuliers stricts d'au moins un des couples formés par η et une composante irréductible locale de H en m.

Fixons un point t d'une variété holomorphe P. Un sous-ensemble analytique W de M est dit *étale au dessus de t via une submersion* $\pi : M \longrightarrow P$ si la restriction de π à W est un difféomorphisme local en chaque point de $\pi^{-1}(t) \cap W$. Cette propriété est ouverte lorque $\pi \mid_W$ est propre. On notera $W_t := \pi^{-1}(t) \cap W$.

Conventions. — Lorsque nous ne précisons pas, les objets considérés ici : variété, courbe, application, champ de vecteurs... sont supposés holomorphes. Nous dirons aussi « 1-forme » pour : « 1-forme différentielle holomorphe ».

1. Introduction

Classiquement la *quasi-homogénéité* d'un germe $f(x, y)$ de fonction holomorphe — ici à l'origine de $\mathbb{C}^2$ — se définit par l'existence de coordonnées $u(x, y), v(x, y)$ dans lesquelles f est *un polynôme quasi-homogène* : son nuage de Newton est contenu dans une droite :

$$f = P(u, v) = \sum_{\alpha i + \beta j = d} P_{ij} u^i v^j.$$

Cette propriété se caractérise algébriquement **[16]** par l'appartenance de f à son *idéal jacobien* $J(f) := (f'_x, f'_y)$. On dispose aussi d'une caractérisation en terme de comparaison des modules de la fonction f à de ceux de la courbe $f^{-1}(0)$. On peut montrer que f est quasi-homogène si et seulement si toute déformation de f

$$F \in \mathcal{O}_{\mathbb{C}^{2+p}, 0}, \quad F_0 = f, \quad F_t(0,0) = 0, \qquad F_t(x, y) := F(x, y; t),$$

qui est topologiquement triviale, satisfait l'équivalence : F est analytiquement triviale $\Longleftrightarrow$ la famille de germes de courbes $\left(F_t^{-1}(0)\right)_t$ est analytiquement triviale ;

La *trivialité topologique*, resp. la *trivialité analytique*, de F signifie ici l'existence de germes d'homéomorphismes, resp. de biholomorphismes, Φ de $\mathbb{C}^{2+p}, 0 \xrightarrow{\sim} \mathbb{C}^{2+p}, 0$ et L de $\mathbb{C}^{1+p}, 0 \xrightarrow{\sim} \mathbb{C}^{1+p}, 0$, qui commutent avec les projections sur $\mathbb{C}^p, 0$, valent l'identité pour $t = 0$ et satisfont : $L \circ (F;t) \circ \Phi = (f;t)$. Cette dernière relation signifie que Φ conjugue les feuilletages de $\mathbb{C}^{2+p}, 0$ définis par dF et par $df \circ \pi$, avec $\pi(x, y; t) := (x, y)$.

L'objet principal de cet article est d'étudier les notions de quasi-homogénéité pour un germe de feuilletage holomorphe $\mathcal{F}_\omega$ à l'origine de $\mathbb{C}^2$ défini par un germe de 1-forme

$$\omega := a(x,y)dx + b(x,y)dy, \quad a, b \in \mathcal{O}_{\mathbb{C}^2,0} \quad \mathrm{Sing}(\omega) = \{0\}.$$

D'après un résultat de C. Camacho et P. Sad **[4]** un tel feuilletage admet toujours une *séparatrice*, c'est à dire un germe à l'origine de courbe analytique irréductible invariante. Il peut en admettre une infinité, on dit alors que ω est *dicritique.* Si ce n'est pas le cas, nous notons $\mathrm{Sep}(\omega)$ ou $\mathrm{Sep}(\mathcal{F}_\omega)$ le germe de l'union des séparatrices de ω.

Définition 1.1. — Nous disons qu'un germe de 1-forme holomorphe non-dicritique ω à l'origine de $\mathbb{C}^2$ est *topologiquement quasi-homogène* si toute déformation topologiquement triviale $\eta := A(x,y;t)dx + B(x,y;t)dy$ de ω satisfait l'équivalence :

- η est analytiquement triviale si et seulement si la déformation de germes de courbes à l'origine $(\mathrm{Sep}(\eta_t))_t$ est analytiquement triviale,

où η_t désigne le germe à l'origine de la restriction de η à $\mathbb{C}^2 \times t$.

Pour les feuilletages quasi-hyperboliques génériques (définis au chapitre 6) on dispose d'un critère différentiel de trivialité topologique **[14] [15]** :

- Une déformation η d'une 1-forme quasi-hyperbolique générique est topologiquement triviale si et seulement si il existe des germes $C_j \in \mathcal{O}_{\mathbb{C}^{2+p},0}$ tels que la 1-forme $\Omega := \eta + \sum_j C_j(x,y;t)dt_j$ est intégrable (*i.e.* $\Omega \wedge d\Omega \equiv 0$) et définit un déploiement équisingulier de $\mathcal{F}_\omega$, *cf.* (6.1).

L'espace (lisse de dimension finie) universel des déploiements équisinguliers construit dans **[10]** s'interprète donc comme l'espace des modules analytiques dans une classe topologique donnée. On est ainsi ammené à définir au chapitre 6 la notion de *quasi-homogénéité au sens des déploiements* ou encore *d-quasi-homogénéité* d'une 1-forme ω par la propriété suivante : « Tout déploiement équisingulier Ω de ω satisfait l'équivalence : Ω est analytiquement trivial $\Longleftrightarrow$ la famille de germes de courbe $(\mathrm{Sep}(\Omega_t))_t$ est analytiquement triviale ». On peut alors affaiblir les hypothèses et considérer la classe plus large des 1-formes non-dicritiques dont la résolution ne comporte aucune singularité de type selle-nœud ; ces formes seront appellées *semi-hyperboliques.* On obtient :

Théorème A. — *Soit $\omega := a(x,y)dx + b(x,y)dy$ un germe de 1-forme holomorphe semi-hyperbolique à l'origine de $\mathbb{C}^2$ et $f(x,y) \in \mathcal{O}_{\mathbb{C}^2,0}$ une équation réduite de* $\mathrm{Sep}(\omega)$. *Alors les assertions suivantes sont équivalentes :*

(1) *ω est d-quasi-homogène,*

(2) *f appartient à l'idéal* $(a, b) \subset \mathcal{O}_{\mathbb{C}^2,0}$,
(3) *df est topologiquement quasi-homogène,*
(4) *f appartient à l'idéal* $(f'_x, f'_y) \subset \mathcal{O}_{\mathbb{C}^2,0}$,
(5) *il existe des coordonnées* u, v *à l'origine, des fonctions* $g, h \in \mathcal{O}_{\mathbb{C}^2,0}$ *avec* $u(0) = v(0) = 0$, $g(0) \neq 0$ *et des entiers* α, β, $d \in \mathbb{N}$ *tels que : f s'exprime comme un polynôme quasi-homogène* $f = \sum_{\alpha i+\beta j=d} P_{ij} u^i v^j \in \mathbb{C}[[u, v]]$ *et* $g\omega = df + h(\beta v du - \alpha u dv)$.

Ce théorème appliqué aux 1-formes quasi-hyperboliques génériques donne directement :

Théorème B. — *Soit* $\omega := a(x, y)dx + b(x, y)dy$ *un germe de 1-forme holomorphe quasi-hyperbolique générique à l'origine de* $\mathbb{C}^2$ *et* $f(x, y) \in \mathcal{O}_{\mathbb{C}^2,0}$ *une équation réduite de* Sep(ω). *Alors les assertions* (2) (3) (4) *et* (5) *du théorème A sont équivalentes à*

(1′) ω *est topologiquement quasi-homogène.*

Pour montrer le théorème A nous serons ammenés (6.7) à prouver le

Théorème. — *Soit* ω *un germe de 1-forme semi-hyperbolique. Alors toute déformation topologiquement triviale* $(X_t)_t$ *de* $X_0 :=$ Sep(ω) *est réalisée comme la famille de séparatrices d'un déploiement équisingulier. En particulier il existe une déformation topologiquement triviale* η *de* ω, *à pseudo-groupe d'holonomie constant, telle que pour t assez petit* $X_t =$ Sep(η_t).

Lorsque ω n'est pas quasi-homogène, nous mettons en évidence sur l'espace universel des déploiements équisinguliers un « feuilletage singulier », défini par un sous-module involutif du module des germes de champs de vecteurs holomorphes, dont « l'espace des feuilles » s'identifie à l'espace des modules locaux du germe de courbe Sep(ω).

Enfin au chapitre 7 nous comparons l'idéal (a, b) de $\mathcal{O}_{\mathbb{C}^2,0}$ engendré par les coefficients de ω, à l'idéal définissant Sep(ω). On obtient le résultat suivant qui généralise — et redémontre — un théorème de J. Briançon **[2]**, **[3]** :

Théorème C. — *Soit* $\omega := a(x, y)dx + b(x, y)dy$ *un germe de 1-forme holomorphe semi-hyperbolique à l'origine de* $\mathbb{C}^2$ *et* $f = 0$ *une équation réduite de* Sep(ω). *Alors* f^2 *appartient à l'idéal* (a, b).

2. Réduction des singularités et équiréduction

Rappelons qu'un germe de 1-forme singulière $\omega := a(x, y)\,dx + b(x, y)\,dy$ à l'origine de $\mathbb{C}^2$ est dit *réduit* s'il existe des coordonnées locales u, v, $u(0) = v(0) = 0$, dans lesquelles le saturé de ω à l'origine s'ecrit $\omega'_0 = udv + \lambda vdu + \cdots$, $\lambda \notin \mathbb{Q}_{<0}$. Lorsque le saturé de $\mathcal{F}_\omega$ est régulier à l'origine nous dirons encore que ω est réduit. Plus généralement, η désignant une 1-forme sur une surface M et $Z \subset M$ une courbe à croisements normaux, on dira que *le couple* (η, Z) *est réduit en* $m \in Z$ si le saturé η'_m de η en m est réduit et satisfait une des deux conditions suivantes :

(1) $\eta'_m(m) = 0$, η est réduit, et Z est strictement invariant,

(2) $\eta'_m(m) \neq 0$ et $m \notin \Sigma(\eta; Z)$.

Nous notons $C(\eta, Z)$ l'ensemble des points de M où (η, Z) n'est pas réduit.

Lorsque ω n'est pas réduit, on lui associe une succession d'applications d'éclatement $E^i_\omega : M^i_\omega \longrightarrow M^{i-1}_\omega$ de la manière suivante :

- $M^0_\omega = \mathbb{C}^2$ et E^1_ω est l'application d'éclatement de l'origine,
- $E^1_\omega, \ldots, E^i_\omega$ étant définis, désignons par $E_{\omega,i}$ le composé $E^1_\omega \circ \cdots \circ E^i_\omega$, par D^i_ω le diviseur exceptionnel ${E_{\omega,i}}^{-1}(0)$ et notons $\omega_i := {E_{\omega,i}}^*(\omega)$. Alors E^{i+1}_ω est l'éclatement simultané des points de $C\left(\omega_i, D^i_\omega\right)$.

Un théorème classique de A. Seidenberg *cf.* **[17]** ou **[11]**, assure que cette succession d'éclatements s'arrête à une certaine hauteur h_ω :

$$C\left(\omega_{h_\omega}, D^{h_\omega}_\omega\right) = \varnothing \qquad \text{et} \qquad C\left(\omega_{h_\omega - 1}, D^{h_\omega - 1}_\omega\right) \neq \varnothing.$$

Nous notons dans toute la suite :

$$\begin{gathered} E_\omega := E_{\omega,h_\omega}, \quad M_\omega := M^{h_\omega}_\omega, \quad D_\omega := D^{h_\omega}_\omega, \\ \widetilde{\omega} := {E_\omega}^*(w), \quad \widetilde{\mathcal{F}}_\omega := \mathcal{F}^{\mathrm{sat}}_{\widetilde{\omega}}, \quad \widetilde{\Sigma}_\omega := \Sigma\left(\widetilde{\omega}; D^{h_\omega}_\omega\right). \end{gathered} \tag{1}$$

Une *déformation de* $\omega := a(x,y)\,dx + b(x,y)\,dy$ *de base* $\mathbb{C}^p, 0$ est un germe de 1-forme du type

$$\eta = A(x,y;\, t)\,dx + B(x,y;\, t)\,dy, \quad A, B \in \mathcal{O}_{\mathbb{C}^{2+p},0}$$

qui vérifie :

$$A(x,y;0) = a(x,y), \quad B(x,y;0) = b(x,y), \quad A(0,0;t) = B(0,0;t) = 0.$$

Ainsi η définit un « germe de famille analytique de germes de 1-formes », en considérant les restrictions η_t de η à $\mathbb{C}^2 \times t, (0,t)$. Nous désignerons souvent la déformation η par $(\eta_t)_{t \in \mathbb{C}^p,0}$.

Définition 2.1. — Une déformation η d'un germe non réduit $\omega \in \Lambda^1_{\mathbb{C}^2,0}$ est dite *équiréductible* s'il existe une succession finie d'éclatements $E^i_\eta : M^i_\eta \longrightarrow M^{i-1}_\eta$ telle que

(1) chaque E^i_η est un éclatement de centre une sous variété-lisse $C^{i-1}_\eta \subset M^{i-1}_\eta$ de codimension 2,

(2) notons encore $E_{\eta,i} := E^1_\eta \circ \cdots \circ E^i_\eta$, $D^i_\eta := {E_{\eta,i}}^{-1}(0)$, $\eta_i := {E_{\eta,i}}^*(\eta)$ et $\pi_{\eta,i} := \pi \circ E_{\eta,i}$ où π est la projection $\pi(x,y;t) := t$. Alors les variétés $\Sigma(\eta_i, D_{\eta,i})$ et $C(\eta_i, D_{\eta,i})$ sont lisses et étales via $\pi_{\eta,i}$ au dessus de $\mathbb{C}^p, 0$ et sont contenues dans D^i_η,

(3) pour $t \in \mathbb{C}^p$ assez petit la succession d'éclatements obtenus par restriction de chaque E^i_η aux fibres de $\pi_{\eta,i}$ et $\pi_{\eta,i-1}$ est exactement la succession d'éclatements de la réduction de η_t.

Lorsque ω est réduit on convient que l'équiréductibilité de η signifie que $\Sigma(\eta) = 0 \times \mathbb{C}^p, 0$ et que pour tout $t \in \mathbb{C}^p$ assez petit, η_t est réduit.

Remarquons que $\eta_t := xdy + (\lambda + t)ydx + \cdots$ avec $\lambda \in \mathbb{R}_{\leq 0}$ n'est pas une déformation équiréductible, puisque η_t n'est pas réduite pour $t \in -\lambda + \mathbb{Q}_{<0}$.

3. Formes semi-hyperboliques

Ce sont les formes qui dans **[5]** s'appelleraient « courbes généralisées non-dicritiques ». Conservons les notations des paragraphes précédents.

Définition 3.1. — On dit qu'un germe à l'origine de $\mathbb{C}^2$ de 1-forme à singularité isolée $\omega := a(x, y)\,dx + b(x, y)\,dy$ est *semi-hyperbolique* si D_ω est un sous ensemble invariant de $\widetilde{\mathcal{F}}_\omega$ et si toutes les singularités de $\widetilde{\mathcal{F}}_\omega$ sont du type : $udv + \lambda vdu + \cdots$ avec $\lambda \notin \mathbb{Q}_{\leq 0}$.

La première condition signifie que ω n'est pas dicritique (*cf.* Introduction). Il est prouvé dans **[5]** que la seconde condition implique les trois propriétés suivantes, où $f \in \mathcal{O}_{\mathbb{C}^2,0}$ désigne une équation réduite de $\mathrm{Sep}(\omega)$:

P1. La réduction de ω est exactement la *réduction de f*, c'est à dire la réduction de df ;

P2. Le *nombre de Milnor* de ω à l'origine, $\mu_0(\omega) := \dim_{\mathbb{C}} \left(\mathcal{O}_{\mathbb{C}^2,0}/(a, b)\right)$ est égal au *nombre de Milnor de f* à l'origine $\mu_0(f) := \dim_{\mathbb{C}} \left(\mathcal{O}_{\mathbb{C}^2,0}/(f'_x, f'_y)\right)$. De plus cette propriété est caractéristique : si ω (non-dicritique) n'est pas semi-hyperbolique, alors $\mu_0(f) < \mu_0(\omega)$;

P3. Associons à chaque composante irréductible D de D_ω les poids suivants :
  (1) *la multiplicité $m_\omega(D)$ de D suivant ω*, c'est à dire le plus grand entier k tel que $u^{-k}\widetilde{\omega}$ soit holomorphe, $u \in \mathcal{O}_{M_\omega,m}$ désignant une équation réduite locale de D en un point $m \in D$ quelconque,
  (2) *la multiplicité $m_f(D)$ de D*, comme composante de $\{f \circ E_\omega = 0\}$, c'est à dire $m_f(D) := m_{df}(D) + 1$;

Alors, pour toute composante irréductible D de D_ω, on a : $m_\omega(D) = m_f(D) - 1$.

Notons maintenant
$$S'_\omega := {E_\omega}^{-1}\left(\mathrm{Sep}(\omega)\right)$$
la courbe formée de D_ω et des transformées strictes de $\mathrm{Sep}(\omega)$. Considérons le faisceau $\mathcal{X}_{M_\omega, S'_\omega}$ des des germes de champs de vecteurs de M_ω tangents à S'_ω en chacun de ses points réguliers et le sous-faisceau $\mathcal{X}_{M_\omega, \widetilde{\mathcal{F}}_\omega}$ des champs tangents au feuilletage $\widetilde{\mathcal{F}}_\omega$. Restreignons ces faisceaux à D_ω

$$\mathcal{X}_{\widetilde{\mathcal{F}}_\omega} := i^{-1}\left(\mathcal{X}_{M_\omega, \widetilde{\mathcal{F}}_\omega}\right), \qquad \mathcal{X}_{S'_\omega} := i^{-1}\left(\mathcal{X}_{M_\omega}, S'_\omega\right), \tag{2}$$

où $i : D_\omega \hookrightarrow M_\omega$ désigne l'application d'inclusion. Ce sont des faisceaux de modules cohérents sur
$$\widetilde{\mathcal{O}} := i^{-1}(\mathcal{O}_{M_\omega}).$$

***Lemme 3.2* (clé).** — *Soit ω un germe de 1-forme semi-hyperbolique à l'origine de $\mathbb{C}^2$. Alors le sous-faisceau d'idéaux de $\widetilde{\mathcal{O}}$ engendré par l'image du morphisme*

$$\widetilde{\omega}\cdot : \mathcal{X}_{S'_\omega} \longrightarrow \widetilde{\mathcal{O}}, \qquad Z \longmapsto \widetilde{\omega}\cdot Z$$

est égal au faisceau d'idéaux $(f \circ E_\omega)$ engendré par la section globale $f \circ E_\omega$ de $\widetilde{\mathcal{O}}$; En d'autres termes, on a la suite exacte :

$$0 \longrightarrow \mathcal{X}_{\widetilde{\mathcal{F}}_\omega} \xrightarrow{\rho} \mathcal{X}_{S'} \xrightarrow{\widetilde{\omega}\cdot} (f \circ E_\omega) \longrightarrow 0, \tag{3}$$

où ρ désigne le morphisme d'inclusion, et $\widetilde{\omega} := E_\omega{}^*(w)$.

Démonstration. — En un point c de D_ω fixons des coordonnées locales u, v dans lesquelles $f \circ E_\omega = u^p v^q$ avec $q \neq 0$. On est nécéssairement dans l'une des deux éventualités suivantes :

(a) $p = 0$. — Alors D_ω et $\widetilde{\mathcal{F}}_\omega$ sont réguliers au point c et d'après la propriété P3 ci-dessus on a

$$\widetilde{\omega} = v^{q-1}\left(A(u,v)vdu + B(u,v)dv\right) \quad \text{avec} \quad B(0,0) \neq 0,$$

$$\mathcal{X}_{S'_\omega,c} = \widetilde{\mathcal{O}}_c \frac{\partial}{\partial u} \oplus \widetilde{\mathcal{O}}_c v \frac{\partial}{\partial v}.$$

(b) $p \neq 0$. — Alors si D_ω est régulier, $p = 1$ car f est réduit. Toujours à l'aide de P3 on obtient :

$$\widetilde{\omega} = u^{p-1}v^{q-1}\left(A(u,v)vdu + B(u,v)udv\right) \quad \text{avec} \quad A(0,0),\;\; B(0,0) \neq 0,$$

$$\mathcal{X}_{S'_\omega,c} = \widetilde{\mathcal{O}}_c u \frac{\partial}{\partial u} \oplus \widetilde{\mathcal{O}}_c v \frac{\partial}{\partial v}.$$

Dans chaque cas on a bien : $\widetilde{\omega} \cdot \mathcal{X}_{S'_\omega,c} = (f \circ E_\omega)_c$. □

La semi-hyperbolicité est une propriété ouverte par déformation équiréductibles. En effet :

Proposition 3.3. — *Soit $(\eta_t)_t$ une déformation équiréductible d'un germe ω de 1-forme semi-hyperbolique à l'origine de $\mathbb{C}^2$. Alors pour tout t assez petit le germe η_t est aussi semi-hyperbolique.*

Démonstration. — En raisonnant par récurrence sur h_ω, on se ramène aisément au cas où ω est réduit. Par continuité des valeurs propres de la partie linéaire de η_t la proposition est alors triviale. □

Les propriétés P1 et P3 peuvent aussi s'exprimer conjointement en terme d'arbre dual.

Définition 3.4. — On appelle *arbre dual associé à ω* le graphe pondéré et fléché $\mathbb{A}^*(\omega)$ construit de la manière suivante :

- à chaque composante irréductible de D_ω on associe biunivoquement un sommet de $\mathbb{A}^*(\omega)$. Deux sommets de $\mathbb{A}^*(\omega)$ sont joints par une arête ssi les composantes de D_ω correspondantes s'intersectent ;
- on munit un sommet de $\mathbb{A}^*(\omega)$ d'une flèche pour chaque point singulier de $\widetilde{\mathcal{F}}_\omega$ qui n'est pas un point singulier de $\mathcal{D}_\omega$;
- les sommets correspondant à une composante dicritique (*i.e.* non-invariante) de $\widetilde{\mathcal{F}}_\omega$ portent une double flèche ;
- chaque sommet est pondéré par l'auto-intersection de la composante de D_ω correspondante.

On aurait pu aussi pondérer $\mathbb{A}^*(\omega)$ *par multiplicités* : le sommet correspondant à une composante D est affecté du poids $m_\omega(D)$. Lorsque ω est la différentielle d'une fonction à singularités isolées il est bien connu que ces deux pondérations sont équivalentes : on passe biunivoquement de l'une à l'autre par une correspondance affine **[8]**. Grâce à la propriété P3 il en est de même lorsque ω est semi-hyperbolique. Les propriétes P1 et P3 des formes semi-hyperboliques se résument ainsi :

P4. Les arbres duaux de ω et de df munis de leurs deux pondérations : auto-intersections et multiplicités, sont égaux, f désignant toujours une équation réduite de $\mathrm{Sep}(\omega)$.

Lorsque ω n'est plus semi-hyperbolique cette propriété n'est pas toujours vérifiée, cela dépend de la transversalité des variétés fortes des singularités de type selle-nœud de $\mathcal{F}_\omega$ aux branches du diviseur exceptionnel. Dans **[12]** **[13]** nous caractérisons et classifions formellement les 1-formes qui satisfont P4. Cependant le théorème 1 de **[5]** permet facilement de montrer la propriété suivante :

P5. Lorsque ω est semi-hyperbolique, les multiplicités $m_\omega(D)$ ne dépendent que de l'arbre dual $\mathbb{A}^*(\omega)$ pondéré par auto-intersections. De plus, dans l'ensemble $\Lambda^1(\mathbb{A}^*)$ des germes de 1-formes ayant toutes le même arbre dual $\mathbb{A}^*$ (pondéré avec les auto-intersections), les formes semi-hyperboliques réalisent pour chaque sommet le minimum des pondérations par multiplicités.

On en déduit le

Théorème 3.5. — *Une déformation η d'un germe en $0 \in \mathbb{C}^2$ de 1-forme semi-hyperbolique est équiréductible si et seulement si, pour t assez petit, on a les égalités des arbres duaux pondérés par auto-intersection :* $\mathbb{A}^*(\eta_t) \equiv \mathbb{A}^*(\omega)$.

Remarquons que dans l'hypothèse de ce théorème, on ne suppose pas l'égalité $\mathrm{Sing}(\eta) = 0 \times \mathbb{C}^p, 0$ mais seulement l'inclusion $\mathrm{Sing}(\eta) \subset 0 \times \mathbb{C}^p, 0$.

Démonstration. — Un sens de cette équivalence est trivial, montrons la suffisance de la condition d'égalité des arbres.

Il est clair que chaque η_t est non-dicritique. On en déduit en particulier que l'équiréductibilité de η est triviale lorsque ω est réduite, car ω ne peut bifurquer en un selle-nœud. Dans ce cas l'équiréduction consiste à ne pas éclater. La démonstration se fait par induction de la manière suivante. Définissons d'abord par récurrence la notion de *n-équiréductibilité* :

- Nous disons que η est *0-équiréductible* si ω est réduite et η est équiréductible : chaque η_t est réduit.
- Nous disons que η est *1-équiréductible* si η n'est pas 0-équiréductible et si après l'éclatement $E^1 : M^1 \longrightarrow \mathbb{C}^{2+p}$ de $C^1 := 0 \times \mathbb{C}^p$ l'ensemble suivant
$$\Sigma^1 := \Sigma(\eta^1; D^1) \cap D^1 \quad \text{avec} \quad \eta^1 := {E^1}^*(\eta), \quad D^1 := (E^1)^{-1}(C^1)$$
est lisse et étale[(1)] au dessus de P.
- Nous disons que η est *n-équiréductible*, $2 \leq n \leq h_\omega$, si η est 1-équiréductible et au voisinage de chaque branche de Σ^1 le saturé de η_1 est k-équiréductible avec $k \leq (n-1)$.

Il suffit de prouver, sous l'hypothèse du théorème, que :

(1) η est 1-équiréductible,
(2) si η est n-équiréductible, alors η est aussi $(n+1)$-équiréductible, $n \leq h_\omega$;

L'équiréduction se construira alors en éclatant successivement les lieux singuliers stricts contenus dans le diviseur exceptionnel le long desquels le transformé saturé de η n'est pas 0-équiréductible.

Montrons 1. Notons comme d'habitude η^1_t, D^1_t, π^1_t les restrictions de η^1, D^1 et π^1 à $M^1_t := (\pi^1)^{-1}(t)$, avec : $\pi : \mathbb{C}^{2+p} \longrightarrow \mathbb{C}^p$ désignant la projection linéaire et $\pi^1 := \pi \circ E^1$. La multiplicité $\nu(t)$ de η_t à l'origine est égale à $m_{\eta_t}(D^1_t)$. La semicontinuité de $\nu(t)$ et la propriété de minimalité de P5 appliquée à D^1_t entrainent que $\nu(t)$ est constant. Ainsi la restriction de π^1 à Σ^1_t est à fibres $\Sigma^1_{\eta_t}$ finies. Le nombre d'éléments de $\Sigma^1_{\eta_t}$ est constant égal au nombre de flèches et d'arêtes portées par le sommet de $\mathbb{A}^*(\eta_t)$ correspondant à D^1_ω. Il est constant puisque $\mathbb{A}^*(\eta_t) = \mathbb{A}^*(\omega)$. On en déduit que Σ^1 est étale via π^1.

Montrons 2. La n-équiréductibilité de ω permet de construire n applications d'éclatements $E^i : M^i \longrightarrow M^{i-1}$ de la manière suivante : E^1 est l'éclatement de $C^1 = 0 \times \mathbb{C}^p$; Notons
$$E_i := E^1 \circ \cdots \circ E^{i-1}, \quad \pi^i := \pi \circ E_i, \quad D^i := {E_i}^{-1}(C^1),$$
$$\eta^i := {E_i}^*(\eta), \quad \Sigma^i := \Sigma(\eta^i, D^i) \cap D^i,$$
et définissons par induction E^i comme l'éclatement de centre l'union $C^i \subset M^{i-1}$ des branches de Σ^i où η^i n'est pas 0-équiréductible. On sait de plus que chaque Σ^i est lisse et étale au dessus de $\mathbb{C}^p$ via π^i. Montrons que η^n est 1-équiréductible le long de chaque branche C^n_r $r = 1, \ldots, k$ de $C^n \subset M^n$. Pour cela paramétrisons chaque C^n_r

(1) Dans cette définition il n'est pas supposé que $\Sigma(\eta^1; D^1)$ est contenu dans D^1.

par une section $c_r(t)$ de π^n. La multiplicité $\nu_r(t)$ au point $c_r(t)$ de la restriction η_t^n de η^n à $M_t^n := (\pi^n)^{-1}(t)$ se décompose en

$$\nu_r(t) = \widetilde{\nu}_r(t) + \sum_j m_{\eta_{t,j}}(D_t^n(r))$$

où $D_{t,j}^n(r)$ désigne les composantes irréductibles de $D_t^n := D^n \cap M_t^n$ qui passent par le point $c_r(t)$ et $\widetilde{\nu}_r(t)$ désigne la multiplicité du saturé local de η_t^n au point $c_r(t)$. Eclatons l'ensemble C^{n+1} des branches de Σ^n où η^n n'est pas 0-équiréductible, $E^{n+1} : M^{n+1} \longrightarrow M^n$. D'après le lemme suivant

Lemme 3.6. — *Les fonctions $t \longmapsto \widetilde{\nu}_r(t)$ sont constantes, $r = 1, \ldots, k$.*

l'ensemble $\Sigma^{n+1} := \Sigma\left(\eta^{n+1}; D^{n+1}\right)$ avec

$$\eta^{n+1} := \left(E^{n+1}\right)^*(\eta^n) \qquad D^{n+1} := \left(E^{n+1}\right)^{-1}(D^n)$$

est fini au-dessus de $\mathbb{C}^p$ via $\pi^{n+1} := \pi^n \circ E^{n+1}$. Pour conclure il suffit de voir qu'il est étale ou, ce qui revient au même, que le nombre $s^n(t)$ de points de sa fibre au dessus de $t \in \mathbb{C}^p$ est indépendant de t. Mais $s^n(t)$ est la somme des nombres de flèches et d'arêtes portées par les sommets de $\mathbb{A}^*(\eta_t)$ correspondant aux diviseurs créés par les éclatements des points $c_r(t)$, $r = 1, \ldots, k$. Ces sommets se détectent sur $\mathbb{A}^*(\eta_t)$ à partir de la pondération par auto-intersection. D'où la conclusion. □

Démonstration du lemme (3.6). — Supposons que $\widetilde{\nu}_1(0) \leq \cdots \leq \widetilde{\nu}_k(0)$ et désignons par $\mathbb{A}_r^*(t)$ l'arbre dual (pondéré par auto-intersections) du saturé de η_t^n au point $c_r(t)$. Fixons pour chaque t une bijection

$$\rho_t : \{1, \ldots, k\} \longrightarrow \{1, \ldots, k\}$$

telle que $\mathbb{A}_r^*(t) = \mathbb{A}_{\rho_r(t)}^*(0) \subset \mathbb{A}^*(\omega)$; De telles applications existent puisque l'arbre dual de η_t est constant et que pour chaque t la suite $\left(E_t^i\right)_{i=1,\ldots,n}$ correspond à la suite $\left(E_{\eta_t}^i\right)_{i=1,\ldots,n}$ extraite de la réduction de η_t ; On sait que cette suite se détecte dans $\mathbb{A}^*(\eta_t)$ à partir des auto-intersections.

Les multiplicités $m_{\eta_t}(D_t^n(r))$ sont constantes puisque Σ^n est fini au-dessus de $\mathbb{C}^p$. D'autre part $\nu_r(t)$ est la multiplicité du diviseur créé par l'éclatement de $c_r(t)$ dans M_t^n. Appliquons pour $t = 0$ la propriété P5 de minimalité des multiplicités de $\mathbb{A}^*(\omega)$ puis la semicontinuité inférieure des $\widetilde{\nu}_r(t)$. Il vient :

$$\widetilde{\nu}_{\rho_t(r)}(0) \leq \widetilde{\nu}_r(t) \leq \widetilde{\nu}_r(0) \qquad r = 1, \ldots, k. \tag{4}$$

En sommant ces inégalités on obtient $\sum_r \widetilde{\nu}_r(0) = \sum_r \widetilde{\nu}_r(t)$. La semi-continuité des $\widetilde{\nu}_r(t)$ permet de conclure. □

4. Déformations équisingulières de germes de courbes planes

Classiquement une *déformation de base* $\mathbb{C}^p, 0$ *d'un germe de courbe* analytique $X_0 \subset \mathbb{C}^2, 0$ est la donnée (X, π) d'un germe d'hypersurface $X \subset \mathbb{C}^2 \times \mathbb{C}^p, 0$ qui intersecte $\mathbb{C}^2 \times 0$ suivant X_0, contient $0 \times \mathbb{C}^p, 0$ et de la projection canonique

$$\pi : \mathbb{C}^2 \times \mathbb{C}^p, 0 \longrightarrow \mathbb{C}^p.$$

On associe ainsi à (X, π) le «germe de famille analytique» de germes en $(0, t)$ de courbe plane $X_t := \pi^{-1}(t) \cap X$. Nous noterons souvent la déformation (X, π) par $(X_t)_{t \in \mathbb{C}^p, 0}$, ou plus simplement par $(X_t)_t$. Tout germe d'application holomorphe λ de $\mathbb{C}^q, 0$ dans $\mathbb{C}^p, 0$ permet de construire la *déformation* $\left(X_{\lambda(s)}\right)_s$ *obtenue par à partir de* $(X_t)_t$ *par le changement des paramètres* λ ; elle est définie par l'hypersurface de $\mathbb{C}^{2+q}, 0$ d'équation $F(x, y; \lambda(s)) = 0$, où $F(x, y; t) = 0$ est une équation réduite de X.

Deux déformations (X, π) et (Y, π) de X_0 de même base $\mathbb{C}^p, 0$ sont dites *topologiquement équivalentes*, resp. *analytiquement équivalentes*, s'il existe un germe d'homéomorphisme, resp. de biholomorphisme de $\mathbb{C}^{2+p}, 0$ qui commute avec π, vaut l'identité sur $\mathbb{C}^2 \times 0$ et transforme X en Y. Une déformation équivalente à la *déformation constante* $(X_0 \times \mathbb{C}^p, \pi)$ est dite *triviale.*

Lorsque la restriction à $\mathbb{C}^2 \times 0$ d'une équation réduite de X est une équation réduite de X_0 on dit que la déformation est *plate.* Cela signifie que *l'ensemble critique* $\mathrm{Crit}(X, \pi)$, constitué des points singuliers de X et des points critiques de la restriction de π à la partie lisse de X, est fini au-dessus de $\mathbb{C}^p, 0$ via π.

On retrouve la notion classique d'équiréductibilité **[18]** **[19]** en posant :

Définition 4.1. — Une déformation (X, π) d'un germe de courbe plane X_0 est dite *équiréductible* si elle est plate et si, $F(x, y; t)$ étant une équation réduite de X, la 1-forme $\eta_F := \dfrac{\partial F}{\partial x} dx + \dfrac{\partial F}{\partial y} dy$ définit une déformation équiréductible de df, avec $f(x, y) := F(x, y; 0)$.

Les résultats classiques de Zariski et Lê Dung Trang - C. P. Ramanujam **[18]** **[19]** **[9]** affirment que les propriétés suivantes, relatives à une déformation (X, π) de X_0 de base $\mathbb{C}^p, 0$, sont équivalentes :

(1) (X, π) est équiréductible,
(2) (X, π) est topologiquement triviale,
(3) $\mathrm{Crit}(X, \pi) = 0 \times \mathbb{C}^p$,
(4) $\mu_{(o,t)}(X_t) = \mu_0(X_0)$ pour t assez petit,

où $\mu_{(o,t)}(X_t)$ désigne le nombre de Milnor d'une équation réduite de X_t. Lorsque ces conditions sont satisfaites la déformation est dite *équisingulière.* L'existence d'un espace semi-universel pour ces déformations est bien connue. Nous allons en donner ici une construction géométrique qui nous sera utile par la suite.

Introduisons d'abord la notion de vitesse de déformation. Pour cela fixons une déformation équiréductible (X, π), $X = F^{-1}(0)$, de $X_0 = f^{-1}(0)$ de base $\mathbb{C}^p, 0$. L'image réciproque X' de X par l'application d'équiréduction $E_X := E_{\eta_F} : M_X \longrightarrow \mathbb{C}^{2+p}, 0$ peut être considérée comme une «déformation» de $X'_0 := X' \cap {\pi_X}^{-1}(0)$, avec $\pi_X := \pi \circ E_X$. En fait, puisque X' est à croisements normaux, son germe en chaque point de $D_{X_0} := D_X \cap {\pi_X}^{-1}(0)$, $D_X := {E_X}^{-1}(0 \times \mathbb{C}^p)$, est une déformation analytiquement triviale du germe de X'_0. On construit ainsi un recouvrement $\mathcal{U} := (U_i)_i$ de D_{X_0} et, le long de chaque U_i, des germes d'applications de trivialisations :

$$\Psi_i : (M_X, U_i) \xrightarrow{\sim} (M_{X_0} \times \mathbb{C}^p, U_i \times 0)\,, \qquad \Psi_i(X', U_i) = X'_0 \times \mathbb{C}^p, U_i \times 0$$

qui valent l'identité en restriction à $M_{X_0} := {\pi_X}^{-1}(0)$. Il est facile de voir, en obtenant par exemple les Ψ_i par intégrations de champs de vecteurs, que l'on peut faire cette construction avec un *recouvrement $\mathcal{U}$ adapté* dans le sens suivant : ses ouverts sont de deux types

- la trace sur D_{X_0} de «polydisques» de M_{X_0}

$$P(m; \varepsilon) := \{\mid u_m \mid < \varepsilon, \mid v_m \mid < \varepsilon\}\,, \quad m \in \operatorname{Sing}(X'_0),$$

  où u_m, v_m sont des coordonnées locales en m telles que $u_m v_m$ définit X'_0,
- les composantes connexes de $D_{X_0} - \bigcup_m P(m; \varepsilon/2)$.

L'intérêt d'un tel recouvrement est de ne pas avoir d'intersection d'ouverts trois à trois non-triviale.

La 1-cocycle défini par la cochaine $\Phi_{ij} := \Psi_i \circ \Psi_j^{-1}$ est à valeurs dans le faisceau de base $D_{X_0} \times 0 \subset D_{X_0} \times \mathbb{C}^p$, noté $\operatorname{Aut}_{\mathbb{C}^p}(M_{X_0}; X'_0)$, des germes aux points de D_{X_0} d'automorphismes analytiques de $M_{X_0} \times \mathbb{C}^p$ valant l'identité au dessus de 0, commutent avec $\pi_{X_0 \times \mathbb{C}^p}$ et laissent $X'_0 \times \mathbb{C}^p$ invariant. Il rend compte de la non-trivialité globale de X' — et donc de X. Ainsi les *«vitesses initiales de déformation»* de X sont les cocycles

$$\left[\frac{\partial X}{\partial t_k}\right]_{t=0} := \left[\left.\frac{\partial \Phi_{ij}}{\partial t_k}\right|_{t=0}\right] \in H^1\left(D_{X_0}; \mathcal{X}_{X'_0}\right)$$

où $\mathcal{X}_{X'_0}$ désigne, comme au paragraphe précédent, faisceau de base D_{X_0} des germes de champs de vecteurs sur M_{X_0} qui sont tangents à X'_0 en chaque point régulier. D'après un théorème [**1**] d'Andreotti-Grauert, $H^1\left(D_{X_0};\ X_{X'_0}\right)$ est un $\mathbb{C}$-espace vectoriel de dimension finie car D_{X_0} est exceptionnel et $\mathcal{X}_{X'_0}$ est un $\mathcal{O}_{M_{X_0}}$-module cohérent.

Théorème 4.2. — *Il existe une déformation équisingulière $\left(X^U_u\right)_u$ de X_0 de base $\mathbb{C}^{\tau(X_0)}, 0$ avec $\tau(X_0) := \dim_{\mathbb{C}} H^1\left(D_{X_0}; \mathcal{X}_{X'_0}\right)$, qui est semi-universelle dans le sens suivant : toute déformation équisingulière $(X_t)_{t \in \mathbb{C}^p, 0}$ de X_0 est analytiquement équivalente à une déformation $\left(X^U_{\lambda(t)}\right)_t$ obtenue à partir de $\left(X^U_u\right)_u$ par un changement de paramètres $\lambda : \mathbb{C}^p, 0 :\longrightarrow \mathbb{C}^{\tau(X_0)}, 0$ dont la dérivée en 0 est unique. De plus on a* *une identification canonique*

$$\sigma_{X_0} : T_0\mathbb{C}^{\tau(X_0)} \xrightarrow{\sim} H^1\left(D_{X_0}; \mathcal{X}_{X_0'}\right)$$

et via cette identification la dérivée $T_0\lambda$ de λ à l'origine s'écrit

$$(5) \qquad \sum_{k=1}^{p} \alpha_k \left.\frac{\partial}{\partial t_k}\right|_{u=0} \longmapsto \sum_{k=1}^{p} \alpha_k \left[\frac{\partial X}{\partial t_k}\right]_{t=0}.$$

Remarque 4.3. — Pour $(X_t)_t = \left(X_u^U\right)_u$ la formule (5) définit l'identification σ_{X_0}.

Démonstration. — Elle se fait par les mêmes techniques que celles utilisées dans **[10]** pour la construction d'un déploiement universel de ω. Aussi nous n'en donnons que les grandes lignes. Notons $\tau := \tau(X_0)$.

Etape 1. — Le recouvrement $\mathcal{U}$ étant adapté, on montre par une technique simple de recollements que tout cocycle $[\Phi_{ij}]$ à valeur dans $\mathrm{Aut}_{\mathbb{C}^p}\left(M_{X_0}; X_0'\right)$ est le cocycle associé à une déformation équiréductible de X_0 de base $\mathbb{C}^q, 0$.

Etape 2. — Donnés des éléments $\left[V_{ij}^1\right], \dots, \left[V_{ij}^q\right]$ de $H^1\left(\mathcal{U}; \mathcal{X}_{X_0'}\right)$ on montre qu'il existe une déformation équiréductible $(X, \pi_{\mathbb{C}^q})$ de X_0 telle que $\left[\frac{\partial X}{\partial t_k}\right]_{t=0} = \left[V_{ij}^k\right]$, $k = 1\dots, q$. Il suffit pour cela d'appliquer ce qui précède au composé des flots ϕ_{ij,t_k} des champs V_{ij}^k.

Etape 3. — On construit $\left(X_u^U\right)_u$ grace à l'étape précédente, en imposant aux cocycles $\left[\frac{\partial X^U}{\partial u_k}\right]_{u=0}$, $k = 1, \dots, \tau$ de former une base de $H^1\left(\mathcal{U}; \mathcal{X}_{X_0'}\right)$.

Etape 4. — Donnée une déformation équiréductible $(Y_t)_t$ de X_0 de base quelconque $\mathbb{C}^p, 0$, on construit grace à 1. une déformation équiréductible $(Z_{u,t})_{u,t}$ de base $\mathbb{C}^\tau \times \mathbb{C}^p, 0$ telle que $(Z_{u,0})_u = \left(X_u^U\right)_u$ et $(Z_{0,t})_t = (Y_t)_t$.

Etape 5. — On construit «l'application de Kodaira-Spencer» de $(Z_{u,t})_{u,t}$. Pour cela on considère un représentant $\underline{X_0}$ du germe X_0 sur une boule B^2 et un representant $\underline{Z}$ de la déformation Z construite à l'étape précédente sur un polydisque $B^2 \times B$, les boules $B^2 \subset \mathbb{C}^2$ et $B \subset \mathbb{C}^{\tau+p}$ étant de rayons assez petits. Soit $E_{\underline{Z}} : M_{\underline{Z}} \longrightarrow B^2 \times B$ l'application d'équiréduction de $\underline{Z}$; notons $\underline{Z}' := E_{\underline{Z}}{}^{-1}(Z)$. Considérons sur $M_{\underline{Z}}$ le faisceau $\mathcal{X}^v_{\underline{Z}'}$ des germes de champs de vecteurs de $M_{\underline{Z}}$ qui sont tangents à $\underline{Z}'$ et verticaux pour la projection $\pi' := \pi \circ E_{\underline{Z}}$, *i.e.* $T\pi' \cdot Z \equiv 0$, où $\pi : B^2 \times B \longrightarrow B$ désigne la projection canonique. D'après un théorème classique de Grauert le faisceau de base $B^2 \times B$

$$R^1(E_{\underline{Z}*})\left(\mathcal{X}^v_{\underline{Z}'}\right) : W \longmapsto H^1\left(E_{\underline{Z}}^{-1}(W); \mathcal{X}^v_{\underline{Z}'}\right)$$

est cohérent sur $\mathcal{O}_{B^2\times B}$. Il est visiblement à support dans $0 \times B$. Ainsi

$$R^1({\pi'}_*)\left(\mathcal{X}^v_{\underline{Z}'}\right) : W \longmapsto H^1\left({\pi'}^{-1}(W); \mathcal{X}^v_{\underline{Z}'}\right)$$

est un faisceau cohérent sur $\mathcal{O}_B$. D'après Andreotti-Grauert on a aussi

$$R^1(\pi'_*)\left(\mathcal{X}^v_{\underline{Z}'}\right)(W) = H^1\left(\pi'^{-1}(W)\cap D_{\underline{Z}};\mathcal{X}^v_{\underline{Z}'}\right), \quad D_{\underline{Z}} := E_{\underline{Z}}^{-1}(0\times B).$$

Visiblement tout champ de vecteur V dans B se relève sur les ouverts U_j d'un recouvrement de $M_{\underline{Z}}$ en des sections champs V_j qui sont tangents à $\underline{Z}'$. *L'application de Kodaira-Spencer de* $(Z_{u,t})_{u,t}$ est le morphisme de $\mathcal{O}_{\mathbb{C}^{2+\tau},0}$-modules de types finis

$$(6) \qquad \left[\frac{\partial Z}{\partial(u,t)}\right] : \mathcal{X}_{\mathbb{C}^{2+\tau},0} \longrightarrow R^1(\pi'_*)\left(\mathcal{X}^v_{\underline{Z}'}\right)_0, \quad V \longmapsto [V_j - V_i].$$

Son noyau $\mathcal{N}$ est constitué des champs V qui admettent un relevé $\widetilde{V}$ global. Ainsi, par intégration des flots de V et de $\widetilde{V}$, la déformation $(Z_{u,t})_{u,t}$ est analytiquement triviale lorsque (u,t) varie sur une orbite de V.

Etape 6. — On construit la factorisation λ. Remarquons d'abord que (6) est surjective. En effet considérons $\mathcal{X}^v_{\underline{Z}'}$ comme un faisceau de modules sur l'anneau $A := \mathcal{O}_B(B)$. Soit $\mathfrak{M} \subset A$ l'idéal des fonctions qui s'annulent à l'origine. A l'aide d'un bon recouvrement de $D_{\underline{Z}}$ sans intersection 3 à 3 non-triviales, on montre facilement l'égalité :

$$(7) \qquad R^1(\pi'_*)\left(\mathcal{X}^v_{\underline{Z}'}\right)(B)\otimes_A A/\mathfrak{M} \simeq H^1\left(D_{\underline{Z}};\mathcal{X}^v_{\underline{Z}'}\otimes_A A/\mathfrak{M}\right) \simeq H^1\left(D_{X_0};\mathcal{X}_{X'_0}\right).$$

Ainsi, avec ces identifications $\sigma_Z := \left[\frac{\partial Z}{\partial(u,t)}\right]\otimes_A A/\mathfrak{M}$ est l'application

$$\sum_{k=1}^{\tau}\alpha_k \left.\frac{\partial}{\partial u_k}\right|_{(0,0)} + \sum_{k=1}^{p}\beta_k \left.\frac{\partial}{\partial t_k}\right|_{(0,0)} \longmapsto \sum_{k=1}^{\tau}\alpha_k\left[\frac{\partial Z}{\partial u_k}\right]_{u=0,t=0} + \sum_{k=1}^{p}\beta_k\left[\frac{\partial Z}{\partial t_k}\right]_{u=0,t=0}.$$

Cette application est visiblement surjective. Le lemme de Nakayama et l'exactitude à droite du produit tensoriel donnent le surjectivité de $\left[\frac{\partial Z}{\partial(u,t)}\right]$.

De manière plus précise, on a la suite exacte

$$0 \longrightarrow \mathcal{N}(0) \longrightarrow T_{(0,0)}\mathbb{C}^\tau \times \mathbb{C}^p \xrightarrow{\sigma_Z} H^1\left(D_{X_0};\mathcal{X}_{X'_0}\right) \longrightarrow 0,$$

où $\mathcal{N}(0)$ est le sous-espace de $T_0\mathbb{C}^{\tau+p}$ formé des valeurs[(2)] à l'origine des éléments de $\mathcal{N}$; La restriction de σ_Z à $T_0\mathbb{C}^\tau\times 0$ est un isomorphisme ; Ainsi $\mathcal{N}(0)$ est de dimension p et transverse à $T_0\mathbb{C}^\tau \times 0$.

Enfin on voit facilement que $\mathcal{N}$ est involutif, *i.e.* $[\mathcal{N},\mathcal{N}]\subset\mathcal{N}$. D'après un résultat de D. Cerveau **[6]** page 266, $\mathcal{N}$ contient un sous-module libre involutif $\mathcal{T}$ de rang p, tel que $\mathcal{T}(0) = \mathcal{N}(0)$. Le feuilletage régulier $\mathcal{F}_{\mathcal{T}}$ définit par $\mathcal{T}$ est transverse à $\mathbb{C}^\tau\times 0$. Lorsque le paramètre (u,v) varie sur une feuille de $\mathcal{F}_{\mathcal{T}}$ la déformation $(Z_{u,t})_{u,t}$ est triviale. On obtient le changement de paramètres désiré en prenant le germe de l'application $\lambda : 0\times\mathbb{C}^p,(0,0) \longrightarrow \mathbb{C}^\tau\times 0,(0,0)$ qui consiste à suivre les feuilles de $\mathcal{F}_{\mathcal{T}}$. □

(2) En général $\mathcal{N}(0)$ n'est pas égal à $\mathcal{N}\otimes_{\mathcal{O}_{\mathbb{C}^{\tau+p},0}}\mathcal{O}_{\mathbb{C}^{2+p},0}/(u,t)$.

Remarque 4.4. — On voit facilement à partir de cette construction que la factorisation λ est nécéssairement identiquement nulle lorsque $(X_t)_t$ est triviale.

5. Familles de séparatrices

Les déformations de courbes qui nous intéressent ici sont les familles de séparatrices d'une déformation de 1-forme. A priori cette notion n'est pas bien définie et la dépendance analytique d'une famille de séparatrices demande à être précisée.

Définition 5.1. — Soit η une déformation de base quelconque $\mathbb{C}^p, 0$ d'un germe de 1-forme semi-hyperbolique ω à l'origine de $\mathbb{C}^2$ et (X, π) une déformation de base $\mathbb{C}^p, 0$ de $X_0 := \mathrm{Sep}(\omega)$. Nous disons que (X, π) est une *famille de séparatrices de* η si pour toute valeur t assez petite du paramètre le germe de 1-forme η_t en $(0, t) \in \mathbb{C}^2 \times t$ est non-dicritique et admet X_t comme courbe invariante (non nécéssairement irréductible).

L'existence de famille de séparatrices permettra d'obtenir de bons critères d'équiréductibilité. Cependant je ne sais pas répondre aux questions suivantes :

(1) Est-ce qu'une déformation η d'une 1-forme semi-hyperbolique η_0, de lieu singulier $\mathrm{Sing}(\eta) = 0 \times \mathbb{C}^p, 0$ admet nécéssairement une famille de séparatrices (X, π) telle que $X_t = \mathrm{Sep}(\eta_t)$?

(2) Même question en supposant de plus que chaque η_t est semi-hyperbolique.

Une réponse affirmative impliquerait d'après ce qui suit que η est équiréductible. Un critère d'équiréductibilité dans ce cadre serait la constance du nombre de Milnor de η_t et cela étendrait à ces feuilletages les théorèmes de Zariski et Lê Dung Trang-Ramanujam. Cela donnerait aussi, lorsque η est la différentielle d'une fonction F, une démonstration de l'implication « $\mu(F_t)$ constant $\Rightarrow m_0(F_t)$ constant » qui n'utiliserait pas le critère topologique de Lê Dung Trang-Ramanujam.

De la même manière que pour l'ouverture de la semi-hyperbolicité on montre :

Proposition 5.2. — *Toute déformation équiréductible η d'un germe de 1-forme semi-hyperbolique ω possède une famille de séparatrices (X, π) qui est une déformation équiréductible de $\mathrm{Sep}(\omega)$ et qui satistait : $X_t = \mathrm{Sep}(\eta_t)$.*

Une « réciproque » à cette proposition est :

Proposition 5.3. — *Soit η une déformation de base quelconque $\mathbb{C}^p, 0$ d'un germe de 1-forme semi-hyperbolique ω qui possède une famille de séparatrices (X, π). Les assertions suivantes sont alors équivalentes :*

(1) *η est équiréductible,*
(2) $\mathrm{Sing}(\eta) = 0 \times \mathbb{C}^p, 0$
(3) $\mu(\eta_t) = \mu(\omega)$,
(4) *(X, π) est équisingulière,*

Démonstration. — L'implication (1) $\Rightarrow$ (2) et l'équivalence (2) $\Leftrightarrow$ (3) sont triviales. Comme $\mathrm{Crit}(X)$ est contenu dans $\mathrm{Sing}(\eta)$, l'implication (3) $\Rightarrow$ (4) résulte de la propriété P2 des formes semi-hyperboliques énoncée au §3 et du critère (4) d'équisingularité donné au chapitre 4.

Supposons (X, π) équisingulière. D'après P2 on a les inégalités :

$$\mu_{(o,t)}(\eta_t) \geq \mu_{(o,t)}\left(\mathrm{Sep}(\eta_t)\right) \geq \mu_{(o,t)}(X_t) = \mu_0\left(\mathrm{Sep}(\omega)\right) = \mu_0(\omega). \tag{8}$$

Par semicontinuité $\mu_0(\omega) \geq \mu_{(o,t)}(\eta_t)$ et, dans (8), on a partout égalité. Ainsi chaque η_t est semi-hyperbolique et $X_t = \mathrm{Sep}(\eta_t)$. La propriété P1 du chapitre 3 permet de conclure à l'équiréduction de η. □

Ces conditions équivalentes n'impliquent pas, bien sûr, la trivialité topologique de la déformation ; c'est déjà faux lorsque ω est réduit. On a cependant :

Corollaire 5.4. — *Soit η une déformation topologiquement triviale d'un germe de 1-forme ω semi-hyperbolique à l'origine de $\mathbb{C}^2$. Alors η est équiréductible.*

Démonstration. — La difficulté vient du fait qu'on ne sait pas à priori que la collection des germes $\mathrm{Sep}(\eta_t)$ est contenue dans une hypersurface de $\mathbb{C}^{2+p}$. Il est clair que chaque η_t est non-dicritique et $\mathrm{Sep}(\eta_t)$ est bien défini. Soit $\Phi := (\Phi_t; t)$ un homéomorphisme trivialisant η, défini sur voisinage assez petit de l'espace des paramètres $P := 0 \times \mathbb{C}^p, 0$ pour que $\mathrm{Sing}(\eta) = P$. Il est clair que $\mu_{(0,t)}(\eta_t)$ est constant. L'image de $\mathrm{Sep}(\omega)$ par Φ_t est une courbe analytique de $\mathbb{C}^2 \times 0$ invariante par η_t. Le même raisonnement avec Φ_t^{-1} prouve que $\Phi_t(\mathrm{Sep}(\omega)) = \mathrm{Sep}(\eta_t)$.

Montrons que chaque η_t est semi-hyperbolique. L'invariance topologique du nombre de Milnor d'un germe de courbe, la propriété P2, la semi-continuité inférieure de $t \longmapsto \mu_{(0,t)}(\eta_t)$ et ce qui précède donnent les relations

$$\mu_0(\mathrm{Sep}(\omega)) = \mu_{(0,t)}(\mathrm{Sep}(\eta_t)) \leq \mu_{(0,t)}(\eta_t) = \mu_0(\omega).$$

La semi-hyperbolicité de ω impliquant l'égalité de $\mu_0(\omega)$ et $\mu_0(\mathrm{Sep}(\omega))$, la ligne précédente ne comporte que des égalités. On conclut de nouveau par P2.

D'après la propriété P4 des 1-formes semi-hyperboliques on a l'égalité des arbres duaux pondérés $\mathbb{A}^*(\eta_t) = \mathbb{A}^*(\mathrm{Sep}(\eta_t))$. On sait d'autre part que l'arbre dual pondéré de réduction d'un germe de courbe plane est un invariant topologique de ce germe. En effet les paires de Puiseux sont des invariants topologique **[20]** et elles donnent, *cf.* **[8]** la réduction des singularités. On conclut par (3.5) □

6. Notions de quasi-homogénéité

Considérons d'abord la quasi-homogénéité au sens des déploiements équisinguliers. Nous conservons les notations des chapitres précédents.

Définition 6.1. — Soit $\omega = a(x,y)\,dx + b(x,y)\,dy$ un germe à l'origine de $\mathbb{C}^2$ de 1-forme semi-hyperbolique. Un *déploiement équisingulier de ω de base $\mathbb{C}^p$* est un germe de 1-forme à l'origine de $\mathbb{C}^{2+p}$

$$\Omega := A(x,y;\,t)\,dx + B(x,y;\,t)\,dy + \sum_{j=1}^{p} C_j(x,y;t)dt_j$$

qui satisfait la condition d'intégrabilité $\Omega \wedge d\Omega \equiv 0$ et tel que :

(1) les germes à l'origine de $\mathbb{C}^{2+p}$ des sous-ensembles

$$\{A = B = 0\} \quad \text{et} \quad \{A = B = C_1 = \cdots = C_p = 0\}$$

sont identiques et égaux à $0 \times \mathbb{C}^p, 0$.

(2) la déformation de ω définie par

$$\eta := A(x,y;\,t)\,dx + B(x,y;\,t)\,dy$$

est équiréductible ; Nous la notons $(\Omega_t)_t$;

(3) le diviseur exceptionnel D_Ω de l'équiréduction de η, que nous notons ici $E_\Omega : M_\Omega \longrightarrow \mathbb{C}^{2+p}$, est une hypersurface invariante du feuilletage singulier saturé de codimension un $\widetilde{\mathcal{F}_\Omega}$ défini par les saturés locaux de $\widetilde{\Omega} := E_\Omega^{-1}(\Omega)$. De plus $\mathrm{Sing}(\widetilde{\mathcal{F}}_\Omega) = \mathrm{Sing}(\widetilde{\mathcal{F}}_\eta)$ et $\Sigma(\widetilde{\mathcal{F}}_\Omega; D_\Omega) = \Sigma(\widetilde{\eta}; D_\Omega)$.

Un tel déploiement définit sur $\mathbb{C}^{2+p}, 0$ un feuilletage $\mathcal{F}_\Omega$ de codimension 1 de lieu singulier $0 \times \mathbb{C}^p, 0$ et transverse en dehors de ce lieu singulier à la projection linéaire $\pi : \mathbb{C}^2 \times \mathbb{C}^p \longrightarrow \mathbb{C}^p$. Une *équivalence* de deux déploiements Ω^1 et Ω^2 est une équivalence des déformations associées qui en plus transforme $\mathcal{F}_{\Omega^1}$ en $\mathcal{F}_{\Omega^2}$. On définit de même le *transformé d'un déploiement par un changement de base* $\lambda := (\lambda_1, \ldots, \lambda_p) : \mathbb{C}^q, 0 \longrightarrow \mathbb{C}^p, 0$ comme la 1-forme $\lambda^\star\Omega$ sur $\mathbb{C}^{2+q}, 0$ obtenue en posant $t_j = \lambda_j(s)$ dans l'expression de Ω.

Un déploiement équisingulier Ω possède d'après (3.3) et (5.2) un *ensemble de séparatrices* : une hypersurface invariante $\mathrm{Sep}(\Omega) \subset \mathbb{C}^{2+p}, 0$ qui est une déformation équisingulière de $\mathrm{Sep}(\omega)$ et telle que $\mathrm{Sep}(\Omega) \cap \mathbb{C}^2 \times t = \mathrm{Sep}(\Omega_t)$.

Théorème 6.2 **([10]).** — *Tout germe ω de 1-forme à l'origine de $\mathbb{C}^2$ possède un déploiement équisingulier Ω^U, de base $\mathbb{C}^{\delta(\omega)}, 0$, qui est universel dans le sens suivant : tout déploiement équisingulier Ω de base quelconque $\mathbb{C}^p, 0$ est analytiquement équivalent à un déploiement du type $\lambda^\star\Omega^U$, avec $\lambda : \mathbb{C}^p \longrightarrow \mathbb{C}^{\delta(\omega)}$. De plus Ω^U est unique à unité multiplicative près et*

$$\delta(\omega) := \sum_c \frac{(\nu_c - 1)(\nu_c - 2)}{2}$$

où c décrit l'ensemble $\bigsqcup_{i=0,\ldots,h_\omega} C^i_\omega$ de tous les points que l'on éclate (0 compris) lors de la réduction de ω et ν_c désigne la multiplicité en $c \in C^i_\omega$ du saturé du transformé ω^i de ω, cf. *ch. 2. De plus on a une identification canonique*

$$\sigma_{\Omega^U} : T_0\mathbb{C}^{\delta(\omega)} \longrightarrow H^1(D_\omega; \mathcal{X}_{\widetilde{\mathcal{F}}_\omega}),$$

où $\mathcal{X}_{\widetilde{\mathcal{F}}_\omega}$ désigne le faisceau de base D_ω des germes (aux points de D_ω) de champs de vecteurs tangents au feuilletage $\widetilde{\mathcal{F}}_\omega$ et au diviseur D_ω.

Avant de prouver le théorème A de l'introduction précisons la notion de quasi-homogénéité au sens des déploiements.

Définition 6.3. — Un germe de 1-forme à l'origine de $\mathbb{C}^2$ est dit *d-quasihomogène* si tout déploiement équisingulier de base quelconque Ω satisfait l'équivalence :

- Ω est holomorphiquement trivial si et seulement si la déformation de courbes $(\mathrm{Sep}(\Omega_t))_t$ est holomorphiquement triviale.

Remarque 6.4. — Lorsque ω est exacte, $\omega = df$, tout déploiement de ω est donné par une 1-forme exacte **[11]**. Dans ce cas la définition ci-dessus est équivalente à la définition classique de quasi-homogénéité de f.

Démonstration du théorème A. — Notons $\delta := \delta(\omega)$ et $\tau := \tau(\omega)$. La déformation

$$(\mathrm{Sep}(\Omega^U); \pi_{\mathbb{C}^\delta}) = (\mathrm{Sep}(\Omega^U_v))_{v\in\mathbb{C}^\delta} \tag{9}$$

de $X_0 := \mathrm{Sep}(\omega)$ se factorise dans la déformation équisingulière semi-universelle $(X^U, \pi_{\mathbb{C}^\tau})$ de X_0. Donnons-nous un changement de base $\Lambda^U : \mathbb{C}^\delta \longrightarrow \mathbb{C}^\tau$ tel que (9) soit analytiquement conjuguée à $\left(X^U_{\Lambda^U(v)}\right)$. Nous allons calculer la dérivée à l'origine de Λ^U.

D'après le lemme (1.1.5) de **[10]** le feuilletage $\widetilde{\mathcal{F}}_{\Omega^U}$ est localement analytiquement trivial et l'on peut construire des germes de trivialisations

$$\Psi_i : M_{\Omega^U}, U_i \xrightarrow{\sim} M_\omega \times 0, U_i \times 0$$

qui sont définis au voisinage de chaque ouvert U_i d'un recouvrement adapté $\mathcal{U}$ (*cf.* ch. 4) de $D_\omega \subset M_\omega \hookrightarrow M_{\Omega^U}$. On dispose maintenant d'un cocycle

$$\left[\frac{\partial \mathcal{F}_{\Omega^U}}{\partial v_k}\right]_{v=0} := \left[\left.\frac{\partial \Psi_i \circ {\Psi_j}^{-1}}{\partial v_k}\right|_{v=0}\right] \in H^1\left(D_\omega; \mathcal{X}_{\widetilde{\mathcal{F}}_\omega}\right). \tag{10}$$

Il résulte de la construction de **[10]** que l'identification σ_{Ω^U} du théorème (6.2) s'explicite par :

$$\sigma_{\Omega^U}\left(\sum_k \alpha_k \left.\frac{\partial}{\partial v_k}\right|_{v=0}\right) = \sum_k \alpha_k \left[\frac{\partial \mathcal{F}_{\Omega^U}}{\partial v_k}\right]_{v=0}. \tag{11}$$

D'autre part chaque Ψ_i induit une trivialisation de $S' := {E_\Omega}^{-1}(\mathrm{Sep}(\Omega^U))$. Ainsi, $\mathcal{X}_{S'}$ désignant la restriction à D_ω du faisceau des champs de vecteurs sur M_ω qui sont tangents à S' en chaque point régulier, le cocycle

$$\left[\frac{\partial \mathrm{Sep}(\Omega^U)}{\partial v_k}\right]_{v=0} \in H^1(D_\omega; \mathcal{X}_{S'})$$

est l'image de (10) par l'application

$$\check{\rho} : H^1\left(D_\omega ; \mathcal{X}_{\widetilde{\mathcal{F}}_\omega}\right) \longrightarrow H^1\left(D_\omega ; \mathcal{X}_{S'}\right)$$

induite par l'inclusion de faisceaux $\rho : \mathcal{X}_{\widetilde{\mathcal{F}}_\omega} \longrightarrow \mathcal{X}_{S'}$. On déduit de (11) et (4.3) l'égalité suivante (modulo les identifications σ_{Ω^U} et σ_{X^U})

$$T_0\Lambda^U = \check{\rho}. \tag{12}$$

Grâce au lemme clé (3.2) on va pouvoir calculer $T_0\Lambda^U$. On considère la suite exacte

$$0 \longrightarrow \mathcal{X}_{\widetilde{\mathcal{F}}_\omega} \xrightarrow{\rho} \mathcal{X}_{S'} \xrightarrow{\widetilde{\omega}\cdot} (f \circ E_\omega) \longrightarrow 0 \tag{13}$$

où f désigne encore une équation réduite de $S := \mathrm{Sep}(\omega)$. La suite longue de cohomologie associée se décompose en

$$0 \longrightarrow H^0\left(D_\omega ; \mathcal{X}_{\widetilde{\mathcal{F}}_\omega}\right) \longrightarrow H^0\left(D_\omega ; \mathcal{X}_{S'}\right) \longrightarrow H^0\left(D_\omega ; (f \circ E_\omega)\right) \longrightarrow N \longrightarrow 0$$

$$0 \longrightarrow N \longrightarrow H^1\left(D_\omega ; \mathcal{X}_{\widetilde{\mathcal{F}}_\omega}\right) \longrightarrow H^1\left(D_\omega ; \mathcal{X}_{S'}\right) \longrightarrow H^1\left(D_\omega ; (f \circ E_\omega)\right) \longrightarrow 0$$

puisque le recouvrement $\mathcal{U}$ n'a pas d'intersection trois à trois non triviale. Le faisceau $(f \circ E_\omega)$ est isomorphe à $\mathcal{O}$ car engendré par une section triviale. Ainsi[(3)] $H^1\left(D_\omega ; (f \circ E_\omega)\right)$ est nul. On en déduit le

Lemme 6.5. — *Si ω est semi-hyperbolique, alors $\Lambda^U : \mathbb{C}^{\delta(\omega)}, 0 \longrightarrow \mathbb{C}^{\tau(\omega)}, 0$ est un germe de submersion.*

Les sections globales des faisceaux de (13) définissent des objets holomorphes sur une petite boule épointée centrée en $0 \in \mathbb{C}^2$; ils se prolongent à l'origine par le théorème classique d'Hartogs. Ainsi, N est le conoyau de l'application $\omega\cdot : V \longmapsto w \cdot V$ qui, à tout élément du module $\mathcal{X}_{S,0} \subset \mathcal{X}_{\mathbb{C}^2,0}$ des germes en 0 de champs de vecteurs de $\mathbb{C}^2, 0$ tangents à S, associe un élément de l'idéal (a, b) de $\mathcal{O}_{\mathbb{C}^2,0}$ engendré par f. Comme tout élément de (a, b) peut s'écrire $\omega \cdot V$ ce conoyau est égal au quotient $(f)/(a, b) \cap (f)$

En conclusion on obtient la suite exacte

$$0 \longrightarrow \frac{(f)}{(a,b) \cap (f)} \longrightarrow T_0\mathbb{C}^\delta \xrightarrow{T_0\Lambda^U} T_0\mathbb{C}^\tau \longrightarrow O \tag{14}$$

D'après la remarque (4.4) la sous variété $(\Lambda^U)^{-1}(0)$ est l'ensemble dans lequel se factorise nécéssairement toute déformation triviale de $\mathrm{Sep}(\omega)$. Par définition ω est d-quasi-homogène si et seulement si $(\Lambda^U)^{-1}(0) = \{0\}$. Puisque Λ^U est une submersion on obtient les équivalences des propriétés :

---

(3) On peut facilement voir que $H^1(D_\omega ; \mathcal{O}) = 0$: Par un simple calcul de séries de Laurent pour un seul éclatement, puis par induction à l'aide de la suite exacte de Mayer-Vietoris.

(i) ω est d-quasi-homogène,
(ii) $\dfrac{(f)}{(a,b)\cap(f)}=0$,
(iii) $\delta(\omega)=\tau(S)$.

Nous sommes maintenant en mesure de montrer les équivalences du théorème.

L'équivalence $(1')\Leftrightarrow(2)$ correspond à (i) $\Leftrightarrow$ (ii). En appliquant (i) $\Leftrightarrow$ (ii) à df on obtient $(3)\Leftrightarrow(4)$. Pour obtenir $(1')\Leftrightarrow(3)$ on applique successivement le critère (i) $\Leftrightarrow$ (iii) à ω et à df, puis on remarque que $\delta(\omega)=\delta(df)$. En effet grâce à la propriété P3 des formes semi-hyperboliques, les multiplicités $m_\omega(D)$ et $m_{df}(D)$ sont toutes égales ; il en est de même des ν_c qui s'en déduisent linéairement. Il reste à prouver $(2)\Leftrightarrow(5)$

Visiblement $\omega\cdot g^{-1}\left(\alpha u\frac{\partial}{\partial u}+\beta v\frac{\partial}{\partial v}\right)=f$; ce qui donne $(2)\Leftarrow(5)$.

Pour l'implication réciproque exprimons que $f^{-1}(0)$ est séparatrice de ω en écrivant

$$\omega\wedge df=fH dx\wedge dy,\qquad H\in\mathcal{O}_{\mathbb{C}^2,0}. \tag{15}$$

Par hypothèse on dipose d'un champ Z qui satisfait $\omega\cdot Z=f$. Ce champ est tangent à $f^{-1}(0)=\mathrm{Sep}(\omega)$ et donc $df\cdot Z$ peut s'écrire fR, $R\in\mathcal{O}_{\mathbb{C}^2,0}$. Le produit intérieur i_Z appliqué à chaque membre de (15) donne

$$df-R\omega=Hi_Z(dx\wedge dy). \tag{16}$$

D'après ce qui est déjà prouvé df est ausi quasi-homogène et on dispose d'un champ Z' qui vérifie $df\cdot Z'=f$. Ce champ est colinéaire à Z le long de la séparatrice commune $f^{-1}(0)$ et l'on a : $dx\wedge dy(Z,Z')=f\Delta'$, ainsi que $\omega\cdot Z'=fR'$, avec $\Delta',R'\in\mathcal{O}_{\mathbb{C}^2,0}$. En appliquant Z' à (16) on obtient

$$RR'=1-H\Delta'. \tag{17}$$

Remarquons que $RR'\neq 0$. En effet la multiplicité à l'origine $\nu_0(H)$ de H est minorée à l'aide de (15) par $\nu_0(\omega)-1$ et R est non nul dès que $\nu_0(\omega)>1$. D'autre part le cas $\nu(\omega)=1, H(0)\neq 0$ ne peut se produire : sinon la fonction f qui est quasi-homogène et de même multiplicité que ω, s'écrirait dans de bonnes coordonnées v^2+u^q ; l'équation (15) se résout alors immédiatement, les relations entre f'_u et f'_v étant toutes triviales ; finalement on obtiendrait

$$\frac{1}{H}\omega=Wdf+\frac{1}{2}vdu-\frac{1}{q}udv,$$

ce qui est exclu car ω n'est pas dicritique.

Pour achever la démonstration prenons de nouveau des coordonnées u,v dans lesquelles f est un polynôme quasi-homogène. Les calculs précédents peuvent alors se refaire avec $Z':=\alpha u\frac{\partial}{\partial u}+\beta v\frac{\partial}{\partial v}$ où α et β sont les poids des coefficients de f. Puisque

$R' := \frac{\omega \cdot Z'}{f}$ est non nul on peut aussi poser $Z := \frac{1}{R'} Z$. L'équation (16) devient

$$df - g\omega = h(\beta v du - \alpha u dv), \qquad g = \frac{f}{\omega \cdot Z'} \neq 0.$$

Ceci qui achève la démonstration. □

Remarque 6.6. — Tout déploiement équisingulier est topologiquement trivial **[10]** et visiblement le pseudo-groupe d'holonomie du feuilletage régulier d'un voisinage épointé $W - t$ de $0 \in \mathbb{C}^2$ est indépendant de t. On déduit alors immédiatement du lemme (6.5) le

Corollaire 6.7. — *Soit ω un germe de 1-forme semi-hyperbolique. Alors toute déformation topologiquement triviale (X, π) de* Sep(ω) *est réalisée comme la séparatrice d'un déploiement équisingulier,* i.e. *il existe un déploiement équisingulier Ω de ω tel que $X =$* Sep(Ω). *En particulier la déformation $(\Omega_t)_t$ de ω est topologiquement triviale, à pseudogroupe d'holonomie constant et pour t assez petit $X_t =$* Sep$(\Omega)_t$.

Démonstration du théorème B. — Rappelons la notion de quasi-hyperbolicité introduite dans **[14] [15]**.

Définition 6.8. — Un germe ω de 1-forme différentielle à l'origine de $\mathbb{C}^2$ est dit *quasi-hyperbolique* si D_ω est un sous-ensemble invariant de $\widetilde{\mathcal{F}}_\omega$ et les singularités de $\widetilde{\mathcal{F}}_\omega$ sont toutes du type $udv + \lambda vdu + \cdots$ avec $\lambda \notin \mathbb{R}_{\leq 0}$. Si de plus il existe une composante irréductible D de D_ω telle que le groupe d'holonomie de $D - \widetilde{\Sigma}_\omega \cap D$ n'est pas résoluble, nous dirons ici que ω est *quasi-hyperbolique générique.*

Cette notion de généricité est plus faible que celle définie dans **[14]**, **[15]** mais elle est suffisante pour ce qui suit. Dans **[12] [13]** nous montrons que dans l'ensemble des 1-formes quasi-hyperboliques celles qui ne satisfont pas cette condition de généricité est contenu dans un ensemble pro-algébrique de codimension infinie. L'intérêt de cette notion est dans le

Théorème 6.9 **([14], [15])**. — *Toute déformation η topologiquement triviale d'un germe ω de 1-forme quasi-hyperbolique générique est sous-jacent à un déploiement équisingulier,* i.e. *il existe un déploiement équisingulier Ω de ω de même base que η tel que $(\eta_t)_t \equiv (\Omega_t)_t$.*

La condition de non-résolubilité d'holonomie de (6.8) implique que ω n'admet pas de *facteur intégrant* c'est à dire de fonction $k \in \mathcal{O}_{\mathbb{C}^2,0}$ telle que $d(\omega/k) = 0$; sinon l'holonomie de chaque composante de D_ω serait linéaire **[7]**. Cela impose[(4)] que deux déploiements de ω qui induisent la même déformation sont égaux, à coefficient multiplicatif près. On en déduit immédiatement le :

[(4)] C'est un calcul facile, d'autant plus que pour le lemme qui suit il suffit de le faire dans le cas d'un déploiement analytiquement trivial.

Lemme 6.10. — *Un germe de 1-forme quasi-hyperbolique générique est topologiquement quasi-homogène (1.1) si et seulement si elle est d-quasi-homogène.*

Ceci achève la démonstration. □

7. Une généralisation d'un théorème de J. Briançon

Nous allons prouver le théorème C de l'introduction, qui généralise un résultat **[2]** de J. Briançon en dimension 2 étendu par la suite en dimension quelconque par cet auteur et H. Skoda **[3]** :

– le carré de toute fonction $f \in \mathcal{O}_{\mathbb{C}^2,0}$ appartient l'idéal jacobien de f.

Fixons un germe en $0 \in \mathbb{C}^2$ de 1-forme semi-hyperbolique ω, une équation réduite f de $\mathrm{Sep}(\omega)$ et conservons les notations (1), (2). Nous allons construire une section globale V du faisceau $\mathcal{X}_{S'}$ des champs de vecteurs tangents à $S' := {E_\omega}^{-1}(\mathrm{Sep}(\omega))$, qui vérifie

$$\widetilde{\omega} \cdot V = f^2 \circ E_\omega. \tag{18}$$

On obtient le théorème C en prenant l'image directe de V sur $\mathbb{C}^2, 0$, prolongée à l'origine par le théorème d'Hartogs.

Dans M_ω, au voisinage de chaque ouvert W_i d'un recouvrement $\mathcal{W}$ assez fin de D_ω, il existe d'après le lemme (3.2) un champ de vecteurs W_i vérifiant $\widetilde{\omega} \cdot V_i = f \circ E_\omega$. Le cocycle $V_{ij} := V_i - V_j$ est à valeurs dans le faisceau $\mathcal{X}_{\widetilde{\mathcal{F}}_\omega}$ des champs tangents à $\widetilde{\mathcal{F}}_\omega$. On va voir que

$$[f \circ E_\omega \cdot V_{ij}] = 0 \in H^1(D_\omega; \mathcal{X}_{\widetilde{\mathcal{F}}_\omega}). \tag{19}$$

Ainsi les champs $f \circ E_\omega \cdot V_i$ se recollent en un champ global V vérifiant (18).

Le théorème ci-dessous, qui peut s'interpréter comme une forme géométrique du théorème de Briançon, donne immédiatement (19). Considérons une succession finie $E^i : M^i \longrightarrow M^{i-1}$, $i = 1, \ldots, h$ d'éclatements de centres finis et au dessus de l'origine de $\mathbb{C}^2 =: M^0$. Notons

$$\widetilde{E} := E^1 \circ \cdots \circ E^h, \quad D^i := {E_i}^{-1}(0), \quad \widetilde{D} := D^h,$$

et désignons par $\mathcal{X}_{\widetilde{\mathcal{F}}}$ le faisceau des champs de vecteurs de M^h tangents au feuilletage saturé $\widetilde{\mathcal{F}}$ défini par $\widetilde{E}^\star(\omega)$. Fixons une équation réduite f de $\mathrm{Sep}(\omega)$ et notons $F := f \circ \widetilde{E}$.

Théorème 7.1. — *Soit ω un germe à l'origine de $\mathbb{C}^2$ de 1-forme semi-hyperbolique. Alors le morphisme*

$$[F]\cdot : H^1\left(\widetilde{D}; \mathcal{X}_{\widetilde{\mathcal{F}}}\right) \longrightarrow H^1\left(\widetilde{D}; \mathcal{X}_{\widetilde{\mathcal{F}}}\right), \qquad [Y_{ij}] \longmapsto [F \cdot Y_{ij}]$$

est identiquement nul.

Démonstration. — Raisonnons par récurrence sur h. Nous avons calculé dans **[10]** une base de $H^1\left(\widetilde{D};\mathcal{X}_{\widetilde{\mathcal{F}}}\right)$ pour $h=1$: on se donne en $0\in\mathbb{C}^2$ un champ de vecteurs Z à singularité isolée qui annule ω ; on recouvre l'éclaté M^1 de l'origine de $\mathbb{C}^2$ par les deux cartes canoniques $(U_1;x_1:=x,y_1:=y/x)$ et $(U_2;x_2:=x/y,y_2:=y)$; le relevé de Z dans la première carte s'écrit $x_1^{\nu-1}Z_1$ ou ν est la multiplicité à l'origine de Z et Z_1 est un champ de vecteurs défini holomorphe et à singularités isolées sur U_1 ; alors les champs

$$Z_{\alpha\beta}:=x_1^{\alpha}y_1^{\beta}\cdot Z_1,\qquad (\alpha,\beta)\in I:=\{\alpha\geq 0\mid \alpha-\nu+1<\beta<0\}$$

induisent une base de $H^1\left(\widetilde{D};\mathcal{X}_{\widetilde{\mathcal{F}}}\right)$. De plus les champs de vecteurs suivants :

$$x_1^{\alpha}y_1^{\beta}\cdot Z_1,\qquad \alpha\geq 0,\qquad \beta\geq 0\quad\text{ou}\quad \beta\leq\alpha-\nu+1$$

sont dans $B^1\left(\widetilde{D};\mathcal{X}_{\widetilde{\mathcal{F}}}\right)$. La multiplicité de f à l'origine est aussi égale à ν car ω est semi-hyperbolique. Ainsi $(F)\subset(x_1^{\nu})$ et les champs $F\cdot Z_{\alpha\beta}$, $(\alpha,\beta)\in I$ sont des cobords ; et le théorème est montré pour $h=1$.

Le pas de récurrence se fait en utilisant la décomposition **[10]** suivante :

$$0\longrightarrow H^1\left(D^1;\mathcal{X}_{\widetilde{\mathcal{F}}^1}\right)\xrightarrow{\ \xi\ }H^1\left(\widetilde{D};\mathcal{X}_{\widetilde{\mathcal{F}}}\right)\xrightarrow{\ \zeta\ }H^1\left(D';\mathcal{X}_{\widetilde{\mathcal{F}}}\right)\longrightarrow 0 \tag{20}$$

où D' est l'union des composantes irréductibles de $\widetilde{D}$ autres que D^1, ζ est le morphisme de restriction et ξ est induit par le plongement naturel de D^1 dans $\widetilde{D}$. Sur le premier terme de cette suite exacte le morphisme $[F]\cdot$ est nul ; c'est le calcul ci-dessus. La nullité de $[F]\cdot$ sur le troisième terme de (20) résulte de l'hypothèse de récurrence appliquée à $\widetilde{\mathcal{F}}^1$ au voisinage de chaque point du centre d'éclatement de E^2. □

Références

[1] A. Andreotti et H. Grauert, *Théorèmes de finitude pour la cohomologie des espaces complexes*, Bulletin de la Société Mathématique de France, **90**, pages 193 à 259, (1962).

[2] J. Briançon, *A propos d'une question de J. Mather*, Preprint Université de Nice, 1973.

[3] J. Briançon et H. Skoda, *Sur la clôture intégrale d'un idéal de germes de fonctions holomorphes n un point de $\mathbb{C}^n$*, Comptes Rendus de l'Académie des Sciences, Paris, **série A 278**, pages 949 à 951, (1974).

[4] C. Camacho et P. Sad, *Invariant varieties through singularities of holomorphic vector fields*, Annals of Mathematics, **115**, pages 579 à 595, (1982).

[5] C. Camacho, A. Lins Neto et P. Sad, *Topological invariants and equidesingularization for holomorphic vector fields*, Journal of Differential Geometry, **20**, pages 143 à 174, (1984).

[6] D. Cerveau, *Distributions involutives singulières*, Annales de l'Institut Fourier, **29, 3**, pages 261 à 294, (1979).

[7] D. CERVEAU ET J.-F. MATTEI, *Formes intégrables holomorphes singulières*, Astérisque, **97**, (1982).

[8] D. EISENBUD ET W. NEUMANN, *Three-dimensional link theory and invariants of plane curve singularities*, Princeton University Press, Annals of mathematics Studies, 110, (1985).

[9] LÊ DUNG TRANG ET C.P. RAMANUJAM, *The invariance of Milnor's number implies the invariance of the topological type*, American Journal of Mathematics, **98, 1**, pages 67 à 78, (1976).

[10] J.-F. MATTEI, *Modules de feuilletages holomorphes singuliers : 1 équisingularité*, Inventiones Mathematicae, **103**, pages 297 à 325, (1991).

[11] J.-F. MATTEI ET R. MOUSSU, *Holonomie et intégrales premières*, Annales Scientifiques de l'Ecole Normale Supérieure, **Série 4, t. 13**, pages 469 à 523, (1980).

[12] J-F. MATTEI ET E. SALEM, *Classification formelle des feuilletages génériques de* $(\mathbb{C}^2, 0)$, preprint.

[13] J-F. MATTEI ET E. SALEM, *Classification formelle et analytique de feuilletages singuliers de* $(\mathbb{C}^2, 0)$, C. R. Académie des Sciences, Paris, **Série 1, t. 352**, pages 773 à 778, (1997).

[14] J-F. MATTEI ET E. SALEM, *Classification topologique et analytique des feuilletages génériques de* $(\mathbb{C}^2, 0)$, en préparation.

[15] J-F. MATTEI ET E. SALEM, *Complete systems of topological and analytical invariants for a generic foliation of* $(\mathbb{C}^2, 0)$, Mathematical Reasearch Letters, **4**, pages 131 à 141, (1997).

[16] K. SAITO, *Quasi-homogene isolierte Singularitäten von Hyperflächen*, Inventiones Mathematicae, **14**, pages 123 à 142, (1971).

[17] A. SEIDENBERG, *Reduction of singularities of the differentiable equation* $AdY = BdX$, American Journal of Mathematics, **90**, pages 248 à 269, (1968).

[18] O. ZARISKI, *Studies in equisingularity I : Equivalent singularities of algebraic curves*, American Journal of Mathematics, **87**, pages 507 à 536, (1965).

[19] O. ZARISKI, *Studies in equisingularity II : Equisingularity in codimension 1 (and characteristic zero)*, American Journal of Mathematics, **87**, pages 972 à 1006, (1965).

[20] O. ZARISKI, *On the topology of algebroid singularities*, American Journal of Mathematics, **54**, pages 433 à 465, (1932).

J.-F. MATTEI, Université Paul Sabatier, Laboratoire de Mathématiques Emile Picard, UMR-CNRS 5580, 118 route de Narbonne, 31 062 Toulouse Cedex, France
E-mail : matteiᴓpicard.ups-tlse.fr

JOHN MILNOR

Periodic orbits, externals rays and the Mandelbrot set: an expository account

Astérisque, tome 261 (2000), p. 277-333
<http://www.numdam.org/item?id=AST_2000__261__277_0>

Astérisque
261, 2000, p. 277–333

PERIODIC ORBITS, EXTERNALS RAYS AND THE MANDELBROT SET: AN EXPOSITORY ACCOUNT

by

John Milnor

Dedicated to Adrien Douady on the occasion of his sixtieth birthday

Abstract. — A presentation of some fundamental results from the Douady-Hubbard theory of the Mandelbrot set, based on the idea of "orbit portrait": the pattern of external rays landing on a periodic orbit for a quadratic polynomial map.

1. Introduction

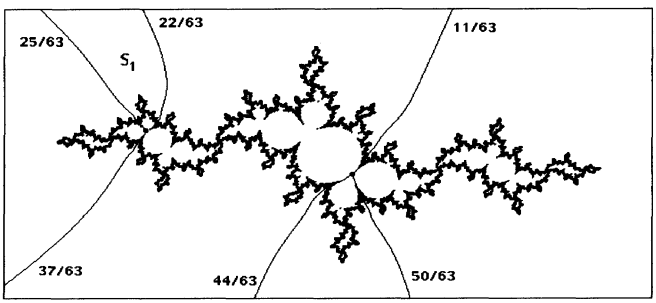


FIGURE 1. Julia set for $z \mapsto z^2 + (\frac{1}{4}e^{2\pi i/3} - 1)$ showing the six rays landing on a period two parabolic orbit. The associated orbit portrait has characteristic arc $\mathcal{I} = (22/63, 25/63)$ and valence $v = 3$ rays per orbit point.

A key point in Douady and Hubbard's study of the Mandelbrot set M is the theorem that every parabolic point $c \neq 1/4$ in M is the landing point for exactly two external rays with angles which are periodic under doubling. (See **[DH2]**. By

1991 ***Mathematics Subject Classification.*** — 30D05.
Key words and phrases. — Mandelbrot set, periodic orbit.

definition, a parameter point is *parabolic* if and only if the corresponding quadratic map has a periodic orbit with some root of unity as multiplier.) This note will try to provide a proof of this result and some of its consequences which relies as much as possible on elementary combinatorics, rather than on more difficult analysis. It was inspired by §2 of the recent thesis of Schleicher **[S1]**, which contains very substantial simplifications of the Douady-Hubbard proofs with a much more compact argument, and is highly recommended. (See also **[S2]**, **[LS]**.) The proofs given here are rather different from those of Schleicher, and are based on a combinatorial study of the angles of external rays for the Julia set which land on periodic orbits. (Compare **[A]**, **[GM]**.) As in **[DH1]**, the basic idea is to find properties of M by a careful study of the dynamics for parameter values outside of M. The results in this paper are mostly well known; there is a particularly strong overlap with **[DH2]**. The only claim to originality is in emphasis, and the organization of the proofs. (Similar methods can be used for higher degree polynomials with only one critical point. Compare **[S3]**, **[E]**, and see **[PR]** for a different approach. For a theory of polynomial maps which may have many critical points, see **[K]**.)

We will assume some familiarity with the classical Fatou-Julia theory, as described for example in **[Be]**, **[CG]**, **[St]**, or **[M2]**.

Standard Definitions. — (Compare Appendix A.) Let $K = K(f_c)$ be the *filled Julia set*, that is the union of all bounded orbits, for the quadratic map

$$f(z) = f_c(z) = z^2 + c.$$

Here both the parameter c and the dynamic variable z range over the complex numbers. The *Mandelbrot set* M can be defined as the compact subset of the *parameter plane* (or c-plane) consisting of all complex numbers c for which $K(f_c)$ is connected. We can also identify the complex number c with one particular point in the *dynamic plane* (or z-plane), namely the *critical value* $f_c(0) = c$ for the map f_c. The parameter c belongs to M if and only if the orbit $f_c : 0 \mapsto c \mapsto c^2 + c \mapsto \cdots$ is bounded, or in other words if and only if $0, c \in K(f_c)$. Associated with each of the compact sets $K = K(f_c)$ in the dynamic plane there is a *potential function* or *Green's function* $G^K : \mathbb{C} \to [0, \infty)$ which vanishes precisely on K, is harmonic off K, and is asymptotic to $\log|z|$ near infinity. The family of *external rays* of K can be described as the orthogonal trajectories of the level curves $G^K =$ constant. Each such ray which extends to infinity can be specified by its angle at infinity $t \in \mathbb{R}/\mathbb{Z}$, and will be denoted by $\mathcal{R}_t^K$. Here c may be either in or outside of the Mandelbrot set. Similarly, we can consider the potential function G^M and the external rays $\mathcal{R}_t^M$ associated with the Mandelbrot set. We will use the term *dynamic ray* (or briefly K-*ray*) for an external ray of the filled Julia set, and *parameter ray* (or briefly M-*ray*) for an external ray of the Mandelbrot set. (Compare **[S1]**, **[S2]**.)

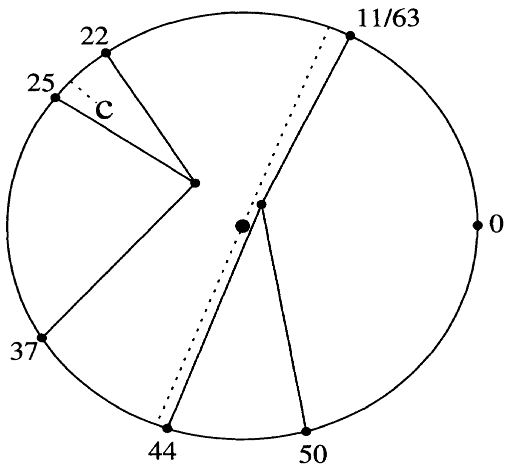


FIGURE 2. Schematic diagram illustrating the orbit portrait (1).

Definition. — Let $\mathcal{O} = \{z_1, \ldots, z_p\}$ be a periodic orbit for f. Suppose that there is some rational angle $t \in \mathbb{Q}/\mathbb{Z}$ so that the dynamic ray $\mathcal{R}_t^{K(f)}$ lands at a point of $\mathcal{O}$. Then for each $z_i \in \mathcal{O}$ the collection A_i consisting of all angles of dynamic rays which land at the point z_i is a finite and non-vacuous subset of $\mathbb{Q}/\mathbb{Z}$. The collection $\{A_1, \ldots, A_p\}$ will be called the *orbit portrait* $\mathcal{P} = \mathcal{P}(\mathcal{O})$. As an example, Figure 1 shows a quadratic Julia set having a parabolic orbit with portrait

$$\mathcal{P} = \{\ \{22/63,\ 25/63,\ 37/63\},\ \ \{11/63,\ 44/63,\ 50/63\}\ \}. \tag{1}$$

It is often convenient to represent such a portrait by a schematic diagram, as shown in Figure 2. (For details, and an abstract characterization of orbit portraits, see §2.)

The number of elements in each A_i (or in other words the number of K-rays which land on each orbit point) will be called the *valence* v. Let us assume that $v \geq 2$. Then the v rays landing at z cut the dynamic plane up into v open regions which will be called the *sectors* based at the orbit point $z \in \mathcal{O}$. The *angular width* of a sector S will mean the length of the open arc I_S consisting of all angles $t \in \mathbb{R}/\mathbb{Z}$ with $\mathcal{R}_t^K \subset S$. (We use the word 'arc' to emphasize that we will identify $\mathbb{R}/\mathbb{Z}$ with the 'circle at infinity' surrounding the plane of complex numbers.) Thus the sum of the angular widths of the v distinct sectors based at an orbit point z is always equal to $+1$. The following result will be proved in 2.11.

***Theorem 1.1* (The Critical Value Sector S_1).** — *Let $\mathcal{O}$ be an orbit of period $p \geq 1$ for $f = f_c$. If there are $v \geq 2$ dynamic rays landing at each point of $\mathcal{O}$, then there is one and only one sector S_1 based at some point $z_1 \in \mathcal{O}$ which contains the critical value $c = f(0)$, and whose closure contains no point other than z_1 of the orbit $\mathcal{O}$. This* *critical value sector S_1 can be characterized, among all of the pv sectors based at the various points of $\mathcal{O}$, as the unique sector of smallest angular width.*

It should be emphasized that this description is correct whether the filled Julia set K is connected or not.

Our main theorem can be stated as follows. Suppose that there exists some polynomial f_{c_0} which admits an orbit $\mathcal{O}$ with portrait $\mathcal{P}$, again having valence $v \geq 2$. Let $0 < t_- < t_+ < 1$ be the angles of the two dynamic rays $\mathcal{R}^K_{t_\pm}$ which bound the critical value sector S_1 for f_{c_0}.

***Theorem 1.2* (The Wake $W_\mathcal{P}$).** — *The two corresponding parameter rays $\mathcal{R}^M_{t_\pm}$ land at a single point $\mathbf{r}_\mathcal{P}$ of the parameter plane. These rays, together with their landing point, cut the plane into two open subsets $W_\mathcal{P}$ and $\mathbb{C} \smallsetminus \overline{W}_\mathcal{P}$ with the following property: A quadratic map f_c has a repelling orbit with portrait $\mathcal{P}$ if and only if $c \in W_\mathcal{P}$, and has a parabolic orbit with portrait $\mathcal{P}$ if and only if $c = \mathbf{r}_\mathcal{P}$.*

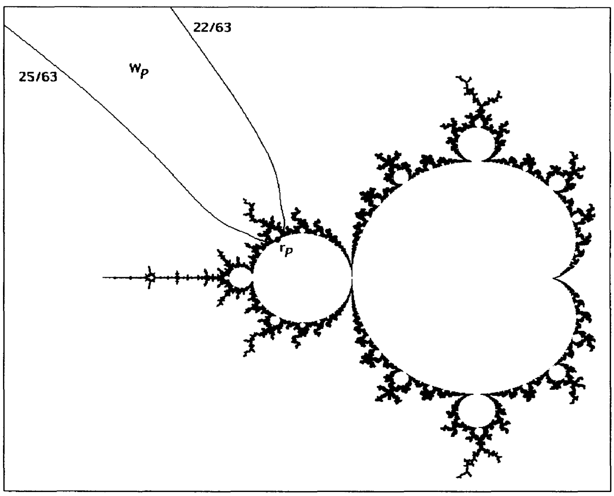


FIGURE 3. The boundary of the Mandelbrot set, showing the wake $W_\mathcal{P}$ and the root point $\mathbf{r}_\mathcal{P} = \frac{1}{4}\, e^{2\pi i/3} - 1$ associated with the orbit portrait of Figure 1, with characteristic arc $\mathcal{I}_\mathcal{P} = (22/63, 25/63)$.

In fact this will follow by combining the assertions 3.1, 4.4, 4.8, and 5.4 below.

Definitions. — This open set $W_\mathcal{P}$ will be called the *$\mathcal{P}$-wake* in parameter space (compare Atela [**A**]), and $\mathbf{r}_\mathcal{P}$ will be called the *root point* of this wake. The intersection $M_\mathcal{P} = M \cap \overline{W}_\mathcal{P}$ will be called the *$\mathcal{P}$-limb* of the Mandelbrot set. The open arc $I_{S_1} = (t_-, t_+)$ consisting of all angles of dynamic rays $\mathcal{R}^K_t$ which are contained in the interior of S_1, or all angles of parameter rays $\mathcal{R}^M_t$ which are contained in $W_\mathcal{P}$, will be called the *characteristic arc* $\mathcal{I} = \mathcal{I}_\mathcal{P}$ for the orbit portrait $\mathcal{P}$. (Compare 2.6.)

In general, the orbit portraits with valence $v = 1$ are of little interest to us. These portraits certainly exist. For example, for the base map $f_0(z) = z^2$ which lies outside of every wake, *every* orbit portrait has valence $v = 1$. As we follow a path in parameter space which crosses into the wake $W_{\mathcal{P}}$ through its root point, either one orbit with a portrait of valence one degenerates to form an orbit of lower period with portrait $\mathcal{P}$, or else two different orbits with portraits of valence one fuse together to form an orbit with portrait $\mathcal{P}$. (If we cross into $W_{\mathcal{P}}$ through a parameter ray $\mathcal{R}^M_{t_\pm}$, the picture is similar except that the landing point of the dynamic ray $\mathcal{R}^K_{t_\pm}$ jumps discontinuously. If t_+ and t_- belong to the same cycle under angle doubling, then the landing points of both of these dynamics rays jump discontinuously.)

However, there is one exceptional portrait of valence one: The *zero portrait* $\mathcal{P} = \{\{0\}\}$ will play an important role. It is not difficult to check that the dynamic ray $\mathcal{R}^K_0$ of angle 0 for f_c lands at a well defined fixed point if and only if the parameter value c lies in the complement of the parameter ray $\mathcal{R}^M_0 = \mathcal{R}^M_1 = (1/4, \infty)$. Furthermore, this fixed point necessarily has portrait $\{\{0\}\}$. Thus the wake, consisting of all $c \in \mathbb{C}$ for which f_c has a repelling fixed point with portrait $\{\{0\}\}$, is just the complementary region $\mathbb{C} \smallsetminus [1/4, \infty)$. The characteristic arc $\mathcal{I}_{\{\{0\}\}}$ for this portrait, consisting of all angles t such that $\mathcal{R}^K_t \subset W_{\{\{0\}\}}$, is the open interval $(0, 1)$, and the root point $\mathbf{r}_{\{\{0\}\}}$, the unique parameter value c such that f_c has a parabolic fixed point with portrait $\{\{0\}\}$, is the landing point $c = 1/4$ for the zero parameter ray.

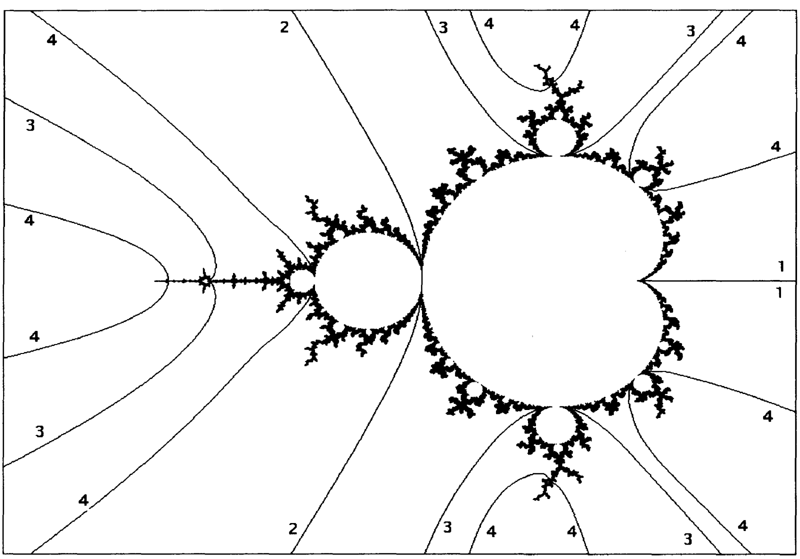


FIGURE 4. Boundaries of the wakes of ray period four or less.

Definition. — It will be convenient to say that a portrait $\mathcal{P}$ is *non-trivial* if it either has valence $v \geq 2$ or is equal to this zero portrait.

Remark. — An alternative characterization would be the following. An orbit portrait $\{A_1, \ldots, A_p\}$ is non-trivial if and only if it is *maximal*, in the sense that there is no orbit portrait $\{A'_1, \ldots, A'_q\}$ with $A'_1 \underset{\neq}{\supset} A_1$. This statement follows easily from 1.5 and 2.7 below. Still another characterization would be that $\mathcal{P}$ is non-trivial if and only if it is the portrait of some parabolic orbit. (See 5.4.)

***Corollary 1.3* (Orbit Forcing).** — *If $\mathcal{P}$ and $\mathcal{Q}$ are two distinct non-trivial orbit portraits, then the boundaries $\partial W_{\mathcal{P}}$ and $\partial W_{\mathcal{Q}}$ of the corresponding wakes are disjoint subsets of $\mathbb{C}$. Hence the closures $\overline{W}_{\mathcal{P}}$ and $\overline{W}_{\mathcal{Q}}$ are either disjoint or strictly nested. In particular, if $\mathcal{I}_{\mathcal{P}} \subset \mathcal{I}_{\mathcal{Q}}$ with $\mathcal{P} \neq \mathcal{Q}$, then it follows that $\overline{W}_{\mathcal{P}} \subset W_{\mathcal{Q}}$.*

Thus whenever $\overline{\mathcal{I}}_{\mathcal{P}} \subset \mathcal{I}_{\mathcal{Q}}$, the existence of a repelling or parabolic orbit with portrait $\mathcal{P}$ *forces* the existence of a *repelling* orbit with portrait $\mathcal{Q}$. We will write briefly $\mathcal{P} \Rightarrow \mathcal{Q}$. On the other hand, if $\mathcal{I}_{\mathcal{P}} \cap \mathcal{I}_{\mathcal{Q}} = \varnothing$ then no f_c can have both an orbit with portrait $\mathcal{P}$ and an orbit with portrait $\mathcal{Q}$.

See Figure 5 for a schematic description of orbit forcing relations for orbits with ray period 4 or less, corresponding to the collection of wakes illustrated in Figure 4. (Evidently this diagram, as well as analogous diagrams in which higher periods are included, has a tree structure, with no loops.)

Proof of 1.3, assuming 1.2. — First note that $W_{\mathcal{P}}$ and $W_{\mathcal{Q}}$ cannot have a boundary ray in common. For the landing point of such a common ray would have to have one parabolic orbit with portrait $\mathcal{P}$ and one parabolic orbit with portrait $\mathcal{Q}$. But a quadratic map, having only one critical point, cannot have two distinct parabolic orbits. In fact this argument shows that $\partial W_{\mathcal{P}} \cap \partial W_{\mathcal{Q}} = \varnothing$. Note that the parameter point $c = 0$ (corresponding to the map $f_0(z) = z^2$) does not belong to any wake $W_{\mathcal{P}}$ with $\mathcal{P} \neq \{\{0\}\}$. Since rays cannot cross each other, it follows easily that either

$$\overline{W}_{\mathcal{P}} \subset W_{\mathcal{Q}}, \quad \text{or} \quad \overline{W}_{\mathcal{Q}} \subset W_{\mathcal{P}}, \quad \text{or} \quad \overline{W}_{\mathcal{P}} \cap \overline{W}_{\mathcal{Q}} = \varnothing,$$

as required. □

For further discussion and a more direct proof, see §7.

To fill out the picture, we also need the following two statements. To any orbit portrait $\mathcal{P} = \{A_1, \ldots, A_p\}$ we associate not only its *orbit period* p but also its *ray period* rp, that is the period of the angles $t \in A_i$ under doubling modulo one. In many cases, rp is a proper multiple of p. (Compare Figure 1.) Suppose in particular that $c \in M$ is a parabolic parameter value, that is suppose that f_c has a periodic orbit where the multiplier is an r-th root of unity, $r \geq 1$. Then one can show that the ray period for the associated portrait is equal to the product rp. (See for example **[GM]**.) This is also the period of the Fatou component containing the critical point.

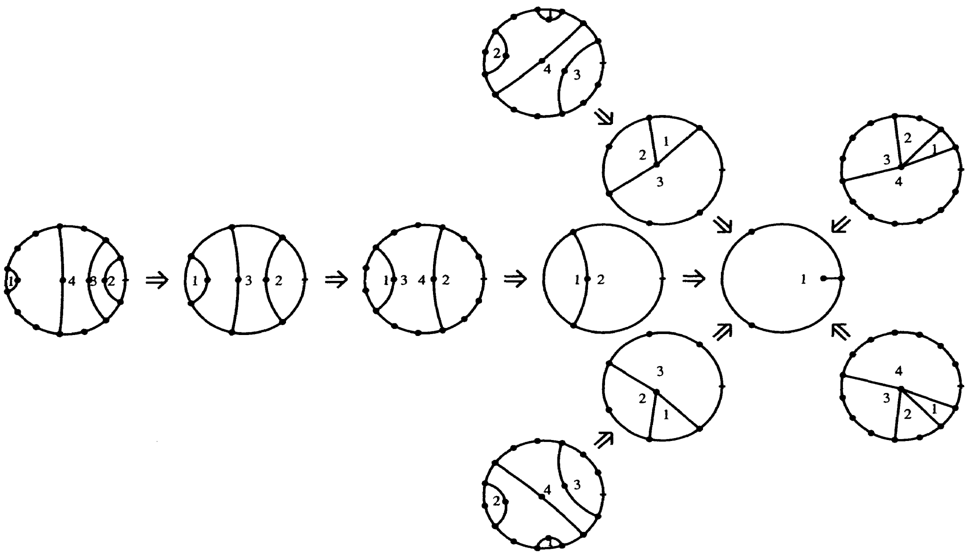


FIGURE 5. Forcing tree for the non-trivial orbit portraits of ray period $n \leq 4$. Each disk in this figure contains a schematic diagram of the corresponding orbit portrait, with the first n forward images of the critical value sector labeled. (Compare Figure 4; and compare the "disked-tree model" for the Mandelbrot set in Douady **[D5]**.)

This ray period rp is the most important parameter associated with a parabolic point c or with a wake $W_{\mathcal{P}}$.

It follows from 1.2 that every non-trivial portrait which occurs at all must occur as the portrait of some uniquely determined parabolic orbit. The converse statement will be proved in 4.8:

***Theorem 1.4* (Parabolic Portraits are Non-Trivial).** — *If c is any parabolic point in M, then the portrait $\mathcal{P} = \mathcal{P}(\mathcal{O})$ of its parabolic orbit is a non-trivial portrait. That is, if we exclude the special case $c = 1/4$, then at least two K-rays must land on each parabolic orbit point.*

It then follows immediately from 1.2 that the parabolic parameter point c must be equal to the root point $\mathbf{r}_{\mathcal{P}}$ of an associated wake. It also follows from 1.2 that the angles of the M-rays which bound a wake $W_{\mathcal{P}}$ are always periodic under doubling. In §5 we use a simple counting argument to prove the converse statement. (This imitates Schleicher, who uses a similar counting argument in a different way.)

***Theorem 1.5* (Every Periodic Angle Occurs).** — *If $t \neq 0$ in $\mathbb{R}/\mathbb{Z}$ is periodic under doubling, then $\mathcal{R}_t^M$ is one of the two boundary rays of some (necessarily unique) wake.*

Further consequences of these ideas will be developed in §6 which shows that each wake contains a uniquely associated hyperbolic component, §8 which describes how each wake contains an associated small copy of the Mandelbrot set, and §9 which shows that each limb is connected even if its root point is removed. There are two appendices giving further supporting details.

Acknowledgement. — I want to thank M. Lyubich and D. Schleicher for their ideas, which play a basic role in this presentation. I am particularly grateful to Schleicher, to S. Zakeri, and to Tan Lei for their extremely helpful criticism of the manuscript. Also, I want to thank both the Gabriella and Paul Rosenbaum Foundation and the National Science Foundation (Grant DMS-9505833) for their support of mathematical activities at Stony Brook.

2. Orbit Portraits

This section will begin the proofs by describing the basic properties of orbit portraits. We will need the following. Let $f(z) = z^2 + c$ with filled Julia set K.

***Lemma 2.1* (Mapping of Rays).** — *If a dynamic ray $\mathcal{R}_t^K$ lands at a point $z \in \partial K$, then the image ray $f(\mathcal{R}_t^K) = \mathcal{R}_{2t}^K$ lands at the image point $f(z)$. Furthermore, if three or more rays $\mathcal{R}_{t_1}^K, \mathcal{R}_{t_2}^K, \ldots, \mathcal{R}_{t_k}^K$ land at $z \neq 0$, then the cyclic order of the angles t_i around the circle $\mathbb{R}/\mathbb{Z}$ is the same as the cyclic order of the doubled angles $2t_i \pmod{\mathbb{Z}}$ around $\mathbb{R}/\mathbb{Z}$.*

Proof. — Since each $\mathcal{R}_{t_j}^K$ is assumed to be a smooth ray, it cannot pass through any precritical point. Hence $\mathcal{R}_{2t_j}^K$ also cannot pass through a precritical point, and must be a smooth ray landing at $f(z)$. Now suppose that we are given three or more rays with angles $0 \leq t_1 < t_2 < \cdots < t_k < 1$, all landing at z. These rays, together with their landing point, cut the plane up into sectors $S_1, \ldots, S_k$, where each S_i is bounded by $\mathcal{R}_{t_i}^K$ and $\mathcal{R}_{t_{i+1}}^K$ (with subscripts modulo k). The cyclic ordering of these various rays can be measured within an arbitrarily small neighborhood of the landing point z, since any transverse arc which crosses $\mathcal{R}_{t_i}^K$ in the positive direction must pass from S_{i-1} to S_i. Since f maps a neighborhood of z to a neighborhood of $f(z)$ by an orientation preserving diffeomorphism, it follows that the image rays must have the same cyclic order. □

Now let us impose the following.

***Standing Hypothesis 2.2.*.** — $\mathcal{O} = \{z_1, \ldots, z_p\}$ is a periodic orbit for a quadratic map $f_c(z) = z^2 + c$, with orbit points numbered so that $f(z_j) = z_{j+1}$, taking subscripts modulo p. Furthermore there is at least one rational angle $t \in \mathbb{Q}/\mathbb{Z}$ so that the dynamic ray $\mathcal{R}_t^K$ associated with f lands at some point of this orbit $\mathcal{O}$.

If c belongs to the Mandelbrot set M, or in other words if the filled Julia set K is connected, then this condition will be satisfied if and only if the orbit $\mathcal{O}$ is either repelling or parabolic. (Compare **[Hu]**, **[M3]**.) On the other hand, for $c \notin M$, all periodic orbit are repelling, but the condition may fail to be satisfied either because the rotation number is irrational (compare [**GM**, Figure 16]), or because the K-rays which 'should' land on $\mathcal{O}$ bounce off precritical points en route ([**GM**, Figure 14]).

As in §1, let $A_j \subset \mathbb{R}/\mathbb{Z}$ be the set of all angles of K-rays which land on the point $z_j \in \mathcal{O}$.

***Lemma 2.3* (Properties of Orbit Portraits).** — *If this Standing Hypothesis 2.2 is satisfied, then:*

(1) *Each A_j is a finite subset of $\mathbb{Q}/\mathbb{Z}$.*
(2) *For each j modulo p, the doubling map $t \mapsto 2t \pmod{\mathbb{Z}}$ carries A_j bijectively onto A_{j+1} preserving cyclic order around the circle,*
(3) *All of the angles in $A_1 \cup \cdots \cup A_p$ are periodic under doubling, with a common period rp, and*
(4) *the sets $A_1, \ldots, A_p$ are* pairwise unlinked; *that is, for each $i \neq j$ the sets A_i and A_j are contained in disjoint sub-intervals of $\mathbb{R}/\mathbb{Z}$.*

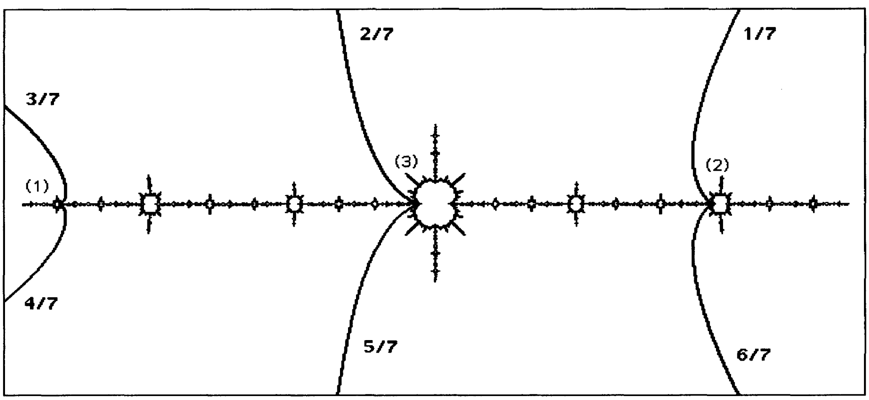


FIGURE 6. Julia set for $z \mapsto z^2 - 7/4$, showing the six K-rays landing on a period three parabolic orbit. Each number (j) in parentheses is close to the orbit point z_j (and also to $f^{\circ j}(0)$).

As in §1, the collection $\mathcal{P} = \{A_1, \ldots, A_p\}$ is called the *orbit portrait* for the orbit $\mathcal{O}$. As examples, Figure 6 shows an orbit of period and ray period three, with portrait

$$\mathcal{P} = \{\ \{3/7, 4/7\},\ \{6/7, 1/7\},\ \{5/7, 2/7\}\ \},$$

Figure 7 shows a period three orbit with ray period six, and with portrait

$$\mathcal{P} = \{\ \{4/9, 5/9\},\ \{8/9, 1/9\},\ \{7/9, 2/9\}\ \},$$

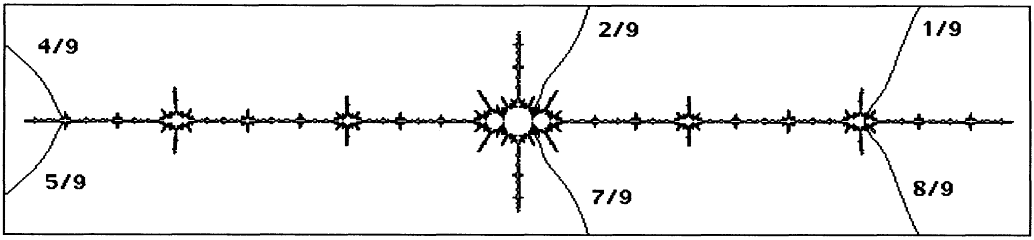


FIGURE 7. Julia set for $z \mapsto z^2 - 1.77$, showing the six K-rays landing on a period three orbit. In contrast to Figure 6, these six rays are permuted cyclically by the map.

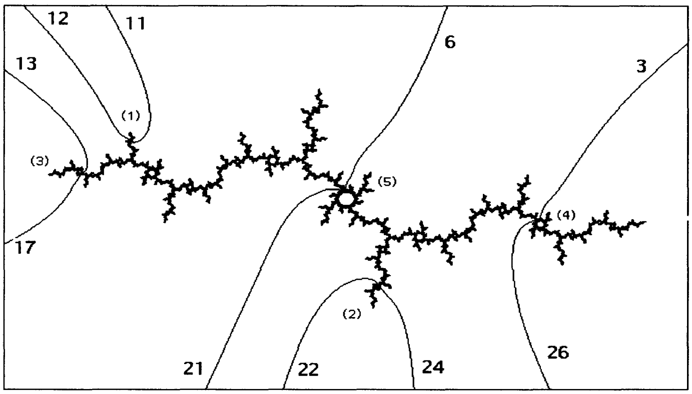


FIGURE 8. Julia set $J(f_c)$ for $c = -1.2564 + .3803\,i$, showing the ten rays landing on a period 5 orbit. Here the angles are in units of $1/31$.

while Figure 8 shows an orbit of period and ray period five, with portrait

$$\mathcal{P} = \left\{ \left\{\frac{11}{31}, \frac{12}{31}\right\}, \left\{\frac{22}{31}, \frac{24}{31}\right\}, \left\{\frac{13}{31}, \frac{17}{31}\right\}, \left\{\frac{26}{31}, \frac{3}{31}\right\}, \left\{\frac{21}{31}, \frac{6}{31}\right\} \right\}.$$

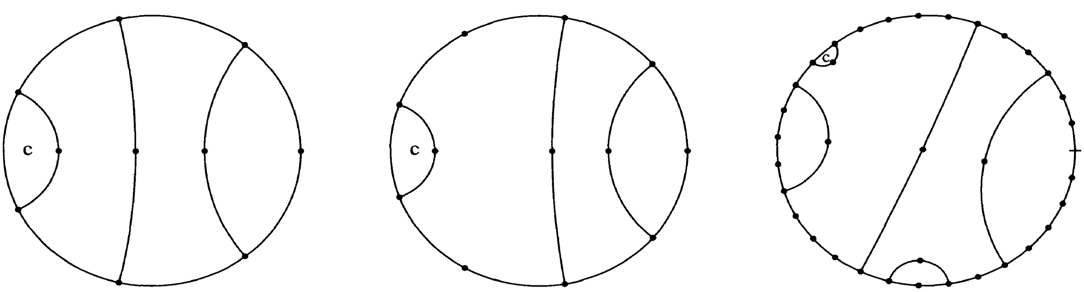


FIGURE 9. Schematic diagrams associated with the orbit portraits of Figures 6, 7, 8. The angles are in units of $1/7$, $1/9$ and $1/31$ respectively.

Proof of 2.3. — Since some A_i contains a rational number modulo $\mathbb{Z}$, it follows from 2.1 that some A_j contains an angle t_0 which is periodic under doubling. Let the period be $n \geq 1$, so that $2^n t_0 \equiv t_0 \pmod{\mathbb{Z}}$. Applying 2.1 n times, we see that the mapping $\eta(t) \equiv 2^n t \pmod{\mathbb{Z}}$ maps the set $A_j \subset \mathbb{R}/\mathbb{Z}$ injectively into itself, preserving cyclic order and fixing t_0. In fact we will show that every element of A_j is fixed by η. For otherwise, if $t \in A_j$ were not fixed, then choosing suitable representatives modulo $\mathbb{Z}$ we would have for example $t_0 = \eta(t_0) < t < \eta(t) < t_0 + 1$. Since η preserves cyclic order, it would then follow inductively that

$$t_0 < t < \eta(t) < \eta^{\circ 2}(t) < \eta^{\circ 3}(t) < \cdots < t_0 + 1.$$

Hence the successive images of t would converge to a fixed point of η. But this is impossible since every fixed point of η is repelling. Thus η fixes every point of A_j. But the fixed points of η are precisely the rational numbers of the form $i/(2^n - 1)$, so it follows that A_j is a finite set of rational numbers. It follows easily that all of the A_k are pointwise fixed by η. This proves (1), (2) and (3) of 2.3; and (4) is clearly true since rays cannot cross each other. □

It is often convenient to compactify the complex numbers by adding a circle of points $e^{2\pi i t}\infty$ at infinity, canonically parametrized by $t \in \mathbb{R}/\mathbb{Z}$. Within the resulting closed topological disk ©, we can form a *diagram* $\mathcal{D}$ illustrating the orbit portrait $\mathcal{P}$ by drawing all of the K-rays joining the circle at infinity to $\mathcal{O}$. These various rays are disjoint, except that each $z \in \mathcal{O}$ is a common endpoint for exactly v of these rays.

Note that this diagram $\mathcal{D}$ deforms continuously, preserving its topology, as we move the parameter point c, provided that the periodic orbit $\mathcal{O}$ remains repelling, and provided that the associated K-rays do not run into precritical points. (Compare [**GM**, Appendix B].)

In fact, given $\mathcal{P}$, we can construct a diagram homeomorphic to $\mathcal{D}$ as follows. Start with the unit circle, and mark all of the points $e(t) = e^{2\pi i t}$ corresponding to angles t in the union $\mathbf{A}_{\mathcal{P}} = A_1 \cup \cdots \cup A_p$. Now for each A_i, let $\widehat{z}_i$ be the center of gravity of the corresponding points $e(t)$, and join each of these points to $\widehat{z}_i$ by a straight line segment. It follows easily from Condition (4) that these line segments will not cross each other. (In practice, in drawing such diagrams, we will not usually use straight lines and centers of gravity, but rather use some topologically equivalent picture, fixing the boundary circle, which is easier to see. Compare Figures 2, 5, 9.)

It will be convenient to temporarily introduce the term *formal orbit portrait* for a collection $\mathcal{P} = \{A_1, \ldots, A_p\}$ of subsets of $\mathbb{R}/\mathbb{Z}$ which satisfies the four conditions of 2.3, whether or not it is actually associated with some periodic orbit. In fact we will prove the following.

***Theorem 2.4* (Characterization of Orbit Portraits).** — *If $\mathcal{P}$ is any formal orbit portrait, then there exists a quadratic polynomial f and an orbit $\mathcal{O}$ for f which realizes this portrait $\mathcal{P}$.*

This will follow from Lemma 2.9 below. To begin the proof, let us study the way in which the angle doubling map acts on a formal orbit portrait. As in §1, the number of angles in each A_j will be called the *valence* v for the formal portrait $\mathcal{P}$. It is easy to see that any formal portrait of valence $v = 1$ can be realized by an appropriate orbit for the map $f(z) = z^2$. Hence it suffices to study the case $v \geq 2$. For each $A_j \in \mathcal{P}$ the v connected components of the complement $\mathbb{R}/\mathbb{Z} \smallsetminus A_j$ are connected open arcs with total length $+1$. These will be called the *complementary arcs* for A_j.

Lemma 2.5 **(The Critical Arcs).** — *For each A_j in the formal orbit portrait $\mathcal{P}$, all but one of the complementary arcs is carried diffeomorphically by the angle doubling map onto a complementary arc for A_{j+1}. However, the remaining complementary arc for A_j has length greater than $1/2$. Its image under the doubling map covers one particular complementary arc for A_{j+1} twice, and every other complementary arc for A_{j+1} just once.*

Definition. — This longest complementary arc will be called the *critical arc* for A_j. The arc which it covers twice under doubling will be called the *critical value arc* for A_{j+1}. (This language will be justified in 2.9 below.)

Proof of 2.5. — If $I \subset \mathbb{R}/\mathbb{Z}$ is a complementary arc for A_j of length less than $1/2$, then clearly the doubling map carries I bijectively onto an arc $2I$ of twice the length, bounded by two points of A_{j+1}. This image arc cannot contain any other point of A_{j+1}, since the doubling map from A_j to A_{j+1} preserves cyclic order. It follows easily that these image arcs cannot overlap. Since we cannot fit v arcs of total length $+2$ into the circle without overlap, and since there cannot be any complementary arc of length exactly $1/2$, it follows that there must be exactly one "critical" complementary arc for A_j which has length greater than $1/2$. Suppose that it has length $(1 + \varepsilon_j)/2$. Then the $v - 1$ non-critical arcs for A_j have total length $(1 - \varepsilon_j)/2$, and their images under doubling form $v - 1$ complementary arcs for A_{j+1} with total length $1 - \varepsilon_j$. Since the doubling map is exactly two-to-one, it follows easily that it maps the critical arc for A_j onto the entire circle, doubly covering one "critical value arc" for A_{j+1} which has length ε_j, and covering every other complementary arc for A_{j+1} just once. $\square$

Lemma 2.6 **(The Characteristic Arc for $\mathcal{P}$).** — *Among the complementary arcs for the various $A_j \in \mathcal{P}$, there exists a unique arc $\mathcal{I}_{\mathcal{P}}$ of shortest length. This shortest arc is a critical value arc for its A_j, and is contained in all of the other critical value arcs.*

Definition. — This shortest complementary arc $\mathcal{I}_{\mathcal{P}}$ will be called the *characteristic arc* for $\mathcal{P}$. (Compare 2.11.)

Proof of 2.6. — There certainly exists at least one complementary arc $\mathcal{I}_{\mathcal{P}}$ of minimal length ℓ among all of the complementary arcs for all of the $A_j \in \mathcal{P}$. This $\mathcal{I}_{\mathcal{P}}$ must be a critical value arc, since otherwise it would have the form $2J$ where J is some complementary arc of length $\ell/2$. Suppose then that $\mathcal{I}_{\mathcal{P}}$ is the critical value arc for

A_{j+1}, doubly covered by the critical arc I_c for A_j. Since $\mathcal{I}_{\mathcal{P}}$ is minimal, it follows from 2.3(4) that this open arc $\mathcal{I}_{\mathcal{P}}$ cannot contain any point of the union $\mathbf{A}_{\mathcal{P}} = A_1 \cup \cdots \cup A_p$. Hence its preimage under doubling also cannot contain any point of $\mathbf{A}_{\mathcal{P}}$. This preimage consists of two arcs I' and $I'' = I' + 1/2$, each of length $\ell/2$. Note that both of these arcs are contained in I_c. In fact the arc I_c of length $(1+\ell)/2$ is covered by these two open arcs of length $\ell/2$ lying at either end, together with the closed arc $I_c \smallsetminus (I' \cup I'')$ of length $(1-\ell)/2$ in the middle.

Now consider any $A_k \in \mathcal{P}$ with $k \not\equiv j$. It follows from the unlinking property 2.3(4) that the entire set A_k must be contained either in the arc $(\mathbb{R}/\mathbb{Z}) \smallsetminus I_c$ of length $(1-\ell)/2$, or in I_c and hence in the arc $I_c \smallsetminus (I' \cup I'')$ which also has length $(1-\ell)/2$. In either case, it follows that the union of all non-critical arcs for A_k is contained in this same arc of length $(1-\ell)/2$, and hence that the image of this union under doubling is contained in the arc

$$2((\mathbb{R}/\mathbb{Z}) \smallsetminus I_c) = 2(I_c \smallsetminus (I' \cup I'')) = (\mathbb{R}/\mathbb{Z}) \smallsetminus \mathcal{I}_{\mathcal{P}}$$

of length $1-\ell$. Therefore, the critical value arc for A_{k+1} contains the complementary arc $\mathcal{I}_{\mathcal{P}}$, as required. It follows that this minimal arc $\mathcal{I}_{\mathcal{P}}$ is unique. For if there were an $\mathcal{I}'_{\mathcal{P}}$ of the same length, then this argument would show that each of these two must contain the other, which is impossible. $\square$

Remark. — This characteristic arc never contains the angle zero. In fact let I_c be the critical arc whose image under doubling covers $I_{\mathcal{P}}$ twice. If $0 \in I_{\mathcal{P}}$, then it is not hard to see that one endpoint of I_c must lie in $\mathcal{I}_{\mathcal{P}}$ and the other endpoint must lie outside, in $1/2 + \mathcal{I}_{\mathcal{P}}$. But this is impossible by 2.3(4) and the minimality of $I_{\mathcal{P}}$.

Recall that the union $\mathbf{A}_{\mathcal{P}} = A_1 \cup \cdots \cup A_p$ contains pv elements, each of which has period rp under doubling. Hence this union splits up into

$$\frac{pv}{rp} = \frac{v}{r}$$

distinct cycles under doubling. If $\mathcal{P}$ is the portrait of a periodic orbit $\mathcal{O}$, then the ratio v/r can be described as the *number of cycles of K-rays* which land on the orbit $\mathcal{O}$. As examples, we have $v = r = 3$ for Figure 1 and $v = r = 2$ for Figure 7 so that there is only one cycle under doubling, but $v = 2$ and $r = 1$ for Figures 6 and 8 so that there are two distinct cycles. In fact we next show that there are at most two cycles in all cases.

***Lemma 2.7* (Primitive versus Satellite)**. — *Any formal orbit portrait of valence $v > r$ must have $v = 2$ and $r = 1$. It follows that there are just two posibilities:*

Primitive Case. *If $r = 1$, so that every ray which lands on the period p orbit is mapped to itself by $f^{\circ p}$, then at most two rays land on each orbit point.*

Satellite Case. *If $r > 1$, then $v = r$ so that exactly r rays land on each orbit point, and all of these rays belong to a single cyclic orbit under angle doubling.*

This terminology will be justified in §6. (Compare Figure 12.)

Proof of 2.7. — Suppose that $v > r$ and $v \geq 3$. Let $\mathcal{I}_\mathcal{P}$ be the characteristic arc. We suppose that $\mathcal{I}_\mathcal{P}$ is the critical value arc in the complement of A_1. Let I_- the complementary arc for A_1 which is just to the left of $I_\mathcal{P}$ and let I_+ be the complementary arc just to the right of $I_\mathcal{P}$. To fix our ideas, suppose that I_- has length $\ell(I_-) \geq \ell(I_+)$. Since I_+ is not the critical value arc for A_1, we see, arguing as in 2.6, that it must be the image under iterated doubling of the critical value arc I' for some A_j. That is, we have $I_+ = 2^m I'$ for some $m \geq 1$. Hence $\ell(I') < \ell(I_+)$.

The hypothesis that $v > r$ implies that the two endpoints of $\mathcal{I}_\mathcal{P}$ belong to different cycles under doubling. Thus the left endpoints of I' and $\mathcal{I}_\mathcal{P}$ belong to distinct cycles, hence $I' \neq \mathcal{I}_\mathcal{P}$. Therefore, by 2.6, I' strictly contains $\mathcal{I}_\mathcal{P}$. This arc I' cannot strictly contain the neighboring arc I_+, since it is shorter than I_+. Hence it must have an endpoint in I_+, and therefore, by 2.3(4), it must have both endpoints in I_+. But this implies that I' contains I_-, which is impossible since $\ell(I') < \ell(I_+) \leq \ell(I_-)$. Thus, if $v > r$ it follows that $v \leq 2$, hence $r = 1$ and $v = 2$, as asserted. □

***Lemma 2.8* (Two Rays determine $\mathcal{P}$).** — *Let $\mathcal{P} = \{A_1, \ldots, A_p\}$ be a formal orbit portrait of valence $v \geq 2$, and let $\mathcal{I}_\mathcal{P} = (t_-, t_+)$ be its characteristic arc, as described above. Then a quadratic polynomial f_c has an orbit with portrait $\mathcal{P}$ if and only if the two K-rays with angles t_- and t_+ for the filled Julia set of f_c land at a common point.*

Proof. — If f_c has an orbit with portrait $\mathcal{P}$, this is true by definition. Conversely, if these rays land at a common point z_1, then the orbit of z_1 is certainly periodic. Let $\mathcal{P}'$ be the portrait for this actual orbit. We will denote its period by p', its valence by v', and so on. Note that the ray period rp is equal to $r'p'$, the common period of the angles t_- and t_+ under doubling.

Primitive Case. — Suppose that $r = 1$ so that $v/r = 2$, and so that each of these angles $t_\pm$ has period exactly p under doubling. If $p' < p$ hence $r' > 1$, then it would follow from 2.7 applied to the portrait $\mathcal{P}'$ that t_- and t_+ must belong to the same cycle under doubling, contradicting the hypothesis that $v/r = 2$.

Satellite Case. — If $r > 1$ hence $v = r$, then t_- and t_+ do belong to the same cycle under doubling, say $2^k t_- \equiv t_+ \pmod{\mathbb{Z}}$. Clearly it follows that $r' > 1$ hence $v' = r'$. Furthermore, it follows easily that multiplication by 2^k acts transitively on A_1, and hence that all of the rays $\mathcal{R}_t^K$ with $t \in A_1$ land at the same point z_1. In other words $A_1 \subset A_1'$. This implies that $r \leq r'$ hence $p \geq p'$. If p were strictly greater than p', then it would follow that $A_{1+p'}$ is also contained in A_1'. But the two sets A_1 and $A_{1+p'}$ are unlinked in $\mathbb{R}/\mathbb{Z}$. Hence there is no way that multiplication by 2^p can act non-trivially on $A_1 \cup A_{1+p'}$ carrying each of these two sets into itself and preserving cyclic order on their union. This contradiction implies that $A_1 = A_1'$ and $p = p'$, and hence that $\mathcal{P} = \mathcal{P}'$, as required. □

Now let c be some parameter value outside the Mandelbrot set. Then, following Douady and Hubbard, the point c, either in the dynamic plane or in the parameter plane, lies on a unique external ray, with the same well defined angle $t(c) \in \mathbb{R}/\mathbb{Z}$ in either case. (Compare Appendix A.)

***Lemma 2.9* (Outside the Mandelbrot Set).** — *Let $\mathcal{P} = \{A_1, \ldots, A_p\}$ be a formal orbit portrait with characteristic arc $\mathcal{I}_{\mathcal{P}}$, and let c be a parameter value outside of the Mandelbrot set. Then the map $f_c(z) = z^2 + c$ admits a periodic orbit with portrait $\mathcal{P}$ if and only if the* external angle *$t(c)$ belongs to this open arc $\mathcal{I}_{\mathcal{P}}$.*

Proof. — The two dynamic rays $\mathcal{R}^K_{t(c)/2}$ and $\mathcal{R}^K_{(1+t(c))/2}$ meet at the critical point 0, and together cut the dynamic plane into two halves. Furthermore, every point of the Julia set $\partial K = K$ is uniquely determined by its symbol sequence with respect to this partition. Correspondingly, the two diametrically opposite points $t(c)/2$ and $(1 + t(c))/2$ on the circle $\mathbb{R}/\mathbb{Z}$ cut the circle into two semicircles, and almost every point $t \in \mathbb{R}/\mathbb{Z}$ has a well defined symbol sequence with respect to this partition under the doubling map. Two rays $\mathcal{R}^K_t$ and $\mathcal{R}^K_u$ land at a common point of K if and only if the external angles t and u have the same symbol sequence.

First suppose that the angle $t(c)$ lies in the characteristic arc $\mathcal{I}_{\mathcal{P}}$. Then, with notation as in the proof of 2.6, the two points $t(c)/2$ and $(1 + t(c))/2$ lie in the two components I' and I'' of the preimage of $\mathcal{I}_{\mathcal{P}}$. For every $A_j \in \mathcal{P}$, all of the points of A_j lie in a single component of $\mathbb{R}/\mathbb{Z} \smallsetminus (I' \cup I'')$. Hence the rays $\mathcal{R}^K_t$ with $t \in A_j$ land at a common point $z_j \in K$. It follows from 2.8 that these points lie in an orbit with portrait $\mathcal{P}$, as required.

On the other hand, if $t(c)$ lies outside of $\overline{\mathcal{I}}_{\mathcal{P}}$, then it is easy to check that the two endpoints of $\mathcal{I}_{\mathcal{P}}$ are separated by the points $t(c)/2$ and $(1 + t(c))/2$. Hence these two endpoints, both belonging to $A_1 \in \mathcal{P}$, land at different points of K. Hence f_c has no orbit with portrait $\mathcal{P}$.

Finally, in the limiting case where $t(c)$ is precisely equal to one of the two endpoints $t_\pm$ of $\mathcal{I}_{\mathcal{P}}$, since these angles are periodic under doubling, it follows that the ray $\mathcal{R}^K_{t_\pm}$ passes through a precritical point, and hence does not have any well defined landing point in K. This completes the proof of 2.9. □

Evidently the Realization Theorem 2.4 is an immediate corollary. Since we have proved 2.4, we can now forget about the distinction between "formal" orbit portraits and portraits which are actually realized. We can describe further properties of portraits and their associated diagrams as follows.

Definition 2.10. — Suppose that we start with any periodic orbit $\mathcal{O}$ with valence $v \geq 2$ and period $p \geq 1$, and fix some point $z_i \in \mathcal{O}$. As in §1, the v rays landing at z_i cut the dynamic plane $\mathbb{C}$ up into v open subsets which we call the *sectors* based at z_i. Evidently there is a one-to-one correspondence between sectors based at z_i and complementary arcs for the corresponding set of angles $A_i \subset \mathbb{R}/\mathbb{Z}$, characterized by the property that $\mathcal{R}_t^K$ is contained in the open sector S if and only if t is contained in the corresponding complementary arc. By definition, the *angular size* $\alpha(S) > 0$ of a sector is the length of the corresponding complementary arc, which we can think of as its "boundary at infinity". It follows that $\sum_S \alpha(S) = 1$, where the sum extends over the v sectors based at some fixed $z_i \in \mathcal{O}$.

Remark. — The angular size of a sector has nothing to do with the angle between the rays at their common landing point, which is often not even defined.

Altogether there are pv rays landing at the various points of the orbit $\mathcal{O}$. Together these rays cut the plane up into $pv - p + 1$ connected components. The closures of these components will be called the pieces of the *preliminary puzzle* associated with the diagram $\mathcal{D}$ or the associated portrait $\mathcal{P}$. Note that every closed sector $\overline{S}$ can be expressed as a union of preliminary puzzle pieces, and that every preliminary puzzle piece is equal to the intersection of the closed sectors containing it. This construction will be modified and developed further in Sections 7 and 8.

For every point z_i of the orbit, note that just one of the v sectors based at z_i contains the critical point 0. We will call this the *critical sector* at z_i, while the others will be called the *non-critical sectors* at z_i. Another noteworthy sector at z_i (not necessarily distinct from the critical sector) is the *critical value sector*, which contains $f(0) = c$.

***Lemma 2.11* (Properties of Sectors)**. — *The diagram $\mathcal{D} \subset$ © associated with any orbit $\mathcal{O}$ of valence $v \geq 2$ has the following properties:*

(a) *For each $z_i \in \mathcal{O}$, the critical sector at z_i has angular size strictly greater than 1/2. It follows that the $v - 1$ non-critical sectors at z_i have total angular size less than 1/2.*

(b) *The map f carries a small neighborhood of z_i diffeomorphically onto a small neighborhood of $z_{i+1} = f(z_i)$, carrying each sector based at z_i locally onto a sector based at z_{i+1}, and preserving the cyclic order of these sectors around their base point. The critical sector at z_i always maps locally, near z_i, onto the critical value sector based at z_{i+1}.*

(c) *Globally, each non-critical sector S at z_i is mapped homeomorphically by f onto a sector $f(S)$ based at z_{i+1}, with angular size given by $\alpha(f(S)) = 2\,\alpha(S)$. However, the critical sector at z_i maps so as to cover the entire plane, covering the critical value sector at z_{i+1} twice with a ramification point at $0 \mapsto c$, and covering every other sector just once.*

(d) *Among all of the pv sectors based at the various points of $\mathcal{O}$, there is a unique sector of smallest angular size, corresponding to the characteristic arc $\mathcal{I}_{\mathcal{P}}$. This smallest sector contains the critical value, and does not contain any other sector.*

(As usual, the index i is to be construed as an integer modulo p.) The proof, based on 2.6 and the fact that f is exactly two-to-one except at its critical point, is straightforward and will be left to the reader. Evidently Theorem 1.1 follows. $\square$

Now let us take a closer look at the dynamics of the diagram $\mathcal{D}$ or of the associated portrait $\mathcal{P}$. The iterated map $f^{\circ p}$ fixes each point $z_i \in \mathcal{O}$, permuting the various rays which land on z_i but preserving their cyclic order. Equivalently, the p-fold iterate of the doubling map carries each finite set $A_i \subset \mathbb{Q}/\mathbb{Z}$ onto itself by a bijection which preserves the cyclic order. For any fixed i mod p, we can number the angles in A_i as $0 \le t^{(1)} < t^{(2)} < \cdots < t^{(v)} < 1$. It then follows that

$$2^p\, t^{(j)} \;\equiv\; t^{(j+k)} \quad (\mathrm{mod}\ \mathbb{Z}),$$

taking superscripts modulo v, where k is some fixed residue class modulo v.

Definition 2.12. — The ratio k/v (mod $\mathbb{Z}$) is called the combinatorial *rotation number* of our orbit portrait. It is easy to check that this rotation number does not depend on the choice of orbit point z_i. Let d be the greatest common divisor of v and k. The we can express the rotation number as a fraction q/r in lowest terms, where $k = qd$ and $v = rd$. (In the special case of rotation number zero, we take $q = 0$ and $r = 1$.)

In all cases, note that the denominator $r \ge 1$ is equal to the period of the angles $t^{(j)} \in A_i$ under the mapping $t \mapsto 2^p\, t$ (mod $\mathbb{Z}$) from A_i to itself. It follows easily that the period of $t^{(j)}$ under angle doubling is equal to the product rp. Thus this definition of r as the denominator of the rotation number is compatible with our earlier notation rp for the ray period.

Notation Summary. — Since we have been accumulating quite a bit of notation, here is a brief summary:

Orbit period p: the number of distinct element in our orbit $\mathcal{O}$,
Ray period rp: the period of each angle $t \in A_1 \cup \cdots \cup A_p$ under doubling.
Rotation number q/r: describes the action of multiplication by 2^p on each set A_i.
Valence v: number of angles in each A_i, for a total of pv angles altogether.
Cycle number v/r: the number of disjoint cycles of size rp in the union $A_1 \cup \cdots \cup A_p$.

According to 2.7, this cycle number is always equal to 1 for a satellite portrait, and is at most 2 in all cases. Thus, in the case $v \ge 2$ there are just two possibilities as follows:

Primitive Case. — The rotation number is zero. There are $v = 2$ rays landing at each orbit point, for a total of $2p$ rays. These split up into two cycles of p rays each under doubling.

Satellite Case. — The rotation number is $q/r \not\equiv 0$. There are $v = r$ rays landing at each orbit point, for a total of $pv = rp$ rays altogether. These rp rays are permuted cyclically under angle doubling, so that the number of cycles is $v/r = 1$.

As examples, Figures 6, 8 illustrate primitive portraits with rotation number zero, while Figures 1, 7 show satellite portraits with rotation number 1/3 and 1/2. We will see in §6 that primitive portraits correspond to primitive hyperbolic components in the Mandelbrot set, that is, to those with a cusp point.

3. Parameter Rays

This section will prove the following preliminary version of Theorem 1.2.

Let $\mathcal{P}$ be any orbit portrait of valence $v \geq 2$, and let $\mathcal{I}_{\mathcal{P}} = (t_-, t_+)$ be its characteristic arc, where $0 < t_- < t_+ < 1$. If the quadratic polynomial $f_c = z^2 + c$ has an orbit $\mathcal{O}$ with portrait $\mathcal{P}$, recall that the two dynamic rays $\mathcal{R}^K_{t_-}$ and $\mathcal{R}^K_{t_+}$ for f_c land at a common orbit point, and together bound a sector S_1 which has minimal angular size among all of the sectors based at points of the orbit $\mathcal{O}$. This S_1 can also be characterized as the smallest of these sectors which contains the critical value c. (Compare Lemmas 2.6, 2.9, 2.11.)

***Theorem 3.1* (Parameter Rays and the Wake).** — *The two parameter rays $\mathcal{R}^M_{t_-}$ and $\mathcal{R}^M_{t_+}$ with these same angles land at a common parabolic point in the Mandelbrot set. Furthermore, these two rays, together with their common landing point, cut the parameter plane into two open subsets $W_{\mathcal{P}}$ and $\mathbb{C} \smallsetminus \overline{W}_{\mathcal{P}}$ with the following property: The quadratic map f_c has a repelling orbit with portrait $\mathcal{P}$ if and only if $c \in W_{\mathcal{P}}$.*

Proof. — Let $\mathbf{A}_{\mathcal{P}} = A_1 \cup \cdots \cup A_p$ be the set of all angles for the orbit portrait $\mathcal{P}$, and let $n = rp$ be the common period of these angles under doubling. The set $F_n \subset M$ of possibly exceptional parameter values will consist of those c for which $f_c^{\circ n}$ has a fixed point of multiplier $+1$. Since $F_n \subset \mathbb{C}$ is an algebraic variety and is not the entire complex plane, it is necessarily a finite set. As noted in **[GM]**, if c belongs to the Mandelbrot set but $c \notin F_n$, then the various dynamic rays $\mathcal{R}^{K(f_c)}_t$ with $t \in \mathbf{A}_{\mathcal{P}}$ all land on repelling periodic points, and the pattern of which of these rays land at a common point remains stable under perturbation of c throughout some open neighborhood within parameter space.

Now suppose that c lies outside of the Mandelbrot set. Then c, considered as a point in parameter space, belongs to some uniquely defined parameter ray $\mathcal{R}^M_{t(c)}$, and considered as a point in the dynamic plane for f_c, belongs to the dynamic ray $\mathcal{R}^K_{t(c)}$ with this same angle. In this case, a dynamic ray $\mathcal{R}^K_t$ for f_c has a well defined landing point in $K = K(f_c)$ if and only if the forward orbit $\{2t, 4t, 8t, \dots\}$ under doubling does not contain this critical value angle $t(c)$. Since the angles in $\mathbf{A}_{\mathcal{P}}$ are periodic, it follows that all of the dynamic rays $\mathcal{R}^K_t$ with $t \in \mathbf{A}_{\mathcal{P}}$ have well defined landing points in K if and only if $t(c) \notin \mathbf{A}_{\mathcal{P}}$.

Let $t \in \mathbf{A}_{\mathcal{P}}$ and let $c_0 \in M$ be any accumulation point for the parameter ray $\mathcal{R}^M_t$. Since every neighborhood of c_0 contains parameter values $c \in \mathcal{R}^M_t$ for which the dynamic ray $\mathcal{R}_t^{K(f_c)}$ does not land, it follows that c_0 must belong to F_n. Thus every accumulation point for $\mathcal{R}_t^M$ belongs to the finite set F_n. Since the set of all accumulation points of a ray is connected, this proves that $\mathcal{R}_t^M$ must actually land at a single point of F_n.

These parameter rays $\mathcal{R}_t^M$ with $t \in \mathbf{A}_\mathcal{P}$, together with the points of F_n, cut the complex parameter plane up into finitely many open sets U_i, and the pattern of which of the corresponding dynamic rays $\mathcal{R}_t^K$ with $t \in \mathbf{A}_\mathcal{P}$ land at a common periodic point remains fixed as c varies through any U_i. Since every U_i is unbounded, it follows from Lemma 2.9 that for $c \in U_i$ the map f_c has an orbit with portrait $\mathcal{P}$ if and only if U_i is that open set which contains the points in $\mathbb{C} \smallsetminus M$ with external angle $t(c)$ in (t_-, t_+). It follows that the two rays $\mathcal{R}_{t_-}^M$ and $\mathcal{R}_{t_+}^M$ must land at a common point of F_n, so as to separate the parameter plane. For otherwise, if they had different landing points, the connected set U_i containing points with external angles in (t_-, t_+) would also contain points with other external angles, which is impossible.

Define the *root point* $\mathbf{r}_\mathcal{P} \in M$ to be this common landing point, and define the *wake* $W_\mathcal{P}$ to be that connected component of $\mathbb{C} \smallsetminus (\mathcal{R}_{t_-}^M \cup \mathcal{R}_{t_+}^M \cup \mathbf{r}_\mathcal{P})$ which does not contain 0. For $c \in W_\mathcal{P} \smallsetminus F_n$, it follows from the discussion above that f_c does have a repelling orbit with portrait $\mathcal{P}$, while for $c \in \mathbb{C} \smallsetminus (W_\mathcal{P} \cup F_n)$ it follows that f_c does not have any repelling orbit with portrait $\mathcal{P}$. Thus, to complete the proof of 3.1, we need only consider those f_c with c in the finite set F_n.

First suppose that some point $c_0 \in F_n \smallsetminus W_\mathcal{P}$ had a repelling orbit with portrait $\mathcal{P}$. Then any nearby parameter value would have a nearby repelling orbit with the same landing pattern for rays with angles in $\mathbf{A}_\mathcal{P}$. A priori it might seem possible that some extra ray, perhaps one landing on a parabolic orbit for f_{c_0}, might land on this same repelling orbit after perturbation. (Compare [**GM**, Fig. 12].) However, this is ruled out by 2.8. Hence all nearby parameter values must belong to $W_\mathcal{P}$, which is impossible.

Now consider parameter points $c \in W_\mathcal{P}$. I am indebted to Tan Lei for pointing out the following very elegant argument due to Peter Haïssinsky which replaces my own more complicated reasoning. As noted above, for every $c \in W_\mathcal{P}$ the rays $\mathcal{R}_{t_-}^{K(f_c)}$ and $\mathcal{R}_{t_+}^{K(f_c)}$ land at well defined periodic points of the Julia set $J(f_c)$. Let $z = z(m, t_\pm, c)$ be the unique point on the ray $\mathcal{R}_{t_\pm}^{K(f_c)}$ which has potential $G(z) = 1/m$. The functions $c \mapsto z(m, t_\pm, c)$ are evidently holomorphic; in fact $z(m,t,c) = \phi_c^{-1}\Big(\exp(2\pi i t + 1/m)\Big)$, where the function

$$\phi_c : \mathbb{C} \smallsetminus K \xrightarrow{\cong} \mathbb{C} \smallsetminus \overline{\mathbb{D}} \qquad \text{can be defined locally as} \qquad \phi_c(z) = \lim_{k \to \infty} \sqrt[2^k]{f_c^{\circ k}(z)}$$

(choosing appropriate branches of the iterated square root) and hence is holomorphic as a function of both variables. These maps $c \mapsto z(m, t_-, c)$ or $c \mapsto z(m, t_+, c)$ form a normal family throughout $W_\mathcal{P}$, since they miss the three points 0, c and ∞. Choosing a convergent subsequence and passing to the limit as $m \to \infty$, we see that the landing points also depend holomorphically on c. Since the landing points for $\mathcal{R}_{t_-}^{K(f_c)}$ and $\mathcal{R}_{t_+}^{K(f_c)}$ coincide for $c \in W_{\mathcal{P}} \setminus F_n$, it follows by continuity that they coincide for all $c \in W_{\mathcal{P}}$. It follows also that the multiplier $\lambda(c)$ of this common landing point depends holomorphically on the parameter $c \in W_{\mathcal{P}}$, with $|\lambda(c)| \geq 1$ since the landing point must be repelling or parabolic. But the absolute value of a non-zero holomorphic function cannot have a local minimum unless it is constant, so all of these landing points must actually be repelling. Together with 2.8, this completes the proof of 3.1. □

We will deal with parabolic orbits with portrait $\mathcal{P}$ in the next two sections.

4. Near Parabolic Maps

Let $\hat{c}$ be a parabolic point in parameter space. This section will study the dynamic behavior of the quadratic map f_c for c in a neighborhood of $\hat{c}$. (Compare **[DH2**, §14(CH)], **[Sh2]**.)

Let $\mathcal{O}$ be the parabolic orbit for $f_{\hat{c}}$ with period $p \geq 1$ and with representative point $\hat{z}$. Then the multiplier $\hat{\lambda} = (f_{\hat{c}}^{\circ p})'(\hat{z})$ is a primitive r-th root of unity for some $r \geq 1$. Let $\mathcal{P}$ be the associated orbit portrait, with ray period $rp \geq p$. We will first prove the following.

***Theorem 4.1* (Deformation Preserving the Orbit Portrait).** — *There exists a smooth path in parameter space ending at the parabolic point $\hat{c}$ and consisting of parameter values c with the following property: The associated map f_c has both a repelling orbit of period p and an attracting orbit of period rp. Furthermore, this repelling orbit has portrait $\mathcal{P}$, and lies on the boundary of the immediate basin for the attracting orbit. As c tends to $\hat{c}$, these two orbits both converge towards the original parabolic orbit $\mathcal{O}$.*

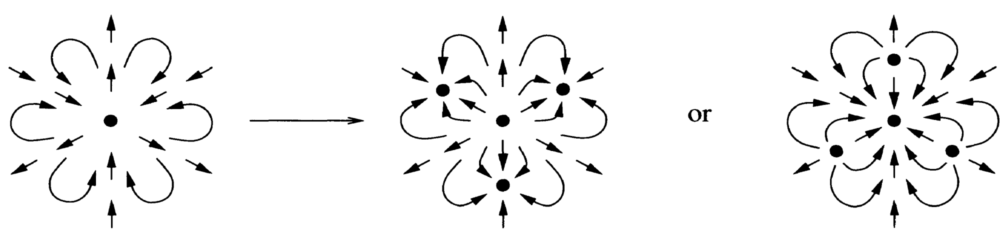


FIGURE 10. (Courtesy of S. Zakeri). The left sketch shows a parabolic fixed point with $r = 3$, the middle shows the modified version with an attracting orbit of period 3, and the right shows a modified version with an attracting fixed point. Here the arrows indicate the action of $f^{\circ 3}$.

(Compare Figure 10, middle.) The proof will depend on the following.

***Lemma 4.2* (Convenient Coordinates).** — *For any complex number λ close to $\widehat{\lambda}$ there exists at least one parameter value c close to $\widehat{c}$ and point z_λ close to $\widehat{z}$ so that z_λ is a periodic point for the map f_c with period p and with multiplier λ. Furthermore there is a local holomorphic change of coordinate $z = \phi_\lambda(w)$ with $z_\lambda = \phi_\lambda(0)$ so that the map $F = F_\lambda = \phi_\lambda^{-1} \circ f_c^{\circ p} \circ \phi_\lambda$ takes the form*

$$F(w) = \lambda\, w \; + \; R(\lambda, w)$$

for w near zero, and so that its r-th iterate takes the form

$$F^{\circ r}(w) = \phi_\lambda^{-1} \circ f_c^{\circ rp} \circ \phi_\lambda = \lambda^r w \left(1 + w^r + R'(\lambda, w)\right), \tag{2}$$

where the remainder terms R and R' satisfy $|R|$, $|R'| \leq \text{constant}\, |w|^{r+1}$ uniformly for λ in some neighborhood of $\widehat{\lambda}$ and for w in some neighborhood of zero.

(In 4.5, we will sharpen this statement by showing that the phrase "at least one" in 4.2 can be replaced by "exactly one".)

Proof of 4.2 in the Primitive Case. — First suppose that $\mathcal{P}$ is a primitive portrait, so that the multiplier $\left(f_{\widehat{c}}^{\circ p}\right)'(\widehat{z})$ is equal to $+1$ for $\widehat{z} \in \mathcal{O}$, with $r = 1$. In this case, $\widehat{z}$ is a fixed point of multiplicity two for the iterate $f_{\widehat{c}}^{\circ p}$, and splits into two nearby fixed points under perturbation. (It cannot have a higher multiplicity, since a fixed point of multiplicity $\mu > 2$ would have $\mu - 1 \geq 2$ attracting Leau-Fatou petals, each with at least one critical point in its basin, which is impossible for a quadratic map.) As c traverses a small loop around $\widehat{c}$, these two fixed points a priori may be (and in practice always will be) interchanged. However, if we loop twice around $\widehat{c}$, then each of these fixed points must return to its original position. Thus, if we introduce a new parameter u by the equation $c = \widehat{c} + u^2$, then we can choose these fixed points as holomorphic functions, $z_\iota = z_\iota(u)$ for $\iota = 1, 2$, with $z_1(0) = z_2(0) = \widehat{z}$. Evidently the u-plane is a two-fold branched cover of the c-parameter plane. Let $\lambda_\iota(u) = \left(f_c^{\circ p}\right)'\left(z_\iota(u)\right)$ be the multiplier for the orbit of z_ι, and note that $\lambda_1(0) = \lambda_2(0) = 1$. Since the holomorphic function $u \mapsto \lambda_1(u)$ cannot be constant, it takes on all values close to $+1$ as u varies through a neighborhood of 0.

Expanding the function $f_c^{\circ p}$ as a power series about its fixed point z_1, we obtain

$$f_c^{\circ p}(z_1(u) + h) \; - \; z_1(u) \; = \; \lambda_1(u)\, h \; + \; a(u)\, h^2 \; + \; (\text{higher terms in } h) \tag{3}$$

for h and u close to zero, where $c = \widehat{c} + u^2$. Here the coefficient $a(u)$ is also a holomorphic function of u, with $a(0) \neq 0$ since the fixed point multiplicity is two. It follows that $a(u) \neq 0$ for u sufficiently small. Denoting the expression (3) by $g_u(h)$, and replacing the variable $h = z - z_1$ by $w = \alpha_u\, h$ where $\alpha_u = a(u)/\lambda_1(u)$, we see easily that the function

$$F_u(w) = \alpha_u \; g_u\left(w/\alpha_u\right)$$

has the required form (2). □

Proof of 4.2 in the Satellite Case. — We now suppose that $\widehat{\lambda}$ is a primitive r-th root of unity, with $r > 1$. Then we can solve for the period p point $z = z(c)$ as a holomorphic function of c for c in some neighborhood of $\widehat{c}$, with $z(\widehat{c}) = \widehat{z}$. Hence the multiplier $\lambda(c) = (f_c^{\circ p})'(z(c))$ will also be a holomorphic function of c, taking the value $\widehat{\lambda} \in \sqrt[r]{1}$ when $c = \widehat{c}$. Similarly $\lambda(c)^r$ is a holomorphic function, taking the value $\lambda(\widehat{c})^r = 1$ when $c = \widehat{c}$. This function $\lambda(c)^r$ clearly cannot be constant, so it takes all values close to $+1$ as c varies through a neighborhood of $\widehat{c}$.

We will construct a sequence of holomorphic changes of variable which conjugate the map $z \mapsto f_c^{\circ p}(z)$ in a neighborhood of $z = z(c)$ to maps $h \mapsto g_{c,k}(h)$ in a neighborhood of $h = 0$, where $1 \le k \le r$, so that

$$g_{c,k}(h) = \lambda(c)h\big(1 + a_k(c)h^k + (\text{higher terms in } h)\big)$$

for some constant $a_k(c)$. Here c can be any point in some neighborhood of $\widehat{c}$. To begin the construction, let

$$g_{c,1}(h) = f^{\circ p}(z(c) + h) \; - \; z(c).$$

This certainly has the required properties. Now inductively set

$$g_{c,k+1}(h) = \phi^{-1} \circ g_{c,k} \circ \phi(h) \qquad \text{where} \qquad \phi(h) = h + bh^{k+1}$$

for $1 \le k < r$. We claim that the constant $b = b(c)$ can be uniquely chosen so that $g_{c,k+1}$ will have the required form. In fact a brief computation shows that

$$g_{c,k+1}(h) = \lambda h\big(1 + (a + b - \lambda^k b)h^k + (\text{higher terms}).$$

But $\lambda^k \ne 1$ since λ is close to $\widehat{\lambda}$, which is a primitive r-th root of unity with $1 \le k < r$. Hence there is a unique choice of b so that $a + b - \lambda^k b = 0$, as required.

In particular, pushing this argument as far as possible, we can take $k = r$ and replace $f_c^{\circ p}$ near $z = z(c)$ by $g_{c,r}(h) = \lambda h\big(1 + ah^r + \cdots\big)$ near $h = 0$. Hence we can replace $f_c^{\circ rp}$ near $z(c)$ by

$$g_{c,r}^{\circ r}(h) = \lambda^r\, h\big(1 + a'\, h^r + \;(\text{higher terms})\big),$$

where computation shows that $a' = \big(1 + \lambda^r + \lambda^{2r} + \cdots + \lambda^{(r-1)r}\big)a$. Here the coefficient a' of h^r must be non-zero when $\lambda = \widehat{\lambda}$, and hence for λ close to $\widehat{\lambda}$. For otherwise, the Leau-Fatou flowers around the points of the parabolic orbit would give rise to more than one periodic cycle of attracting petals for f_c. This is impossible, since each such cycle must contain a critical point, and a quadratic polynomial has only one critical point. Finally, after a scale change, replacing $g_{c,r+1}(h)$ by $F_c(w) = \alpha_c\, g_{c,r+1}^{\circ r}(w/\alpha_c)$ for suitably chosen α_c, we obtain simply

$$F_c^{\circ r}(w) = \lambda^r w\big(1 + w^r + (\text{higher terms in } w)\big),$$

as required. $\square$

Proof of 4.1. — First note that we can choose a smooth path in parameter space so that the multiplier λ^r of Lemma 4.2 is real and belongs to some interval $(1,\, 1+\eta)$. This follows easily from the fact that λ is a non-constant holomorphic function of c in the case $r > 1$, or of $u = \sqrt{c - \widehat{c}}$ in the case $r = 1$. Note that the map $F^{\circ r}$ of 4.2 satisfies

$$|F^{\circ r}(w)| = \lambda^r \cdot |w| \cdot \big(1 + \mathrm{Re}(w^r) + \text{ (higher terms)}\big) \tag{4}$$

and

$$\arg(F^{\circ r}(w)) = \arg(w) + \mathrm{Im}(w^r) + \text{ (higher terms)} \tag{5}$$

whenever λ^r is real and positive; and note also that $F^{\circ r}$ has a locally defined holomorphic inverse of the form

$$F^{-r}(w) = w\,\big(1\ -\ w^r/\lambda^{2r} + \text{ (higher terms)}\big)/\lambda^r,$$

which satisfies

$$|F^{-r}(w)| = |w|\,\big(1 - \mathrm{Re}(w^r)/\lambda^{2r} + \text{ (higher terms)}\big)/\lambda^r \tag{4$'$}$$

and

$$\arg(F^{-r}(w)) = \arg(w)\ -\ \mathrm{Im}(w^r)/\lambda^{2r} + \text{ (higher terms)}. \tag{5$'$}$$

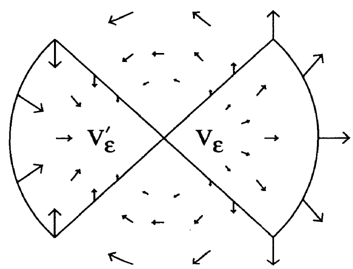


FIGURE 11. A repelling petal V_ε and attracting petal V'_ε for the map $F(w) \approx w + w^2$ (illustrating the primitive case, before perturbation).

As a representative repelling petal for $F^{\circ r}$ let us choose a small wedge shaped region V_ε described in polar coordinates by setting $w = \rho\, e^{2\pi i t}$ with $0 \le \rho \le \varepsilon$ and $|t| < 1/(8r)$. (Compare Figure 11 for the case $r = 1$.) If $\lambda^r \ge 1$ with λ^r sufficiently close to 1, it follows easily from (4$'$) and (5$'$) that V_ε maps into itself under F^{-r}, with all orbits converging towards the boundary fixed point at $w = 0$. If a dynamic ray for $f_{\widehat{c}}$ lands at $\widehat{z}$, then it must land through one of the r repelling petals, for example through the image of V_ε in the z-plane. For c sufficiently close to $\widehat{c}$, this image must still contain a full segment, from some point z to $f_c^{\circ rp}(z)$, of the perturbed ray, hence this perturbed ray must still land at the repelling point which corresponds to $w = 0$.

Note that no new rays land at this point, after perturbation. There are only finitely many rays which have period p. But every dynamic ray of period p for $f_{\widehat{c}}$ with angle not in the set $\mathbf{A}_{\mathcal{P}}$ of angles for $\mathcal{P}$ must land on some disjoint repelling point, and this condition will be preserved under perturbation. Thus the perturbed orbit, for $\lambda^r > 1$, still has portrait $\mathcal{P}$.

As an attracting petal for $F^{\circ r}$ we can choose the set $V'_\varepsilon = e^{\pi i/r}V_\varepsilon$ consisting of all $w = \rho e^{2\pi i t}$ with $0 \le \rho \le \varepsilon$ and $\frac{3}{8r} \le t \le \frac{5}{8r}$. If $\lambda^r > 1$ with λ^r close to 1, then using (4) and (5) we can check that $F^{\circ r}$ maps V'_ε into itself. However, the origin is a repelling point, so orbits cannot converge to it. In fact, if K is the compact set obtained from V'_ε by removing a very small neighborhood of the origin, then $F^{\circ r}$ maps K into its own interior. It follows easily that all orbits in $V'_\varepsilon \smallsetminus \{0\}$ converge to an interior fixed point. This must be a strictly attracting point, and must correspond to an attracting orbit of period rp for the map f_c. □

***Corollary 4.3* (Parabolic Points as Root Points).** — *If $f_{\widehat{c}}$ has a parabolic orbit whose portrait $\mathcal{P}$ is non-trivial, then $\widehat{c}$ must be equal to the root point $\mathbf{r}_{\mathcal{P}}$ of the $\mathcal{P}$-wake.*

Note. — The hypothesis that $\mathcal{P}$ is non-trivial is actually redundant. (See 4.8.) It will be shown in 5.4 that every parabolic point is the root point of only one wake, so that the root point of the $\mathcal{P}$-wake always has portrait equal to $\mathcal{P}$.

Proof of 4.3. — Since $f_{\widehat{c}}$ has a parabolic orbit with portrait $\mathcal{P}$, it certainly cannot have a *repelling* orbit with portrait $\mathcal{P}$. Hence it cannot be inside the $\mathcal{P}$-wake by 3.1. On the other hand, by 4.1 it must belong to the boundary of the $\mathcal{P}$-wake. By construction, the root point $\mathbf{r}_{\mathcal{P}}$ is the only boundary point of $W_{\mathcal{P}}$ which belongs to the Mandelbrot set. □

Here is a complementary statement to 4.1, in the case $r > 1$.

***Lemma 4.4* (A Deformation Breaking the Portrait).** — *Under the hypothesis of 4.1, there also exists a smooth path of parameter values c, converging to $\widehat{c}$, so that each f_c has an attracting orbit of period p, and a repelling orbit of period rp which lies on the boundary of its immediate basin. Furthermore, the dynamic rays with angles in $\mathbf{A}_{\mathcal{P}} = A_1 \cup \cdots \cup A_p$ all land on this repelling orbit.*

(Compare Figure 10, right.) For such values of c (still assuming that $r > 1$), it follows that there is *no* periodic orbit with portrait $\mathcal{P}$. Together with 4.1, this gives an alternative proof that $\widehat{c}$ is on the boundary of the $\mathcal{P}$-wake.

Proof. — The proof of 4.4 is completely analogous to the proof of 4.1, and will be left to the reader: One simply deforms so that $\lambda^r < 1$, instead of $\lambda^r > 1$. □

The following assertion helps to make the statement of 4.2 more precise.

***Lemma 4.5* (Local Uniqueness).** — *Under the hypothesis of 4.2, there exist unique single valued functions $c = c(\lambda)$ and $z = z(\lambda)$, defined and holomorphic for λ in a neighborhood of $\widehat{\lambda}$, so that $z(\lambda)$ is a periodic point of period p and multiplier λ for the* *map $f_{c(\lambda)}$, with $\widehat{c} = c(\widehat{\lambda})$ and $\widehat{z} = z(\widehat{\lambda})$. This function $c(\lambda)$ is univalent in the satellite case, but has a simple critical point at $\widehat{\lambda}$ in the primitive case.*

The implications of this lemma for the geometry of the Mandelbrot set will be described in 6.1 and 6.2.

Proof of 4.5. — First consider the satellite case, with $\widehat{\lambda} \neq 1$. Then clearly the period p orbit and its multiplier $\lambda(c)$ depend smoothly on c throughout some neighborhood of $\widehat{c}$. We will show that the derivative $d\lambda/dc$ is non-zero at $\widehat{c}$. For otherwise, we could write

$$\lambda^r(c) = 1 + a(c - \widehat{c})^k + \text{ (higher terms)}$$

with $k \geq 2$. Hence we could vary c from $\widehat{c}$ in two or more different directions so that $\lambda^r > 1$ and in two or more intermediate directions so that $\lambda^r < 1$. The former points would be within the $\mathcal{P}$-wake and the later points would be outside it; but this configuration is impossible by 3.1. Thus $d\lambda/dc \neq 0$, and it follows by the Inverse Function Theorem that the inverse mapping $\lambda \mapsto c(\lambda)$ is well defined and holomorphic throughout a neighborhood of $\widehat{\lambda}$, as required.

In the primitive case, the situation is different, but the proof is similar. In this case, setting $c = \widehat{c} + u^2$, we must express the multiplier λ_1 for one of the two nearby period p points as a holomorphic function of u, and show that the derivative $d\lambda_1/du$ is non-zero at $u = 0$. Otherwise, if the derivative $d\lambda_1(u)/du$ were equal to zero for $u = 0$, then we could write

$$\lambda_1(u) = 1 + a\,u^k + \text{ (higher terms)}$$

for some $k \geq 2$. It would follow that we could vary u from 0 in two or more different directions so that $\lambda_1 > 1$ and in two or more separating directions so that $\lambda_2 > 1$. All of these points would be within the $\mathcal{P}$-wake, but the rays landing on the periodic point z_1 would have to jump discontinuously so as to land on z_2 as we pass from $\lambda_1 > 1$ to $\lambda_2 > 1$, and such points of discontinuity must be outside the $\mathcal{P}$-wake. Even allowing for the fact that the u-plane is a two-fold covering of the c-plane, such a configuration is incompatible with 3.1. Therefore, λ_1 and u must determine each other holomorphically in a neighborhood of $\widehat{\lambda} \leftrightarrow 0$. In particular, it follows that the parameter value $c = \widehat{c} + u^2$ can be expressed as a holomorphic function of λ_1, with a simple critical point at $\lambda_1 = \widehat{\lambda}$. □

To conclude this section, we will prove that the portrait of a parabolic periodic point is always non-trivial. We will use a somewhat simplified form of the Hubbard tree construction to show that every parabolic orbit with ray period $rp \geq 2$ must have portrait with valence $v \geq 2$. First some general remarks about locally connected subsets of the plane.

***Lemma 4.6* (A Canonical Retraction).** — *Let $K \subset \mathbb{C}$ be compact, connected, locally connected, and full, and let U be a connected component of the interior of K. Then* *the closure $\overline{U}$ is homeomorphic to the closed unit disk, and there is a unique retraction ρ_U from $\mathbb{C}$ onto $\overline{U}$ which carries each external ray, and also each connected component of the complement $K \smallsetminus \overline{U}$, to a single point of the circle ∂U. There are at least two distinct external rays landing at a point $z_0 \in \partial U$ if and only if $K \smallsetminus \{z_0\}$ is disconnected, or if and only if there is some connected component X of $K \smallsetminus \overline{U}$ with $\rho_U(X) = \{z_0\}$.*

Proof. — (Compare [**D5**].) The statement that $\overline{U}$ is a disk follows easily from well known results of Carathéodory. Furthermore, according to Carathéodory, there is a unique retraction from $\mathbb{C}$ onto K which maps each external ray to its landing point. Composing this with the retraction $K \to \overline{U}$ which maps each component X of $K \smallsetminus \overline{U}$ to the unique intersection point $z_0 \in \overline{X} \cap \overline{U}$, we obtain the required retraction ρ_U.

For any such X, note that there must be at least one maximal open interval of angles t such that the ray $\mathcal{R}_t^K$ lands in X. The endpoints of such a maximal interval are the angles for the required pair of rays landing on z_0. Conversely, if there were two rays landing on z_0 but no component X attached in between, then there would be an entire open interval of angles t so that $\mathcal{R}_t^K$ lands at z_0. But this is impossible by a classical theorem of F. and M. Riesz. (See for example [**M2**, App. A].) □

In particular, let $K = K(f)$ be the filled Julia set for a hyperbolic quadratic polynomial. (We are actually interested in the parabolic case, but will work first with the hyperbolic case, since that will suffice for our purposes, and since it is much easier to prove local connectivity in the hyperbolic case.)

***Lemma 4.7* (The Dynamic Root Point).** — *Suppose that $f = f_c$ has an attracting orbit of period $n \geq 2$. Let K be its filled Julia set, and let U_0 and $U_1 \subset K$ be the Fatou components containing the critical point* 0 *and the critical value c respectively. Then the canonical retraction $\rho_{U_1} : \mathbb{C} \to \overline{U}_1$ carries the component U_0 to the unique point $\mathbf{r}_c \in \partial U_1$ which is fixed by $f^{\circ n}$. Hence at least two dynamic rays land at this point.*

(See for example Figures 1, 6.) Following Schleicher, I will call $\mathbf{r}_c$ the *dynamic root point* for the Fatou component U_1.

Proof. — Let $U_0 \to U_1 \xrightarrow{\sim} U_2 \xrightarrow{\sim} \cdots \xrightarrow{\sim} U_n = U_0$ be the Fatou components containing the critical orbit. Then $f^{\circ n}$ maps each circle ∂U_j onto itself by an expanding map of degree two. Hence there is a canonical homeomorphism $a_j : \partial U_j \to \mathbb{R}/\mathbb{Z}$ which conjugates $f^{\circ n}$ to the angle doubling map on the standard circle. For each $z \in \mathbb{C} \smallsetminus U_j$, the image $a_j(\rho_{U_j}(z))$ will be called the *internal angle* of the point z with respect to U_j. The map f from ∂U_j to ∂U_{j+1} preserves the internal angles of boundary points for $0 < j < n$, but doubles them for the case $j = 0$ of the critical component.

Define the *t-wake* $L_t(U_j)$ to be the set of all $z \in \mathbb{C} \smallsetminus U_j$ with $a_j(\rho_{U_j}(z)) = t \in \mathbb{R}/\mathbb{Z}$. These wakes are pairwise disjoint sets with union equal to $\mathbb{C} \smallsetminus U_j$. In general f maps to t-wake of U_j homeomorphically onto the t-wake of U_{j+1} for $0 < j < n$, and onto the $2t$-wake of U_{j+1} when $j = 0$. However, there is one exceptional value of t for each

U_j with $0 < j < n$. Namely, if the wake $L_t(U_j)$ contains the critical component U_0 then it certainly cannot map homeomorphically, and its image may be much larger than $L_t(U_{j+1})$.

Let $\mathcal{A}_j \subset \mathbb{R}/\mathbb{Z}$ be the finite set consisting of all angles $t \in \mathbb{R}/\mathbb{Z}$ such that the wake $L_t(U_j)$ contains one of the components U_k (where necessarily $j \neq k$). Then it follows that $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \cdots \subset \mathcal{A}_n$ and $2\mathcal{A}_n = \mathcal{A}_0$. On the other hand, since K is full, the various U_j must be connected together in a tree-like arrangement (the *Hubbard tree*). There cannot be any cycles. Hence at least one of the $\mathcal{A}_i$ must consist of a single angle. It follows easily that $\mathcal{A}_1 = \{0\}$, and the conclusion follows. □

***Corollary 4.8* (Parabolic Orbit Portraits are Non-Trivial).** — *If c is any parabolic point of the Mandelbrot set other than $c = 1/4$, and if $\mathcal{O}$ is the parabolic orbit for f_c, then at least two dynamic rays land on each point of $\mathcal{O}$.*

(This is just a restatement of Theorem 1.4 of §1.)

Proof. — In the satellite case this is trivially true, while in the primitive case it follows from 4.7, using 4.1 to pass from the parabolic to the hyperbolic case. This completes the proof of Theorem 1.4. □

5. The Period n Curve in (Parameter×Dynamic) Space

It is convenient to define a sequence of numbers $\nu_2(n)$ inductively by the formula

$$2^k = \sum_{n|k} \nu_2(n), \qquad \text{or} \qquad \nu_2(k) = \sum_{n|k} \mu(k/n) 2^n,$$

to be summed over all divisors $n \geq 1$ of k, where $\mu(k/n) \in \{\pm 1, 0\}$ is the Möbius function. In fact we will be mainly interested in the quotients $\nu_2(n)/2$ and $\nu_2(n)/n$. The first few values are

n	1	2	3	4	5	6	7	8	9	10
$\nu_2(n)/2$	1	1	3	6	15	27	63	120	252	495
$\nu_2(n)/n$	2	1	2	3	6	9	18	30	56	99 .

Define the *period n curve* $\mathrm{Per}_n \subset \mathbb{C}^2$ to be the locus of zeros of the polynomial $Q_n(c,z)$ which is defined by the formula

$$f_c^{\circ k}(z) - z = \prod_{n|k} Q_n(c,z), \quad \text{or} \quad Q_k(c,z) = \prod_{n|k} \left(f_c^{\circ k}(z) - z\right)^{\mu(k/n)},$$

taking the product over all divisors n of k. For example,

$$Q_1(c,z) = z^2 + c - z, \qquad Q_2(c,z) = \frac{(z^2+c)^2 + c - z}{z^2 + c - z} = z^2 + z + c + 1.$$

Note that each point $(c,z) \in \mathrm{Per}_n$ determines a periodic orbit

$$z = z_0 \mapsto z_1 \mapsto \cdots \mapsto z_n = z_0$$

for the map f_c. Let $\lambda_n = \lambda_n(c, z) = \partial f_c^{\circ n}(z)/\partial z = 2^n z_1 \cdots z_n$. For a generic choice of c, this orbit has period exactly n, and λ_n is the multiplier. However, if z is a parabolic periodic point for f_c with ray period $n = rp > p$, then (c, z) belongs both to Per_n with $\lambda_n = 1$, and to Per_p with $\lambda_p \in \sqrt[r]{1}$. (In fact, the two curves Per_n and Per_p intersect transversally at (c, z).)

Remarks. — Compare **[M4]** for a somewhat analogous discussion for cubic polynomials. The fact that Q_n is really a polynomial can be verified by expressing $f_c^{\circ k}(z) - z$ as a product of irreducible polynomials, and checking that each of these irreducible factors has a well defined period n dividing k. The factors are all distinct since $\partial(f_c^{\circ j}(z) - z)/\partial z \neq 0$ at every zero of this polynomial when $|c|$ is large. It is shown in **[Bou]**, and also in **[S1]**, **[LS]**, that the algebraic curve Per_n (or the polynomial Q_n) is actually irreducible; however, we will not make any use of that fact.

***Lemma 5.1* (Properties of the Period n Curve).** — *This algebraic curve $\mathrm{Per}_n \subset \mathbb{C}^2$ is non-singular. The projection $(c, z) \mapsto c$ is a proper map of degree $\nu_2(n)$ from Per_n to the parameter plane, while the projection $(c, z) \mapsto z$ is a proper map of degree $\nu_2(n)/2$ to the dynamic plane. Finally, the function $(c, z) \mapsto \lambda_n(c, z)$ is a proper map of degree $n\nu_2(n)/2$ to the λ_n-plane.*

Note that the cyclic group of order n, which we will denote by $\mathbb{Z}_n$, acts on Per_n, a generator carrying (c, z) to $(c,\ f_c(z))$.

***Lemma 5.2* (Properties of $\mathrm{Per}_n / \mathbb{Z}_n$).** — *The quotient $\mathrm{Per}_n / \mathbb{Z}_n$ is a smooth algebraic curve consisting of all pairs $(c, \mathcal{O})$ where $\mathcal{O}$ is a periodic orbit for f_c which is either non-parabolic of period n, or parabolic with attracting petals of period n. At any point where $\lambda_n \neq 1$, the coordinate c can be used as local uniformizing parameter, while in a neighborhood of a point with $\lambda_n = 1$, the multiplier $\lambda_n = \lambda_n(c, z)$ serves as a local uniformizing parameter for this curve. The projection maps $(c, \mathcal{O}) \mapsto c$ and $(c, \mathcal{O}) \mapsto \lambda_n$ are proper, with degrees $\nu_2(n)/n$ and $\nu_2(n)/2$ respectively.*

The proof that Per_n and $\mathrm{Per}_n / \mathbb{Z}_n$ are non-singular will be divided into three cases, as follows.

Generic Case. — First consider a point $(\widehat{c}, \widehat{z}) \in \mathrm{Per}_n$ with $\lambda_n(\widehat{c}, \widehat{z}) \neq 1$. Then, by the Implicit Function Theorem, we can solve the equation $f_c^{\circ n}(z) = z$ locally for z as a smooth function of c. It follows that both of the curves Per_n and $\mathrm{Per}_n / \mathbb{Z}_n$ are locally smooth, with c as local uniformizing parameter.

Primitive Parabolic Case. — Now consider a point $(\widehat{c}, \widehat{z}) \in \mathrm{Per}_n$ with $\lambda_n(\widehat{c}, \widehat{z}) = 1$, where $\widehat{z}$ has period exactly n under $f_{\widehat{c}}$. According to the proof of 4.5, if we set $c = \widehat{c} + u^2$, then both z and $\lambda_n = \lambda_n(c, z)$ can be expressed locally as smooth functions of u with $d\lambda_n/du \neq 0$. It follows that both Per_n and $\mathrm{Per}_n / \mathbb{Z}_n$ are locally smooth at this point, and that we can use either u or λ_n as local uniformizing parameter.

(Similarly $dz/du \neq 0$, so we could use z as local uniformizing parameter for Per_n. However dc/du is zero when $u = 0$, so c cannot be used as local parameter.)

Satellite Parabolic Case. — Again suppose that $\lambda_n(\widehat{c}, \widehat{z}) = 1$, but now assume that the period p of $\widehat{z}$ is strictly less than the ray period $n = rp$. For c near $\widehat{c}$, let $z = z(c)$ be the equation of the unique period p point near $\widehat{z}$. Using the change of variable $w = \alpha(z - z(c)) +$ (higher terms) of 4.2, the map $f_c^{\circ n}$ corresponds to

$$w \mapsto F^{\circ r}(w) = \lambda^r w\big(1 + w^r + \text{ (higher terms)}\big), \tag{6}$$

where $\lambda = \lambda(w)$ is the multiplier of this period p orbit. The equation for a fixed point is $w = \lambda^r w\,(1 + w^r +$ (higher terms)). Dividing by w (since we want the fixed point with $w \neq 0$ or with $z \neq z(c)$), this becomes

$$1 = \lambda^r(1 + w^r + \text{ (higher terms)}) \qquad \text{or} \qquad \lambda^r = 1 - w^r + \text{ (higher terms)}.$$

Thus we can express λ as a holomorphic function of w, with a critical point at $w = 0$. Therefore, by 4.5, we can also express c as a holomorphic function of w. Since w is defined as a holomorphic function of z and c with $\partial w/\partial z \neq 0$, it follows that Per_n is locally smooth with local uniformizing parameter z or w.

Now note that there is a unique local change of coordinate $w \mapsto \phi(w)$ with $\phi'(0) = 1$ so that $\lambda^r = 1 - \phi(w)^r$. Since the expression $\phi(w)^r$ is invariant under the $\mathbb{Z}_n$ action of Per_n, it follows easily that this action can be described by the formula $\phi(w) \mapsto \widehat{\lambda}\phi(w)$. It follows that $\phi(w)^r = 1 - \lambda^r$ is a local uniformizing parameter for the quotient curve $\mathrm{Per}_n/\mathbb{Z}_n$. Therefore, either λ or c can also be taken as local uniformizing parameter. In particular, it follows that the multiplier λ_n of the period $n = rp$ orbit can be expressed as a smooth function of the multiplier $\lambda = \lambda_p$ of the period p orbit. Note that

$$d\lambda_n/d(\lambda^r) = -r \tag{7}$$

at the parabolic point. (Compare **[CM**, (4.3)**]**.) This can be verified by direct computation from (6), or by using the holomorphic fixed point formula **[M2]** for the function $f_c^{\circ n}$ to show that the expression

$$\frac{r}{1 - \lambda_n} + \frac{1}{1 - \lambda^r}$$

depends smoothly on the parameter c throughout some neighborhood of the parabolic point. Therefore λ_n can also be used as local uniformizing parameter for $\mathrm{Per}_n/\mathbb{Z}_n$.

The degrees of the various projection maps can easily be computed algebraically, by counting solutions to the appropriate polynomial equations. Here is a more geometric argument, which also provides a quite explicit description of the ends of the curve Per_n, and hence proves that these mappings are proper. Let us consider the limiting case as $|c| \to \infty$. Setting $c = -v^2$ with $|v| > 2$, let $\pm\Delta$ be the open disk of radius 1 centered at $\pm v$. It is not difficult to check that both Δ and $-\Delta$ map holomorphically onto a disk $f(\Delta)$ which contains $\overline{\Delta} \cup (-\overline{\Delta})$. The (filled) Julia set K can then be described explicitly as follows. Given an arbitrary sequence of signs $\varepsilon_0, \varepsilon_1, \ldots$, there is one an only one orbit $z_0 \mapsto z_1 \mapsto \cdots$ in K with $z_j \in \varepsilon_j \Delta$ for every $j \geq 0$. This is proved using the Poincaré metric for the inverse maps $f(\Delta) \to \pm\Delta \subset f(\Delta)$. In particular, the number of solutions of period n is equal to the number of sign sequences of period n, which is easily seen to be $\nu_2(n)$. Thus the degree of the projection to the c-plane is $\nu_2(n)$. It follows also that the product $z_1 \cdots z_n = \lambda/2^n$ is given asymptotically by

$$\lambda/2^n \sim \pm z^n \sim \pm v^n = \pm(-c)^{n/2} \qquad \text{as} \qquad |v| \to \infty.$$

Thus the degree of the projection to the λ-plane is n times the degree of the projection to the z-plane, and is $n/2$ times the degree of the projection to the c-plane. □

Thus we have a diagram of smooth algebraic curves and proper holomorphic maps with degrees as indicated:

$$\begin{array}{ccccc}
\mathrm{Per}_n & \xrightarrow{n} & \mathrm{Per}_n/\mathbb{Z}_n & \xrightarrow{\nu_2(n)/2} & \lambda_n\text{-plane} \\
\downarrow{\scriptstyle \nu_2(n)/2} & & \downarrow{\scriptstyle \nu_2(n)/n} & & \\
z\text{-plane} & & c\text{-plane} & &
\end{array}$$

For a generic choice of c, it follows that the map f_c has exactly $\nu_2(n)/n$ periodic orbits of period n, while for generic choice of λ_n there are exactly $\nu_2(n)/2$ pairs $(c, \mathcal{O})$ consisting of a parameter value c and a period n *orbit* of multiplier λ_n for the map f_c. The discussion shows that the correspondence $(c, \mathcal{O}) \mapsto (c, \lambda_n)$ yields a smooth immersion of $\mathrm{Per}_n/\mathbb{Z}_n$ into $\mathbb{C}^2$. (Caution: Presumably some f_c may have two different period n orbits with the same multiplier, so this immersion may have self-intersections.)

***Corollary 5.3* (Counting Parabolic Points).** — *The number of parabolic points in the Mandelbrot set with ray period $rp = n$ is equal to $\nu_2(n)/2$.*

Proof. — This is the same as the number of points in the pre-image of $+1$ under the projection $(c, \mathcal{O}) \mapsto \lambda_n(c, \mathcal{O})$ from $\mathrm{Per}_n/\mathbb{Z}_n$ to the λ_n-plane. According to 5.2, the degree of this projection is $\nu_2(n)/2$, and $+1$ is a regular value. The conclusion follows. □

We are now ready to prove the main results, as stated in §1.

Corollary 5.4. — *There are exactly two parameter rays which angles which are periodic under doubling landing at each parabolic point $\widehat{c} \neq 1/4$. Hence distinct wakes have distinct root points; and for each non-trivial portrait $\mathcal{P}$, the root point of the $\mathcal{P}$-wake has a parabolic orbit with portrait $\mathcal{P}$.*

(For angles which are not periodic, compare 9.4.)

Corollary 5.5. — *Every parameter ray $\mathcal{R}_t^M$ whose angle has period $n \geq 2$ under doubling forms one of the two boundary rays for one and only one wake $W_{\mathcal{P}}$, where $\mathcal{P}$ is some portrait with ray period n.*

Proof of 5.4 and 5.5. — According to 5.3, the number of parabolic points $\widehat{c}$ with ray period $n \geq 2$ is equal to $\nu_2(n)/2$, and according to Theorem 1.4 each such point is the landing point of at least two rays, which necessarily have ray period n. Thus altogether there are at least $\nu_2(n)$ distinct rays of period n. On the other hand, since the map $t \mapsto 2^n t \pmod{\mathbb{Z}}$ has $2^n - 1$ fixed points, it follows inductively that the number of angles with period exactly $n \geq 2$ is precisely equal to $\nu_2(n)$. Thus there cannot be more than two rays landing at any such point $\widehat{c}$. It follows that $\widehat{c}$ is the root point of at most one wake. For if $\widehat{c}$ were the root point of two different wakes, then (even if they shared a boundary ray) it would be the landing point for at least three different parameter rays. Using 4.3, it now follows that each such $\widehat{c}$ is the root point $\mathbf{r}_{\mathcal{P}}$ for exactly one wake $W_{\mathcal{P}}$, and furthermore that each $f_{\mathbf{r}_{\mathcal{P}}}$ has a parabolic orbit with portrait $\mathcal{P}$.

Here we have assumed that $n \geq 2$. However, for $n = 1$ there is clearly just one parameter ray $\mathcal{R}_0^M = (1/4, \infty)$ which is fixed under doubling, and its landing point $\widehat{c} = 1/4$ is the unique parabolic point with ray period $n = 1$. This completes the proof of 5.4 and 5.5. Clearly Theorems 1.2 and 1.5, as stated in §1, follow immediately. □

To conclude this section, here is a more explicit description of the first few period n curves:

Period 1. — The curve $\mathrm{Per}_1 = \mathrm{Per}_1/\mathbb{Z}_1 \cong \mathbb{C}$ can be identified with the λ_1-plane. It is a 2-fold branched cover of the c-plane, ramified at the root point $r_{\{\{0\}\}} = 1/4$, and can be described by the equations $z = \lambda_1/2$, $c = z - z^2$. Note that the unit disk $|\lambda_1| < 1$ in the λ_1-plane maps homeomorphically onto the region bounded by the cardioid in the c-plane.

Period 2. — The quotient $\mathrm{Per}_2/\mathbb{Z}_2 \cong \mathbb{C}$ can be identified either with the λ_2-plane or with the c-plane, where $\lambda_2 = 4(1+c)$. The curve $\mathrm{Per}_2 \cong \mathbb{C}$ is a 2-fold branched cover with coordinate z, branched at the point $\lambda_2 = 1$ which corresponds to the period 2 root point $c = r_{\mathcal{P}} = -3/4$ with portrait $\mathcal{P} = \{\{1/3, 2/3\}\}$. It is described by the equation $z^2 + z + (c+1) = 0$, with $\mathbb{Z}_2$-action $z \leftrightarrow f_c(z) = -z - 1$.

Period 3. — (See **[GF]**.) The quotient $\mathrm{Per}_3/\mathbb{Z}_3 \cong \mathbb{C}$ can be identified with a 2-fold branched cover of the c-plane, branched at the root point $r_{\mathcal{P}} = -7/4$ of the real period 3 component, where $\mathcal{P} = \{\{3/7, 4/7\}, \{6/7, 1/7\}, \{5/7, 2/7\}\}$. If we choose a parameter u on this quotient by setting $c = -(u^2 + 7)/4$, then computation shows that the multiplier is given by the cubic expression $\lambda_3 = u^3 - u^2 + 7u + 1$. The curve Per_3 itself is conformally isomorphic to a thrice punctured Riemann sphere. It can be described as a 3-fold cyclic branched cover of this u-plane, branched with ramification index 3 at the two points $u = (1 \pm \sqrt{-27})/2$ where $\lambda_3 = 1$.

6. Hyperbolic Components

By definition, a *hyperbolic component* H of *period* n in the Mandelbrot set is a connected component of the open set consisting of all parameter values c such that f_c has a (necessarily unique) attracting orbit of period n. We will first study the geometry of a hyperbolic component near a parabolic boundary point.

***Lemma 6.1* (Geometry near a Satellite Boundary Point).** — *Let $\widehat{c}$ be a parabolic point with orbit portrait $\mathcal{P}$ having ray period $rp > p$. Then $\widehat{c}$ lies on the boundary of exactly two hyperbolic components. One of these has period rp and lies inside the $\mathcal{P}$-wake, while the other has period p and lies outside the $\mathcal{P}$-wake. Locally the boundaries of these components are smooth curves which meet tangentially at $\widehat{c}$.*

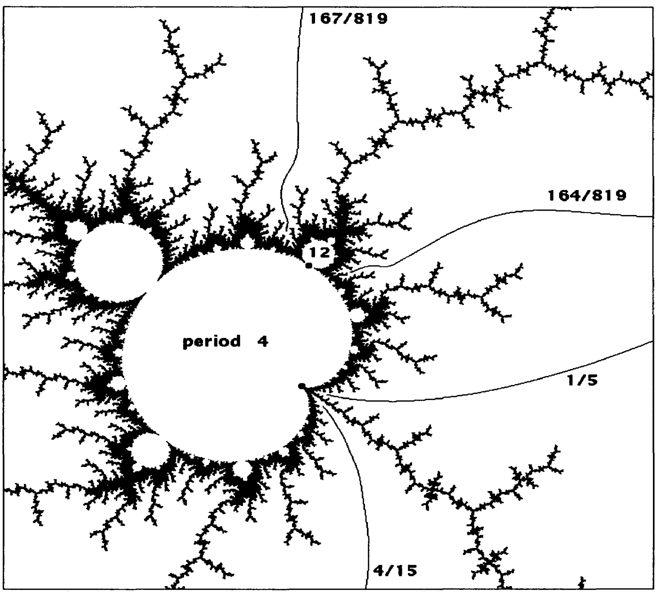


FIGURE 12. Detail of the Mandelbrot boundary, showing the rays landing at the root points of a primitive period 4 component and a satellite period 12 component.

Proof. — According to 4.1, $\widehat{c}$ lies on the boundary of a hyperbolic component H_{rp} of period rp which lies inside the $\mathcal{P}$-wake, while according to 4.4 it lies on the boundary of a component H_p of period p which lies outside the $\mathcal{P}$-wake. Let $\mathcal{O}_{rp}$ and $\mathcal{O}_p$ be the associated periodic orbits, with multipliers λ_{rp} and λ_p. According to 4.5, the multiplier λ_p can be used as a local uniformizing parameter for the c-plane near $\widehat{c}$. Therefore the boundary ∂H_p, with equation $|\lambda_p| = 1$, is locally smooth. Similarly, it follows from equation (7) of §5, that we can take λ_{rp} as local uniformizing parameter, so the locus $|\lambda_{rp}| = 1$ is also locally smooth. These two boundary curves are necessarily tangent to each other since the two hyperbolic components cannot overlap, or by direct computation from (7).

To see that there are no other components with $\widehat{c}$ as boundary point, first note that all periodic orbits for the map $f_{\widehat{c}}$, other than its designated parabolic orbit, must be strictly repelling. For any orbit with multiplier $|\lambda| \leq 1$ must either attract the critical orbit (in the attracting or parabolic case) or at least be in the ω-limit set of the critical orbit (in the Cremer case), or have Fatou component boundary in this ω-limit set (in the Siegel disk case). Since the unique critical orbit converges to the parabolic orbit, all other periodic orbits must be repelling.

Now choose some large integer N. If we choose c sufficiently close to $\widehat{c}$, then all repelling periodic orbits of period $\leq N$ for $f_{\widehat{c}}$ will deform to repelling periodic orbits of the same period for f_c. Thus any non-repelling orbit of period $\leq N$ for f_c must be one of the two orbits $\mathcal{O}_p$ and $\mathcal{O}_{rp}$ which arise from perturbation of the parabolic orbit. In other words, any hyperbolic component H' of period $\leq N$ which intersects some small neighborhood of $\widehat{c}$ must be either H_p or H_{rp}. In particular, any hyperbolic component which has $\widehat{c}$ as boundary point must coincide with either H_p or H_{rp}. □

By definition, the component H_{rp} is a *satellite* of H_p, attached at the parabolic point $\widehat{c}$. (It follows from (7) that $|d\lambda_{rp}/d\lambda_p| = r^2$ at $\widehat{c}$, so to a first approximation the component H_p is r^2 times as big as its satellite H_{rp}. Compare **[CM]**.)

Lemma 6.2 **(Geometry near a Primitive Boundary Point).** — *If the portrait $\mathcal{P}$ of the parabolic point $\widehat{c}$ has ray period $rp = p$, then $\widehat{c}$ lies on the boundary of just one hyperbolic component H, which has period p and lies inside the $\mathcal{P}$-wake. The boundary of H near $\widehat{c}$ is a smooth curve, except for a cusp at the point $\widehat{c}$ itself.*

Proof. — As in the proof of 4.2, we set $c = \widehat{c} + u^2$ and find a period p point $z(u)$ with multiplier $\lambda(u)$ which depends smoothly on u, with $d\lambda/du \neq 0$. Hence the locus $|\lambda(u)| = 1$ is a smooth curve in the u-plane, while its image in the c-plane has a cusp at $c = \widehat{c}$. The rest of the argument is completely analogous to the proof of 6.1. □

Lemma 6.3 **(The Root Point of a Hyperbolic Component).** — *Every parabolic point of ray period $n = rp$ is on the boundary of one and only one hyperbolic component of period n. Conversely, every hyperbolic component of period n has one and only one parabolic point of ray period n on its boundary. In this way, we obtain a canonical one-to-one correspondence between parabolic points and hyperbolic components in parameter space.*

Proof. — The first statement follows immediately from 6.1 and 6.2. Conversely, if H is a hyperbolic component of period n, then we can map H holomorphically into the open unit disk $\mathbf{D}$ by sending each $c \in H$ to the multiplier of the unique attracting orbit for f_c. In order to extend to the closure $\overline{H}$, it is convenient to lift to the curve $\mathrm{Per}_n/\mathbb{Z}_n$, using the proper holomorphic map $(c, \mathcal{O}) \mapsto c$ of §5. Evidently H lifts biholomorphically to an open set $H^\natural \subset \mathrm{Per}_n/\mathbb{Z}_n$, which then maps holomorphically to the λ_n-plane under the projection $(c, \mathcal{O}) \mapsto \lambda_n(c, \mathcal{O})$. (Here $H^\natural$ is a connected component of the set of $(c, \mathcal{O})$ such that $\mathcal{O}$ is an attracting period n orbit for f_c.) Since the projection to the λ_n-plane is open and proper, it follows easily that the closure $\overline{H}^\natural$ maps *onto* the closed disk $\overline{\mathbf{D}}$. In particular, there exists a point $(\widehat{c}, \widehat{\mathcal{O}})$ of $\overline{H}^\natural$ with $\lambda_n(\widehat{c}, \widehat{\mathcal{O}}) = +1$. Evidently this $\widehat{c}$ is a parabolic boundary point of H with ray period dividing n, and it follows from 6.1 and 6.2 that it must have ray period precisely n.

According to 4.7, for each $c \in H$ there is a unique repelling orbit of lowest period on the boundary of the immediate basin for the attracting orbit of f_c. Furthermore, according to 4.1, the portrait $\mathcal{P} = \mathcal{P}_H$ for this orbit is the same as the portrait for the parabolic orbit of $f_{\widehat{c}}$. Since there is only one parabolic point with specified portrait by Theorem 1.2, this proves that there can only one such point $\widehat{c} \in \partial H$. □

Definition. — This distinguished parabolic point on the boundary ∂H of a hyperbolic component is called the *root point* of the hyperbolic component H. We know from 1.2 and 1.4 that the parabolic points of ray period n can be indexed by the non-trivial orbit portraits of ray period n. *Hence the hyperbolic components of period n can also be indexed by non-trivial portraits of ray period n.* We will write $H = H_{\mathcal{P}}$ (or $\mathcal{P} = \mathcal{P}_H$) if H is the hyperbolic component with root point $r_{\mathcal{P}}$. We will say that H is a *primitive component* or a *satellite component* according as the associated portrait is primitive or satellite.

Remark 6.4. — Of course there are many other parabolic points in ∂H. For each root of unity $\mu = e^{2\pi i q/s} \neq 1$ a similar argument shows that there is at least one point $(\widehat{c}_\mu, \mathcal{O}_\mu) \in \partial H^\natural$ with $\lambda_n(\widehat{c}_\mu, \mathcal{O}_\mu) = \mu$. In fact the following theorem implies that $\widehat{c}_\mu$ is unique. This $\widehat{c}_\mu$ is the root point for a hyperbolic component H' of period $sn > n$, with associated orbit portrait $\mathcal{P}'$ of period n and rotation number q/s. By definition, $\mathcal{P}'$ is the (q/s)-*satellite* of $\mathcal{P}$, and H' is the (q/s)-*satellite* of H.

We next prove the following basic result of Douady and Hubbard. Again let H be a hyperbolic component of period n and let $H^\natural \subset \mathrm{Per}_n/\mathbb{Z}_n$ be the set of pairs $(c, \mathcal{O})$ with $c \in H$, where $\mathcal{O}$ is the attracting orbit for f_c.

***Theorem 6.5* (Uniformization of Hyperbolic Components).** — *The closure $\overline{H}$ is homeomorphic to the closed unit disk $\overline{\mathbf{D}}$. In fact there is a canonical homeomorphism*

$$\overline{\mathbb{D}} \cong \overline{H}^\natural \to \overline{H}$$

which carries each point $\lambda \neq 1$ *in* $\overline{\mathbb{D}}$ *to the unique point* $c \in \overline{H}$ *such that* f_c *has a period* n *orbit of multiplier* λ. *This homeomorphism extends holomorphically over a neighborhood of* $\overline{\mathbb{D}}$, *with just one critical point* $1 \in \overline{\mathbb{D}}$ *mapping to the root point* $\widehat{c} \in H$ *in the primitive case, and with no critical points in the satellite case. The closures of the various hyperbolic components are pairwise disjoint, except for the tangential contact between a component and its satellite as described in* 6.1.

Proof. — Recall that $\lambda_n : \mathrm{Per}_n / \mathbb{Z}_n \to \mathbb{C}$ is a proper holomorphic map of degree $\nu_2(n)/2$. We will first show that there are no critical values of λ_n within the closed unit disk $\overline{\mathbb{D}}$. This will imply that the inverse image $\lambda_n^{-1}(\overline{\mathbb{D}})$ is the disjoint union of $\nu_2(n)/2$ disjoint sets $\overline{H}^\natural$, each of which maps diffeomorphically onto $\overline{\mathbb{D}}$. First note that there are no critical values of λ_n on the boundary circle $\partial\,\mathbb{D}$. In the case of a root of unity $\mu \in \partial\,\mathbb{D}$, every $(c, \mathcal{O})$ with $\lambda_n(c, \mathcal{O}) = \mu$ must be parabolic, and it follows from 6.1 and 6.2 that the derivative of λ_n at $(c, \mathcal{O})$ is non-zero. Consider then a point $(\widehat{c}, \mathcal{O}) \in \partial H^\natural$ such that $\lambda_n(\widehat{c}, \mathcal{O})$ is not a root of unity. According to 5.2, we can use c as local uniformizing parameter throughout a neighborhood of $(\widehat{c}, \mathcal{O})$. If this were a critical point of λ_n, then it would follow that we could find two different line segments emerging from $\widehat{c}$ which map into $\mathbb{D}$, separated by two line segments which map outside of $\mathbb{D}$. In other words, one of the following two possibilities would have to occur.

Case 1. — There are two different hyperbolic components with $\widehat{c}$ as non-root boundary point. Each of these components must have a root point, and be contained in its associated wake. But these two components cannot be separated by any rational parameter ray, hence each one must be contained in the wake of the other, which is impossible.

Case 2. — The single hyperbolic component H must approach $\widehat{c}$ from two different directions, separated by two directions which lie outside of H. In other words. There must be a simple closed loop $L \subset \overline{H}$ which encloses points lying outside of $\overline{H}$. Now the collection of iterates $f_c^{\circ k}(0)$ must be uniformly bounded for $c \in L$, and hence also for all c in the region bounded by L. Thus this entire region must lie within the interior of the Mandelbrot set, which is impossible since this region contains parabolic points.

Thus both cases are impossible, and λ_n must be locally injective near the boundary of $H^\natural$. It follows easily that $H^\natural$ maps onto $\mathbb{D}$ by a proper map of some degree $d \geq 1$, and similarly that the boundary $\partial H^\natural$ wraps around the boundary circle $\partial\,\mathbb{D}$ exactly d times. Now a counting argument shows that this degree is $+1$. In fact the number of H or $H^\natural$ of period n is equal to $\nu_2(n)/2$ by 6.3 and 5.3. Since the degree of the map λ_n on $\mathrm{Per}_n / \mathbb{Z}_n$ is also $\nu_2(n)/2$ by 5.2, it follows that each $H^\natural$ must map with degree $d = 1$. Therefore λ_n maps each $\overline{H}^\natural$ biholomorphically onto $\overline{\mathbb{D}}$.

Next consider the projection $(c, \mathcal{O}) \mapsto c$ from the compact set $\overline{H}^{\natural}$ onto $\overline{H}$. This is one-to-one, and hence a homeomorphism, by a theorem of Douady and Hubbard which asserts that a polynomial of degree d can have at most $d-1$ non-repelling cycles. (Compare **[Sh1]**. Alternatively, it follows from the classical Fatou-Julia theory that a polynomial with one critical point can have at most one attracting cycle. If two distinct points of $\partial H^{\natural}$ mapped to a single point of ∂H, then, as in Case 2 above, a path between these points in $H^{\natural}$ would map to a loop in $\overline{H}$ which could enclose no boundary points of H, leading to a contradiction.)

According to 5.2, the parameter c can be used as local uniformizing parameter for $\mathrm{Per}_n / \mathbb{Z}_n$ unless $\lambda_n = 1$. Hence the only possible critical value for the projection $\overline{H}^{\natural} \to \overline{H}$ is the root point. In fact, by 6.1 and 6.2, the root point is actually a critical value if and only if H is a primitive component.

Finally suppose that two different hyperbolic components have a common boundary point. If this boundary point is parabolic, then one of these components must be a satellite of the other by 6.1 and 6.2. If the point were non-parabolic, then the argument of Case 1 above would yield a contradiction. This completes the proof of 6.5. □

7. Orbit Forcing

Recall that an orbit portrait is *non-trivial* if either it has valence $v \geq 2$, or it is the zero portrait $\{\{0\}\}$. The following statement follows easily from 1.3. However, it seems of interest to give a direct and more constructive proof; and the methods used will be useful in the next section.

***Lemma 7.1* (Orbit Forcing).** — *Let $\mathcal{P}$ and $\mathcal{Q}$ be distinct non-trivial orbit portraits. If their characteristic arcs satisfy $I(\mathcal{P}) \subset I(\mathcal{Q})$, then every f_c with a (repelling or parabolic) orbit of portrait $\mathcal{P}$ must also have a repelling orbit of portrait $\mathcal{Q}$.*

Compare Figure 5, and see 1.3 and for further discussion. The proof of 7.1 begins as follows.

Puzzle Pieces. — Recall from 2.10 that the pv rays landing on a periodic orbit for $f = f_c$ separate the dynamic plane into $pv - p + 1$ connected components, the closures of which are called the (unbounded) *preliminary puzzle pieces* associated with the given orbit portrait. (As in **[K]**, we work with puzzle pieces which are closed but not compact. The associated *bounded* pieces can be obtained by intersecting each unbounded puzzle piece with the compact region enclosed by some fixed equipotential curve.)

Most of these preliminary puzzle pieces Π have the *Markov property* that f maps Π homeomorphically onto some union of preliminary puzzle pieces. However, the puzzle piece containing the critical point is exceptional: Its image under f covers the critical value puzzle piece twice, and also covers some further puzzle pieces once. To obtain a modified puzzle with more convenient properties, we will subdivide this exceptional piece into two connected sub-pieces.

Let Π_1 be the preliminary puzzle piece containing the critical value. Then $\partial\Pi_1$ consists of the two rays whose angles bound the characteristic arc for $\mathcal{P}$, together with their common landing point, say z_1. The pre-image $\Pi_0 = f^{-1}(\Pi_1)$ is bounded by two rays landing at the point $z_0 = f^{-1}(z_1) \cap \mathcal{O}$, together with two rays landing at the symmetric point $-z_0$. Note that Π_0 is a connected set containing the critical point, and that the map f from Π_0 onto Π_1 is exactly two-to-one, except at the critical point 0, which maps to c.

The pv rays landing on $\mathcal{O}$, together with these two additional rays landing on $-z_0$, cut the complex plane up into $pv - p + 2$ closed subsets which we will call the pieces of the corrected *puzzle* associated with $\mathcal{P}$. These will be numbered as $\Pi_0, \Pi_1, \ldots, \Pi_{pv-p+1}$, with Π_0 and Π_1 as above. The central piece Π_0 will be called the *critical puzzle piece*, and Π_1 will be called the *critical value puzzle piece*. This corrected puzzle satisfies the following.

Modified Markov Property. — *The puzzle piece Π_0 maps onto Π_1 by a 2-fold branched covering, while every other puzzle piece maps homeomorphically onto a finite union of puzzle pieces.*

We can represent the allowed transitions by a *Markov matrix* M_{ij}, where

$$M_{ij} = \begin{cases} 1 & \text{if } \Pi_i \text{ maps homeomorphically, with } f(\Pi_i) \supset \Pi_j \\ 0 & \text{if } f(\Pi_i) \text{ and } \Pi_j \text{ have no interior points in common,} \end{cases}$$

and where $M_{01} = 2$ since Π_0 double covers Π_1. Since f is quadratic, note that the sum of entries in any column is equal to 2. Equivalently, this same data can be represented by a *Markov graph*, with one vertex for each puzzle piece, and with M_{ij} arrows from the i-th vertex to the j-th.

As an example, for the puzzle shown in Figure 13, we obtain the Markov graph of Figure 14, or the following Markov matrix

$$[M_{ij}] = \begin{bmatrix} 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \tag{8}$$

To illustrate the idea of the proof of 7.1, let us show that any f having an orbit with this portrait $\mathcal{P}$ must also have a repelling orbit with portrait $\mathcal{Q} = \{\{1/7, 2/7, 4/7\}\}$. (Compare the top implication in Figure 5.) Inspecting the next to last row of the matrix (8), we see that $f(\Pi_4) = \Pi_0 \cup \Pi_4 \cup \Pi_5$. Therefore, there is a branch g of f^{-1} which maps the interior of Π_4 holomorphically onto some proper subset of itself.

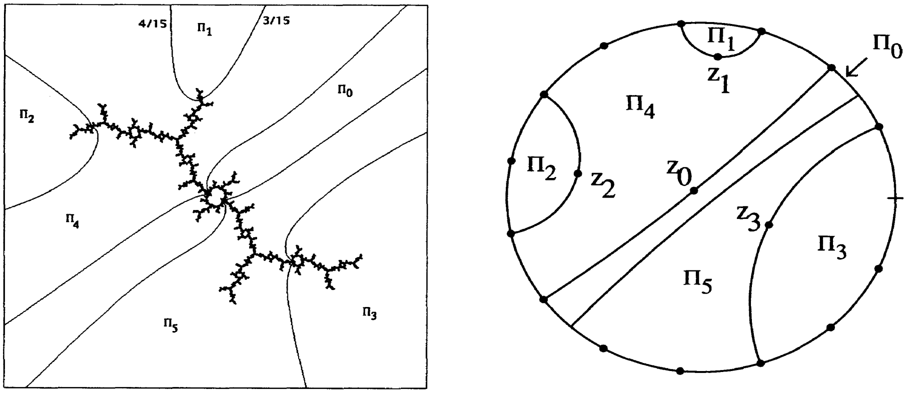


FIGURE 13. Julia set with a parabolic orbit of period four with characteristic arc $I(\mathcal{P}) = (3/15, 4/15)$, showing the six corrected puzzle pieces; and a corresponding schematic diagram. (For the corresponding preliminary puzzle, see the top of Figure 5.)

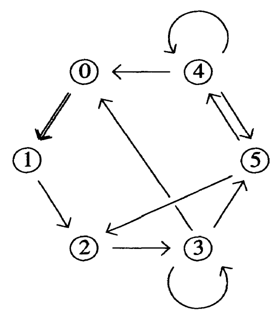


FIGURE 14. Markov graph associated with the matrix (8), with one vertex for each puzzle piece. Since f is quadratic, there are two arrows pointing to each vertex.

This mapping g must strictly decrease the Poincaré metric for the interior of Π_4. On the other hand, it is easy to check that the $1/7$, $2/7$ and $4/7$ rays are all contained in the interior of Π_4. Hence their landing points, call them w_1, w_2 and w_3, are also contained in Π_4, necessarily in the interior, since the points of $K \cap \partial\Pi_4$ have period four. Now

$$g : w_1 \mapsto w_3 \mapsto w_2 \mapsto w_1,$$

and all positive distances are strictly decreased. Thus if the distance from w_i to w_j were greater than zero, then applying g three times we would obtain a contradiction. This proves that $w_1 = w_2 = w_3$, as required. This fixed point must be repelling, since g clearly cannot be an isometry.

A similar argument proves the following statement. Suppose that $f = f_c$ has an orbit $\mathcal{O}$ with some given portrait $\mathcal{P}$. By a *Markov cycle* for $\mathcal{P}$ we will mean an infinite sequence of non-critical puzzle pieces $\Pi_{i_1}, \Pi_{i_2} \ldots$ which is periodic, $i_j = i_{j+m}$ with period $m \geq 1$, and which satisfies $f(\Pi_{i_\alpha}) \supset \Pi_{i_{\alpha+1}}$, so that $M_{i_\alpha\, i_{\alpha+1}} = 1$, for every α modulo m.

***Lemma 7.2* (Realizing Markov Cycles).** — *Given such a Markov cycle, there is one and only one periodic orbit $z_1 \mapsto \cdots \mapsto z_m$ for f_c with period dividing m so that each z_α belongs to Π_{i_α}, and this orbit is necessarily repelling unless it coincides with the given orbit $\mathcal{O}$ (which may be parabolic). In particular, for any angle t which is periodic under doubling, if the dynamic ray with angle $2^\alpha t$ lies in Π_{i_α} for all integers α, then this ray must land at the point z_α.*

(Note that the period of t may well be some multiple of m, as in the example just discussed.)

Proof. — There is a unique branch of f_c^{-1} which carries the interior of $\Pi_{i_{\alpha+1}}$ holomorphically onto a subset of Π_{i_α}. Let g_{i_α} be the composition of these m maps, in the appropriate reversed order so as to carry the interior of Π_{i_α} into itself.

A similar construction applies to the associated external angles. Let $J_i \subset \mathbb{R}/\mathbb{Z}$ be the set of all angles of dynamic rays which are contained in Π_i. Thus each J_i is a finite union of closed arcs, and together the J_i cover $\mathbb{R}/\mathbb{Z}$ without overlap. Now there is a unique branch of the 2-valued map $t \mapsto t/2$ which carries $J_{i_{\alpha+1}}$ into J_{i_α} with derivative $1/2$ everywhere. Taking an m-fold composition, we map each J_{i_α} into itself with derivative $1/2^m$. This composition may well permute the various connected components of J_{i_α}. However, some iterate must carry some component of J_{i_α} into itself, and hence have a unique fixed point t in that component. The landing point of the corresponding dynamic ray will be a periodic point $z_\alpha \in \Pi_{i_\alpha}$.

Case 1. — If this landing point belongs to the interior of Π_{i_α}, then it is fixed by some iterate of our map g_{i_α}. This map g_{i_α} cannot be an isometry, hence it must contract the Poincaré metric. Therefore every orbit under g_{i_α} must converge towards z_α. Thus z_α is an attracting fixed point for g_{i_α}, and hence is a repelling periodic point for f.

Case 2. — If the landing point belongs to the boundary of Π_{i_α} then it must belong to $\mathcal{O} \cup \{-z_0\}$, and hence to the original orbit $\mathcal{O}$ since $-z_0$ is not periodic. Evidently this case will occur only when the angle t belongs to the union $A_1 \cup \cdots \cup A_p$ of angles in the given portrait $\mathcal{P}$. □

Note. — It is essential for this argument that our given Markov cycle $\{\Pi_{i_\alpha}\}$ does not involve the critical puzzle piece Π_0. In fact, as an immediate corollary we get the following statement:

***Corollary 7.3* (Non-Repelling Cycles).** — *Any non-repelling periodic orbit for f must intersect the critical puzzle piece Π_0 as well as the critical value puzzle piece Π_1.*

Proof of 7.1. — If $I(\mathcal{P}) \subset I(\mathcal{Q})$, then it follows from Lemma 2.9 that there exists a map $\widehat{f}$ having both an orbit with portrait $\mathcal{P}$ and an orbit with portrait $\mathcal{Q}$. The latter orbit determines a Markov cycle in the puzzle associated with $\mathcal{P}$. (The condition $I(\mathcal{P}) \subset I(\mathcal{Q})$ guarantees that this cycle avoids the critical puzzle piece.) Now for any map f with an orbit of portrait $\mathcal{P}$, we can use this Markov cycle, together with 7.2, to construct the required periodic orbit and to guarantee that the rays associated with the portrait $\mathcal{Q}$ land on it, as required. □

In fact an argument similar to the proof of 7.2 proves a much sharper statement. Let $\mathcal{O}$ be a repelling periodic orbit with non-trivial portrait $\mathcal{P}$.

***Lemma 7.4* (Orbits Bounded Away From Zero).** — *Given an infinite sequence of non-critical puzzle pieces $\{\Pi_{i_k}\}$ for $k \geq 0$ with $f(\Pi_{i_k}) \supset \Pi_{i_{k+1}}$, there is one and only one point $w_0 \in K(f)$ so that the orbit $w_0 \mapsto w_1 \mapsto \cdots$ satisfies $w_k \in \Pi_{i_k}$ for every $k \geq 0$. It follows that the action of f on the compact set $K_{\mathcal{P}}$ consisting of all $w_0 \in K(f)$ such that the forward orbit $\{w_k\}$ never hits the interior of Π_0 is topologically conjugate to the one-sided subshift of finite type, associated to the matrix $[M_{ij}]$ with 0-th row and column deleted. In particular, the topology of $K_{\mathcal{P}}$ depends only on $\mathcal{P}$, and not on the particular choice of f within the $\mathcal{P}$-wake.*

Proof Outline. — First replace each puzzle piece Π_i by a slightly thickened puzzle piece, as described in **[M3]**. (Compare §8, Figure 18.) The interior of this thickened piece is an open neighborhood $N_i \supset \Pi_i$, with the property that $f(N_i) \supset \overline{N}_j$ whenever $f(\Pi_i) \supset \Pi_j$. It then follows that there is a branch of f^{-1} which maps N_j into N_i, carrying $K \cap \Pi_j$ into $K \cap \Pi_i$, and reducing distances by at least some fixed ratio $r < 1$ throughout the compact set $K \cap \Pi_j$. Further details are straightforward. □

Presumably this statement remains true for a parabolic orbit, although the present proof does not work in the parabolic case. (Compare **[Ha]**.)

8. Renormalization

One remarkable property of the Mandelbrot boundary is that it is densely filled with small copies of itself. (See Figures 11, 14 for a magnified picture of one such small copy.) This section will provide a rough outline, without proofs, of the Douady-Hubbard theory of renormalization, or the inverse operation of tuning, which provides a dynamical explanation for these small copies. It is based on **[D4]** as well as **[DH3]**, **[D3]**. (Compare **[D1]**, **[M1]**. For the Yoccoz interpretation of this construction, see **[Hu]**, **[M3]**, **[Mc]**, **[Ly]**. For a more general form of renormalization, see **[Mc]**, **[RS]**.)

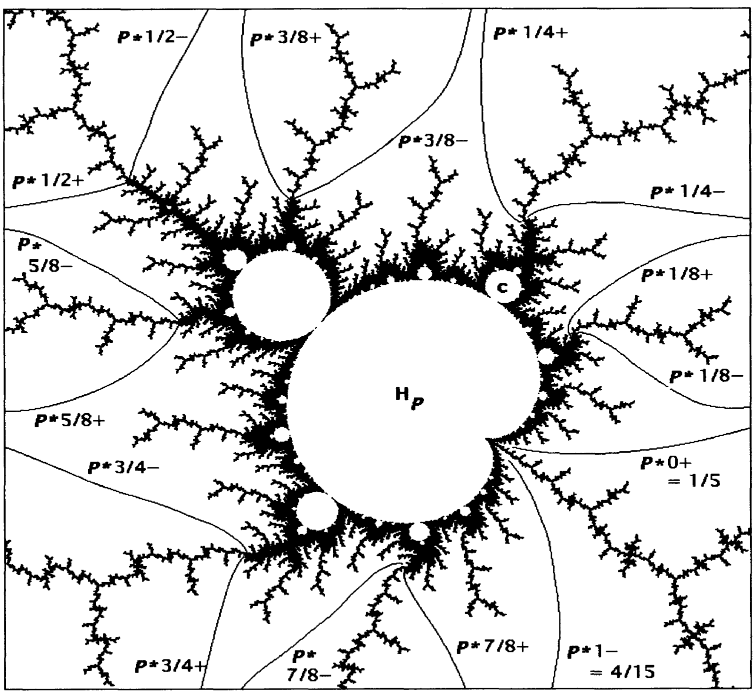


FIGURE 15. Detail near the period 4 hyperbolic component $H_{\mathcal{P}}$ of Figure 12, where $\mathcal{P} = \mathcal{P}(1/5, 4/15)$, showing the first eight of the parameter sectors which must be pruned away from M to leave the small Mandelbrot set consisting of $\mathcal{P}$-renormalizable parameter values.

To begin the construction, consider any orbit portrait $\mathcal{P}$ of ray period $n \geq 2$ and valence $v \geq 2$. Let c be a parameter value in $W_{\mathcal{P}} \cup \{r_{\mathcal{P}}\}$, so that $f = f_c$ has a periodic orbit $\mathcal{O}$ with portrait $\mathcal{P}$, and let $S = S(f)$ be the critical value sector for this orbit (so that $\overline{S}$ is the critical value puzzle piece). To a first approximation, we could try to say that f is "$\mathcal{P}$-renormalizable" if the orbit of c under $f^{\circ n}$ is completely contained in S. In fact this is a necessary and sufficient condition whenever the map $f^{\circ n-1}|_S$ is univalent. However, in examples such as that of Figures 1, 2 one needs a slightly sharper condition.

Let $\mathcal{I}_{\mathcal{P}} = (t_-, t_+)$ be the characteristic arc for this portrait, so that ∂S consists of the dynamic rays of angle t_- and t_+ together with their common landing point z_1, and let $\ell = t_+ - t_-$ be the length of this arc.

***Lemma 8.1* (A (Nearly) Quadratic-Like Map).** — *The dynamic rays of angle $t'_1 = t_- + \ell/2^n$ and $t'_2 = t_+ - \ell/2^n$ land at a common point $z' \neq z_1$ in $S \cap f^{-n}(z_1)$. Let $S' \subset S$ be the region bounded by ∂S together with these two rays and their common landing*

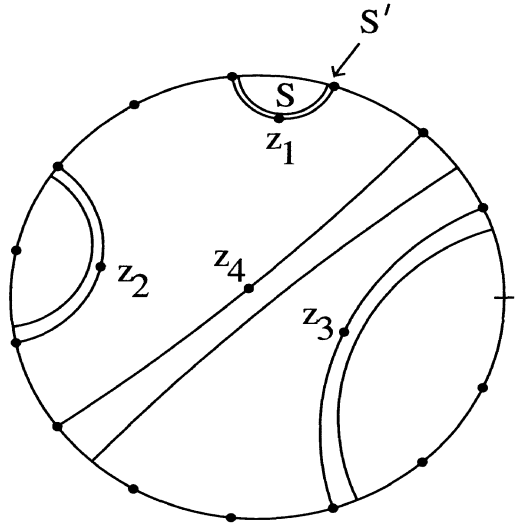


FIGURE 16. The n-fold pull-back of the critical sector S along the orbit $\mathcal{O}$, illustrated schematically for the orbit diagram of period $n = 4$ which has characteristic arc $(1/5, 4/15)$. Compare Figures 5 (top), 12, 16.

point. Then the map $f^{\circ n}$ carries S' onto S by a proper map of degree two, with critical value equal to the critical value $f(0) = c$.

This region S' can be described as the n-fold "pull-back" of S along the orbit $\mathcal{O}$. (Compare Figure 16, which also shows the first three forward images of S'.)

Proof of 8.1. — First suppose that $c \in W_{\mathcal{P}}$ is outside the Mandelbrot set. Then, following Appendix A, we can bisect the complex plane by the two rays leading from infinity to the critical point. (Compare the proof of 2.9.) In order to check that the two rays of angle t'_1 and t'_2 have a common landing point, we need only show that they have the same symbol sequence with respect to the resulting partition. In other words, we must show, for every $k \geq 0$, that the $2^k t'_1$ and $2^k t'_2$ rays lie on the same side of the bisecting critical ray pair. For $k \geq n$ this is clear since $2^n t'_1 \equiv t_+$ and $2^n t'_2 \equiv t_-$ modulo $\mathbb{Z}$.

Now consider the critical puzzle piece Π_0 of §7. Evidently Π_0 is a neighborhood, of angular radius $\ell/4$, of the bisecting critical ray pair. For $k < n-1$ the dynamic rays with angle $2^k t_-$ and $2^k t_+$ both lie in the same component of $\mathbb{C} \smallsetminus \Pi_0$. Since $2^k t'_j$ differs from $2^k t_j$ by at most $\ell/4$, it follows that the $2^k t'_1$ and $2^k t'_2$ rays have the same symbol. Finally, for $k = n-1$, it is not difficult to check that the $2^k t'_1$ and $2^k t'_2$ rays both land at the same point $-z_0 \neq z_0$. This proves that the t'_1 and t'_2 rays land at the same point, different from z_1, when $c \notin M$. A straightforward continuity argument now proves the same statement for all $c \in W_{\mathcal{P}}$.

Thus we obtain the required region $S' \subset S$. As in §2, it will be convenient to complete the complex plane by adjoining a circle of points at infinity. Note that the boundary of S' within this circled plane © consists of two arcs of length $\ell/2^n$ at infinity, together with two ray pairs and their common landing points. As we traverse this boundary once in the positive direction, the image under $f^{\circ n}$ evidently traverses the boundary of S twice in the positive direction. Using the Argument Principle, it follows that the image of S' is contained in S, and covers every point of S twice, as required. Thus $f^{\circ n}|_{S'}$ must have exactly one critical point, which can only be c. □

Thus we have an object somewhat like a quadratic-like map, as studied in **[DH3]**. Note however that S' is not compactly contained in S.

Definition. — We will say that f is *$\mathcal{P}$-renormalizable* if $f(0) = c$ is contained in the closure $\overline{S}'$, and furthermore the entire forward orbit of c under the map $f^{\circ n}$ is contained in $\overline{S}'$. If this condition is satisfied, and the orbit of c is also bounded so that $c \in M$, then we will say that c belongs to the "*small copy*" $\mathcal{P} * M$ of the Mandelbrot set which is associated with $\mathcal{P}$. (This terminology will be justified in 8.2. If the orbit is unbounded, then we may say that c belongs to a *$\mathcal{P}$-renormalizable external ray.*)

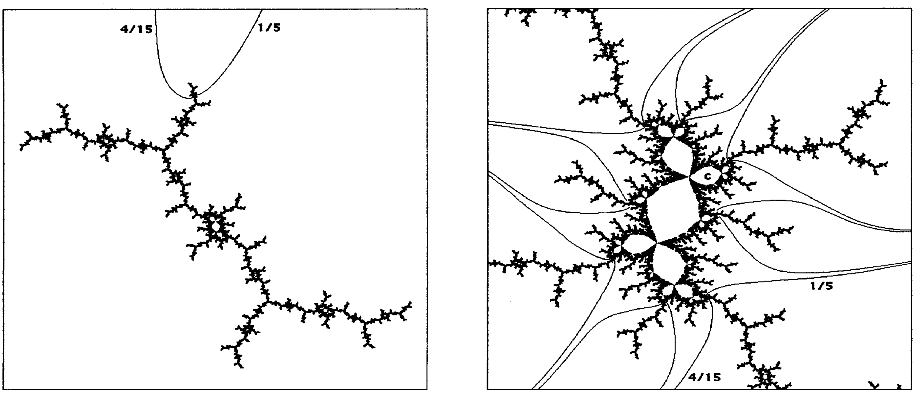


FIGURE 17. Julia set for the center point of the period 12 satellite component of Figure 12 (the point c of Figure 15), and a detail near the critical value c, showing the first eight of the sectors of the dynamic plane which must be pruned away to leave the small Julia set associated with $\mathcal{P}$-renormalization, with $\mathcal{P}$ as in Figures 12, 15. (Here the right hand figure has been magnified by a factor of 75.) This can be described as the Julia set of Figure 13 (left) tuned by a "Douady rabbit" Julia set.

Closely associated is the "*small filled Julia set*" $K' = K(f^{\circ n}|S')$ consisting of all $z \in \overline{S}'$ such that the entire forward orbit of z under $f^{\circ n}$ is bounded and contained in $\overline{S}'$. (Compare Figure 17.) Thus the critical value $f(0) = c$ belongs to K' if and only if f is $\mathcal{P}$-renormalizable, with $c \in M$. As in the classical Fatou-Julia theory, c belongs to K' if and only if K' is connected.

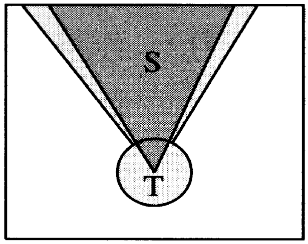


FIGURE 18. The sector S and the thickened sector T.

In order to tie this construction up with Douady and Hubbard's theory of polynomial-like mappings, we need to thicken the sector S, and then cut it down to a bounded set. (Compare **[M3]**.) We exclude the exceptional special case where c is the root point $r_{\mathcal{P}}$. Thus we will suppose that the periodic point $z_1 \in \partial S$ is repelling. Choose a small disk D_ε about z_1 which is mapped univalently by $f_c^{\circ n}$ and is compactly contained in $f_c^{\circ n}(D_\varepsilon)$. Choose also a very small $\eta > 0$, and consider the dynamic rays with angle $t_- - \eta$ and $t_+ + \eta$. Following these rays until they first meet D_ε, they delineate an open region $T \supset S \cup D_\varepsilon$ in $\mathbb{C}$. (Compare Figure 18.) Now let T' be the connected component of $f_c^{-n}(T)$ which contains S'. It is not difficult to check that $\overline{T}' \subset T$, and that $f_c^{\circ n}$ carries T' onto T by a proper map of degree two.

To obtain a bounded region, we let U be the intersection of T with the set $\{z \in \mathbb{C}\ ;\ G^K(z) < 1\}$, where G^K is the Green's function for $K = K(f_c)$. Similarly, let U' be the intersection T' with $\{z\ ;\ G^K(z) < 1/2^n\}$. Then U' is compactly contained in U, and $f_c^{\circ n}$ carries U' onto U by a proper map of degree two. *In other words,* $f_c^{\circ n}|U'$ *is a quadratic-like map.*

Evidently the forward orbit of a point $z \in U'$ under $f_c^{\circ n}$ is contained in U' if and only if z belongs to the small filled Julia set K'. In particular, for $c \in M$, the map f_c is $\mathcal{P}$-renormalizable if and only if $c \in K'$, or if and only if K' is connected.

If these conditions are satisfied, then according to **[DH3]** the map $f_c^{\circ n}$ restricted to a neighborhood of K' is "hybrid equivalent" to some uniquely defined quadratic map $f_{c'}$, with $c' \in M$. Briefly, we will write $c = \mathcal{P} * c'$, or say that c equals $\mathcal{P}$ *tuned by* c'. Douady and Hubbard show also that this correspondence

$$c' \mapsto \mathcal{P} * c'$$

is a well defined continuous embedding of $M \setminus \{1/4\}$ onto a proper subset of itself. As an example, as c' varies over the hyperbolic component $H_{\{\{0\}\}}$ which is bounded by the cardioid, they show that $\mathcal{P} * c'$ varies over the hyperbolic component $H_{\mathcal{P}}$.

It is convenient to supplement this construction, by defining the operation $\mathcal{P},\ c' \mapsto \mathcal{P} * c'$ in two further special cases. If c' is the root point $1/4 = r_{\{\{0\}\}}$ of M, then we define

$$\mathcal{P} * (1/4) = r_{\mathcal{P}}$$

to be the root point of the $\mathcal{P}$-wake. Furthermore, if $\mathcal{P} = \{\{0\}\}$ is the zero orbit portrait, then we define $\{\{0\}\}*$ to be the identity map,

$$\{\{0\}\} * c' = c'$$

for all $c' \in M$. With these definitions, we have the following basic result of Douady and Hubbard.

***Theorem 8.2* (Tuning).** — *For each non-trivial orbit portrait $\mathcal{P}$, the correspondence $c \mapsto \mathcal{P} * c$ defines a continuous embedding of the Mandelbrot set M into itself. The image of this embedding is just the "small Mandelbrot set" $\mathcal{P} * M \subset M$ described earlier. Furthermore, there is a unique composition operation $\mathcal{P}$, $\mathcal{Q} \mapsto \mathcal{P} * \mathcal{Q}$ between non-trivial orbit portraits so that the associative law is valid,*

$$(\mathcal{P} * \mathcal{Q}) * c = \mathcal{P} * (\mathcal{Q} * c)$$

for all $\mathcal{P}$, $\mathcal{Q}$ and c'. Under this $*$ composition operation, the collection of all non-trivial orbit portraits forms a free (associative but noncommutative) monoid, with the zero orbit portrait as identity element.

The proof is beyond the scope of this note.

We can better understand this construction by introducing a nested sequence of open sets

$$S = S^{(0)} \supset S' = S^{(1)} \supset S^{(2)} \supset \cdots$$

in the dynamic plane for f, where $S^{(k+1)}$ is defined inductively as $S^{(k)} \cap f^{-n}(S^{(k)})$ for $k \geq 1$. Thus $S = S^{(0)}$ is bounded by the dynamic rays of angle t_- and t_+, together with their common landing point z_1. Similarly, $S^{(1)}$ is bounded by $\partial S^{(0)}$ together with the rays of angle $t_- + \ell/2^n$ and $t_+ - \ell/2^n$, together with their common landing point, which is an n-fold pre-image of z_1. If $c \in S^{(1)}$, so that $S^{(2)}$ is a 2-fold branched covering of $S^{(1)}$, then $S^{(2)}$ has two further boundary components, namely the rays of angle $t_- + \ell/2^{2n}$ and $t_- + \ell/2^n - \ell/2^{2n}$ and their common landing point, together with the rays of angle $t_+ - \ell/2^n + \ell/2^{2n}$ and $t_+ - \ell/2^{2n}$ and their common landing point, for a total of 4 boundary components. Similarly, if $c \in S^{(2)}$, then $S^{(3)}$ has 8 boundary components, as illustrated in Figure 17.

The angles which are left, after we have cut away the angles in all of these (open) sectors, form a standard middle fraction Cantor set $\mathcal{K}$, which can be described as follows. Let ϕ be the fraction $1 - 2/2^n$. Start with the closure $[t_-, t_+]$ of the characteristic arc for $\mathcal{P}$, with length ℓ. First remove the open middle segment of length $\phi\,\ell$, leaving two arcs of length $\ell/2^n$. Then, from each of these two remaining closed arcs, remove the middle segment of length $\phi\,\ell/2^n$, leaving four segments of length $\ell/2^{2n}$, and continue inductively. The intersection of all of the sets obtained in this way is the required Cantor set $\mathcal{K} \subset [t_-, t_+]$ of angles. These are precisely the angles of the dynamic rays which land on the small Julia set $\partial K'$ (at least if we assume that these Julia sets are locally connected).

There is a completely analogous construction in parameter space, as illustrated in Figure 15. As noted earlier, parameter rays of angle t_- and t_+ land on a common point $r_{\mathcal{P}}$, and together form the boundary of the $\mathcal{P}$-wake. Similarly, the parameter rays of angle $t_- + \ell/2^n$ and $t_+ - \ell/2^n$ must land at a common point. These rays, together with their landing point, cut $W_{\mathcal{P}}$ into two halves. For c in the inner half, with boundary point $r_{\mathcal{P}}$, the critical value of f_c lies in $S' = S^{(1)}$, while for c in the outer half, this is not true. Similarly, for each pair of dynamic rays with a common landing point in ∂K, forming part of the boundary of $S^{(k)}$, there is a pair of parameter rays with the same angles which have a common landing point in ∂M and form part of the boundary of a corresponding region $W_{\mathcal{P}}^{(k)}$ in parameter space. The basic property is that $c \in W_{\mathcal{P}}^{(k)}$ if and only if c belongs to the corresponding region $S^{(k)}$ in the dynamic plane for f_c.

Dynamically, the Cantor set $\mathcal{K} \subset \mathbb{R}/\mathbb{Z}$ can be described as the set of angles in

$$[t_-,\ t_- + \ell/2^n] \cup [t_+ - \ell/2^n,\ t_+]$$

such that the entire forward orbit under multiplication by 2^n is contained in this set. Evidently the resulting dynamical system is topologically isomorphic to the one-sided two-shift. Thus each element $t \in \mathcal{K}$ can be coded by an infinite sequence $(b_0, b_1, \ldots)$ of bits, where each b_k is zero or one according as $2^{nk}t$ belongs to the left or right subarc. We will write $t = \mathcal{P} * (b_0 b_1 b_2 \cdots)$. Intuitively, we can identify this sequence of bits b_i with the angle $.b_0 b_1 b_2 \cdots = \sum b_k/2^{k+1}$. However, some care is needed since the correspondence $.b_0 b_1 b_2 \cdots \mapsto \mathcal{P} * (b_0 b_1 \cdots)$has a jump discontinuity at every dyadic rational angle, i.e., at those angles corresponding to gaps in the Cantor set $\mathcal{K}$. Thus we must distinguish between the left hand limit $\mathcal{P} * \alpha-$ and the right hand limit $\mathcal{P} * \alpha+$ when α is a dyadic rational.

With this notation, the angles of the bounding rays for the various open sets $S^{(k)}$, or for the corresponding sets $W_{\mathcal{P}}^{(k)}$ in parameter space, are just these left and right hand limits $\mathcal{P} * \alpha\pm$, where α varies over the dyadic rationals; and the composition operation between non-trivial orbit portraits can be described as follows: *If $\mathcal{Q}$ has characteristic arc (t_-, t_+), then $\mathcal{P} * \mathcal{Q}$ has characteristic arc $(\mathcal{P} * t_-, \mathcal{P} * t_+)$.* For further details, see **[D3]**.

9. Limbs and the Satellite Orbit

Let $\mathcal{P}$ be a non-trivial orbit portrait with period $p \geq 1$ and ray period $rp \geq p$. (Thus $\mathcal{P}$ may be either a primitive or a satellite portrait.) Recall that the *limb* $M_{\mathcal{P}}$ consists of all points which belong both to the Mandelbrot set M and to the closure $\overline{W}_{\mathcal{P}}$ of the $\mathcal{P}$-wake. By definition, a limb $M_{\mathcal{Q}}$ with $\mathcal{Q} \neq \mathcal{P}$ is a *satellite* of $M_{\mathcal{P}}$ if its root point $\mathbf{r}_\mathcal{Q}$ belongs to the boundary of the associated hyperbolic component $H_\mathcal{P}$. (See 6.4.) We will prove the following two statements. (Compare **[Hu]**, **[Sø]**, **[S3]**.)

***Theorem 9.1* (Limb Structure).** — *Every point in the limb $M_\mathcal{P}$ either belongs to the closure $\overline{H}_\mathcal{P}$ of the associated hyperbolic component, or else belongs to some satellite limb $M_\mathcal{Q}$.*

(For a typical example, see Figure 12.) For any parameter value c in the wake $W_\mathcal{P}$, let $\mathcal{O}(c) = \mathcal{O}_\mathcal{P}(c)$ be the repelling orbit for f_c which has period p and portrait $\mathcal{P}$. Clearly this orbit $\mathcal{O}(c)$ varies holomorphically with the parameter value c.

***Corollary 9.2* (The Satellite Orbit).** — *To any $c \in W_\mathcal{P}$ there is associated another orbit $\mathcal{O}^\star(c) = \mathcal{O}^\star_\mathcal{P}(c)$, distinct from $\mathcal{O}(c)$, which has period $n = rp$ and which also varies holomorphically with the parameter value c. As c tends to the root point $\mathbf{r}_\mathcal{P}$, the two orbits $\mathcal{O}(c)$ and $\mathcal{O}^\star(c)$ converge towards a common parabolic orbit of portrait $\mathcal{P}$. (Compare 4.1.) This associated orbit $\mathcal{O}^\star(c)$ is attracting if c belongs to the hyperbolic component $H_\mathcal{P} \subset W_\mathcal{P}$, indifferent for $c \in \partial H_\mathcal{P}$, and is repelling for $c \in W_\mathcal{P} \smallsetminus \overline{H}_\mathcal{P}$, with portrait equal to $\mathcal{Q}$ if c belongs to the satellite wake $W_\mathcal{Q}$.*

As an example, both statements apply to the zero portrait, with $M_{\{\{0\}\}}$ equal to the entire Mandelbrot set, with $W_{\{\{0\}\}} = \mathbb{C} \smallsetminus (1/4, +\infty)$, and with $H_{\{\{0\}\}}$ bounded by the cardioid. In this case, for any $c \in W_{\{\{0\}\}}$, the orbit $\mathcal{O}(c)$ consists of the *beta fixed point* $(1 + \sqrt{1-4c})/2$ while $\mathcal{O}^\star(c)$ consists of the *alpha fixed point* $(1 - \sqrt{1-4c})/2$, taking that branch of the square root function with $\sqrt{1} = 1$.

Proof of 9.1. — For each $c \in H_\mathcal{P}$ let $\mathcal{O}^\star(c)$ be the unique attracting periodic orbit. By the discussion in §6, this orbit extends analytically as we vary c over some neighborhood of the closure $\overline{H}_\mathcal{P}$, provided that we stay within the wake $W_\mathcal{P}$. Furthermore, this orbit becomes strictly repelling as we cross out of $\overline{H}_\mathcal{P}$. Therefore we can choose a neighborhood N of $\overline{H}_\mathcal{P}$ which is small enough so that this analytically continued orbit $\mathcal{O}^\star(c)$ will be strictly repelling for all $c \in N \cap W_\mathcal{P} \smallsetminus \overline{H}_\mathcal{P}$. If c also belongs to the Mandelbrot set, so that $c \in N \cap M_\mathcal{P} \smallsetminus \overline{H}_\mathcal{P}$, it follows that at least one rational dynamic ray lands on the orbit $\mathcal{O}^\star(c)$; hence there is an orbit portrait $\mathcal{Q} = \mathcal{Q}(c)$ of period n associated with $\mathcal{O}^\star(c)$. Choosing the neighborhood N even smaller if necessary, we will show that the rotation number of $\mathcal{Q}(c)$ is non-zero, and hence that this portrait $\mathcal{Q}(c)$ is non-trivial. In other words, we will prove that c belongs to a limb $M_\mathcal{Q}$ which is associated to the orbit $\mathcal{O}^\star(c)$.

First consider a point $\widehat{c}$ which belongs to the boundary $\partial H_\mathcal{P}$. Then $\mathcal{O}^\star(\widehat{c})$ is an indifferent periodic orbit, with multiplier on the unit circle. Consider some dynamic ray $\mathcal{R}_t^K$ which has period n, but does not participate in the portrait $\mathcal{P}$, and hence does not land on the original orbit $\mathcal{O}(\widehat{c})$. Such a ray certainly cannot land on $\mathcal{O}^\star(\widehat{c})$, for that would imply that $\mathcal{O}^\star(\widehat{c})$ was a repelling or parabolic orbit of rotation number zero. However, for $\widehat{c}$ in the boundary of $H_\mathcal{P}$ the orbit $\mathcal{O}^\star(\widehat{c})$ is never repelling, and is parabolic of rotation number zero only when $\widehat{c}$ is the root point of $H_{\mathcal{P}}$, so that $\mathcal{O}^\star(\widehat{c}) = \mathcal{O}(\widehat{c})$. Since we have assumed that the ray $\mathcal{R}_t^K$ does not land on $\mathcal{O}(\widehat{c})$, it must land on some repelling or parabolic periodic point which is disjoint from $\mathcal{O}^\star(\widehat{c})$. In fact it must land on a repelling orbit, since a quadratic map cannot have a parabolic orbit and also a disjoint indifferent orbit. (Compare §6.) Now as we perturb c throughout some neighborhood of $\widehat{c}$ it follows that the corresponding ray still lands on a repelling periodic point disjoint from $\mathcal{O}^\star(c)$. Since $\mathcal{O}^\star(c)$ has period n, but no ray of period n can land on it, this proves that the rotation number of the associated portrait $\mathcal{Q}(c)$ is non-zero, as asserted.

Let X be any connected component of $M_{\mathcal{P}} \smallsetminus \overline{H}_{\mathcal{P}}$. Since the Mandelbrot set is connected, X must have some limit point in $\partial H_{\mathcal{P}}$. Therefore, by the argument above, some point $c \in X$ must belong to a wake $W_{\mathcal{Q}}$ associated with the orbit $\mathcal{O}^\star(c)$. Since the portrait $\mathcal{Q}$ has period n, the root point $\mathbf{r}_{\mathcal{Q}}$ of its wake must lie on the boundary of some hyperbolic component H' which has period n and is contained in $W_{\mathcal{P}}$. In fact, for suitable choice of c, we claim that H' can only be $H_{\mathcal{P}}$ itself. There are finitely many other components of period n, but these others are all bounded away from $H_{\mathcal{P}}$, while the point $c \in X$ can be chosen arbitrarily close to $\overline{H}_{\mathcal{P}}$. Thus we may assume that $W_{\mathcal{Q}}$ is rooted at a point of $\partial H_{\mathcal{P}}$, and hence is a satellite wake. Since the connected set X cannot cross the boundary of $W_{\mathcal{Q}}$, it follows that X is completely contained within $W_{\mathcal{Q}}$, which completes the proof of 9.1. □

Proof of 9.2. — As in the argument above, the orbit $\mathcal{O}^\star(c)$ is well defined for c in some neighborhood of $W_{\mathcal{P}} \cap \overline{H}_{\mathcal{P}}$, and we can try to extend analytically throughout the simply connected region $W_{\mathcal{P}}$. There is a potential obstruction if we ever reach a point in $W_{\mathcal{P}}$ where the multiplier λ_n of this analytically extended orbit is equal to $+1$. However, this can never happen. In fact such a point would have to belong to the Mandelbrot set, and hence to some satellite limb $M_{\mathcal{Q}}$. But we can extend analytically throughout the associated wake $W_{\mathcal{Q}}$, taking $\mathcal{O}^\star(c)$ to be the repelling orbit $\mathcal{O}_{\mathcal{Q}}(c)$ for every $c \in W_{\mathcal{Q}}$. Thus there is no obstruction. It follows similarly that the analytically extended orbit must be repelling everywhere in $W_{\mathcal{P}} \smallsetminus \overline{H}_{\mathcal{P}}$. For if it became non-repelling at some point c, then again c would have to belong to some satellite limb $M_{\mathcal{Q}}$, but $\mathcal{O}^\star(c)$ is repelling throughout the wake $W_{\mathcal{Q}}$. □

***Corollary 9.3* (Limb Connectedness).** — *Each limb $M_{\mathcal{P}} = M \cap \overline{W}_{\mathcal{P}}$ is connected, even if we remove its root point $r_{\mathcal{P}}$.*

Proof. — The entire Mandelbrot set is connected by **[DH1]**. It follows that each $M_{\mathcal{P}}$ is connected. For if some limb $M_{\mathcal{P}}$ could be expressed as the union of two disjoint non-vacuous compact subsets, then only one of these two could contain the root point $r_{\mathcal{P}}$. The other would be a non-trivial open-and-closed subset of M, which is impossible.

Now consider the open subset $M_{\mathcal{P}} \smallsetminus \{r_{\mathcal{P}}\}$. This is a union of the connected set $\overline{H}_{\mathcal{P}} \smallsetminus \{r_{\mathcal{P}}\}$, together with the various satellite limbs $M_{\mathcal{Q}}$, where each $M_{\mathcal{Q}}$ has root point $r_{\mathcal{Q}}$ belonging to $\overline{H}_{\mathcal{P}} \smallsetminus \{r_{\mathcal{P}}\}$. Since each $M_{\mathcal{Q}}$ is connected, the conclusion follows. □

Remark 9.4.. — It follows easily that every satellite root point separates the Mandelbrot set into exactly two connected components, and hence that exactly two parameter rays land at every such point. For a proof of the corresponding statement for a primitive root (other than 1/4) see [**Ta**] or [**S3**].

Appendix A
Totally Disconnected Julia Sets and the Mandelbrot set

This appendix will be a brief review of well known material. For any parameter value c, let $K = K(f_c)$ be the filled Julia set for the map $f_c(z) = z^2 + c$, and let

$$G(z) = G^K(z) = \lim_{n\to\infty} \frac{1}{2^n} \log |f^{\circ n}(z)|$$

be the canonical potential function or *Green's function*, which vanishes only on K, and satisfies $G(f(z)) = 2G(z)$. The level sets $\{z;\ G(z) = G_0\}$ are called *equipotential curves* for K, and the orthogonal trajectories which extend to infinity are called the *dynamic rays* $\mathcal{R}_t^K$, where $t \in \mathbb{R}/\mathbb{Z}$ is the angle at infinity.

Now suppose that K is totally disconnected (and hence coincides with the Julia set $J = \partial K$). Then the value $G(0) = G(c)/2 > 0$ plays a special role. In fact there is a canonical conformal isomorphism ψ_c from the open set $\{z;\ G(z) > G(0)\}$ to the region $\{w;\ \log|w| > G(0)\}$. The map $z \mapsto f(z)$ on this region is conjugate under ψ_c to the map $w \mapsto w^2$, and the equipotentials and dynamic rays in the z-plane correspond to concentric circles and straight half-lines through the origin respectively in the w-plane. In particular, if we choose a constant $G_0 > G(0)$, then the locus $\{z;\ G(z) = G_0\}$ is a simple closed curve, canonically parametrized by the angle of the corresponding dynamic ray. In particular, the critical value $c \in \mathbb{C} \smallsetminus K$ has a well defined external angle, which we denote by $t(c) \in \mathbb{R}/\mathbb{Z}$. Thus $\psi_c(c)/|\psi_c(c)| = e^{2\pi i t(c)}$, and c belongs to the dynamic ray $\mathcal{R}_{t(c)} = \mathcal{R}_{t(c)}^K$.

However, for $G_0 = G(0)$ this locus $\{z;\ G(z) = G(0)\}$ is a figure eight curve. The open set $\{z;\ G(z) < G(0)\}$ splits as a disjoint union $U_0 \cup U_1$, where the U_b are the regions enclosed by the two lobes of this figure eight. (We can express this splitting in terms of dynamic rays as follows. The ray $\mathcal{R}_{t(c)}^K \subset \mathbb{C} \smallsetminus K$ has two preimage rays under f_c, with angles $t(c)/2$ and $(1 + t(c))/2$ respectively. Each of these joins the critical point 0 to the circle at infinity, and together they cut $\mathbb{C}$ into two open subsets, say $V_0 \supset U_0$ and $V_1 \supset U_1$. If c does not belong to the positive real axis, then we can choose the labels for these open sets so that the zero ray is contained in V_0, and $c \in V_1$.) We then cut the filled Julia set K into two disjoint compact subsets $K_b = K \cap U_b$. These constitute a *Bernoulli partition.* That is, for any one-sided-infinite sequence of

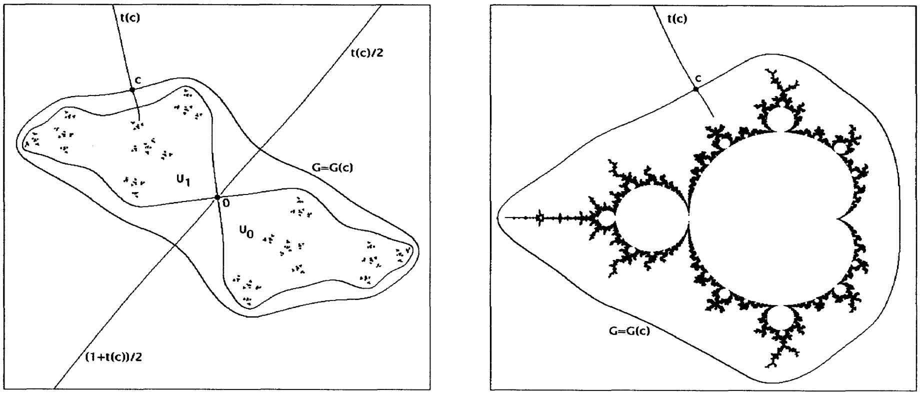


FIGURE 19. Picture in the dynamic plane for a polynomial f_c with $c \notin M$, and a corresponding picture in the parameter plane.

bits $b_0, b_1, \cdots \in \{0,1\}$, there is one and only one point $z \in K$ with $f_c^{\circ k}(z) \in K_{b_k}$ for every $k \geq 0$. To prove this statement, let U be the region $\{z;\ G(z) < G(c)\}$ and let $\phi_b : U \to U_b$ be the branch of f^{-1} which maps U diffeomorphically onto U_b. Using the Poincaré metric for U, we see that each ϕ_b shrinks distances by a factor bounded away from one, and it follows easily that the diameter of the image

$$\phi_{b_0} \circ \phi_{b_1} \circ \cdots \circ \phi_{b_n}(U)$$

shrinks to zero, so that this intersection shrinks to a single point $z \in K$, as $n \to \infty$. *Thus each point of* $J = K$ *can be uniquely characterized by an infinite sequence of symbols* $(b_0, b_1, \ldots)$ *with* $b_j \in \{0,1\}$. In particular, K is homeomorphic to the infinite cartesian product $\{0,1\}^{\mathbf{N}}$, where the symbol $\mathbf{N}$ stands for the set $\{0,1,2,\ldots\}$ of natural numbers. We say that the dynamical system $(K,\ f_c|_K)$ is a *one sided shift* on two symbols.

Similarly, given any angle $t \in \mathbb{R}/\mathbb{Z}$, if none of the successive images $2^k\, t \pmod{\mathbb{Z}}$ under doubling is precisely equal to $t(c)/2$ or $(1 + t(c))/2$, then t has an associated symbol sequence, called its $t(c)$-itinerary, and the ray $\mathcal{R}_t^K$ lands precisely at that point of K which has this symbol sequence. For the special case $t = t(c)$, this symbol sequence characterizes the point $c \in K$, and is called the *kneading sequence* for c or for $t(c)$. (However, if $t(c)$ is periodic, there is some ambiguity since the symbols $b_{n-1}, b_{2n-1}, \ldots$ of the kneading sequence are not uniquely defined in the period n case.)

If t is periodic under doubling, then the itinerary is periodic (if uniquely defined), and the ray $\mathcal{R}_t^K$ lands at a periodic point of K. For further discussion, see **[LS]**, as well as Appendix B.

Here we have been thinking of $c = f(0)$ as a point in the dynamic plane (the z-plane), but we can also think of $c \in \mathbb{C} \setminus M$ as a point in the parameter plane (the

c-plane). In fact Douady and Hubbard construct a conformal isomorphism from the complement of M onto the complement of the closed unit disk by mapping $c \in \mathbb{C} \setminus M$ to the point $\psi_c(c) = \exp(G^K(c) + 2\pi i t(c)) \in \mathbb{C} \setminus \overline{\mathbf{D}}$. *Thus they show that the value of the Green's function on c and the external angle $t(c)$ of c are the same whether c is considered as a point of $\mathbb{C} \setminus K(f_c)$ or as a point of $\mathbb{C} \setminus M$.* In particular, the point $c \in \mathbb{C} \setminus M$ lies on the external ray $\mathcal{R}^M_{t(c)}$ for the Mandelbrot set.

Appendix B
Computing Rotation Numbers

This appendix will outline how to actually compute the rotation number q/r of a periodic point for a map f_c with $c \notin M$. Let $\tau = t(c) \in \mathbb{R}/\mathbb{Z}$ be the angle of the external ray which passes through c. We may identify this critical value angle with a number in the interval $0 < \tau \leq 1$. The two preimages of τ under the angle doubling map $m_2 : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ separate the circle $\mathbb{R}/\mathbb{Z}$ into the two open arcs

$$I(0) = I_\tau(0) = \left(\frac{\tau - 1}{2}, \frac{\tau}{2}\right) \quad \text{and} \quad I(1) = I_\tau(1) = \left(\frac{\tau}{2}, \frac{\tau + 1}{2}\right).$$

(We will write I_τ instead of I whenever we want to emphasize dependence on the critical value angle τ.) For any finite sequence $b_0, b_1, \ldots, b_k$ of zeros and ones, let $\overline{I}(b_0, b_1, \ldots, b_k)$ be the closure of the open set

$$I(b_0, b_1, \ldots, b_k) \;=\; I(b_0) \cap m_2^{-1} I(b_1) \cap \cdots \cap m_2^{-k} I(b_k)$$

consisting of all $t \in \mathbb{R}/\mathbb{Z}$ with $m_2^{\circ i}(t) \in I(b_i)$ for $0 \leq i \leq k$. (Caution: This is not the same as the intersection of the corresponding closures $m_2^{-i}\overline{I}(b_i)$, which may contain additional isolated points.) An easy induction shows that $\overline{I}(b_0, b_1, \ldots, b_k)$ is a finite union of closed arcs with total length $1/2^{k+1}$. *If $\sigma = (b_0, b_1, \ldots)$ is any infinite sequence of zeros and ones, it follows that the intersection*

$$\overline{I}(\sigma) = \bigcap_k \overline{I}(b_0, b_1, \ldots, b_k)$$

is a compact non-vacuous set of measure zero. For each angle $t \in \mathbb{R}/\mathbb{Z}$ there are two possibilities:

Precritical Case. — If t satisfies $m_2^{\circ i}(t) \equiv \tau$ for some $i > 0$, then there will be two distinct infinite symbol sequences with $t \in \overline{I}(b_0, b_1, b_2 \cdots)$. In this case, the associated dynamic ray $\mathcal{R}^K_t$ does not land, but rather bounces off some precritical point for the map f_c. (Compare **[GM]**.)

Generic Case. — Otherwise there will be a unique infinite symbol sequence with $t \in \overline{I}(b_0, b_1, \cdots)$. The corresponding ray $\mathcal{R}^K_t$ will land at the unique point of the Julia set for f_c which has this same symbol sequence, as described in Appendix A. In particular, if t is periodic under doubling, then $\mathcal{R}^K_t$ must land at a periodic point of the Julia set, possibly with smaller period.

***Lemma B.1* (Symbol Sequences and Rotation Numbers).** — *For any symbol sequence $\sigma = (b_0, b_1, \dots) \in \{0,1\}^{\mathbf{N}}$ which is periodic of period p, the map $m_2^{\circ p}$ on the compact set $\overline{I}_\tau(\sigma) \subset \mathbb{R}/\mathbb{Z}$ has a well defined rotation number* $\mathrm{rot}(b_0, \dots, b_{p-1}\,;\,\tau) \in \mathbb{R}/\mathbb{Z}$ *which is invariant under cyclic permutation of the bits b_i. This number increases monotonically with τ, and winds $b_0 + \cdots + b_{p-1}$ times around the circle as τ increases from* 0 *to* 1.

To see this, we introduce an auxiliary monotone degree one map which is defined on the entire circle and agrees with $m_2^{\circ p}$ on $\overline{I}_\tau(\sigma)$. (Compare **[GM]**.) By definition, a *monotone degree one circle map* $\psi : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ is the reduction modulo $\mathbb{Z}$ of a map $\Psi : \mathbb{R} \to \mathbb{R}$ which is monotone increasing and satisfies the identity $\Psi(u+1) = \Psi(u)+1$. Such a Ψ, called a *lift* of ψ, is unique up to addition of an integer constant. The *translation number* of such a map Ψ is defined to be the real number

$$\mathrm{Trans}(\Psi) = \lim_{k\to\infty} \big(\Psi^{\circ k}(u) - u\big)/k.$$

This always exists, and is independent of u. The *rotation number* $\mathrm{rot}(\psi)$ of the associated circle map is now defined to be the image of this real number $\mathrm{Trans}(\Psi)$ under the projection $\mathbb{R} \to \mathbb{R}/\mathbb{Z}$. This is well defined, since $\mathrm{Trans}(\Psi+1) = \mathrm{Trans}(\Psi)+1$. One important property is the identity

$$\mathrm{Trans}(\Psi_1 \circ \Psi_2) = \mathrm{Trans}(\Psi_2 \circ \Psi_1), \tag{9}$$

where Ψ_1 and Ψ_2 are the lifts of two different monotone degree one circle maps. If Ψ_1 is a homeomorphism, this is just invariance under a suitable change of coordinates, and the general case follows by continuity.

Given any $b \in \{0,1\}$, and given a critical value angle τ, define an auxiliary monotone map $\Phi_{b,\tau}$ by the formula

$$\Phi_{b,\tau}(u) = \begin{cases} \min(2u,\, \tau) & \text{if } b = 0, \\ \max(2u,\, \tau) & \text{if } b = 1, \end{cases}$$

for u between $(\tau-1)/2$ and $(\tau+1)/2$, extending by the identity $\Phi(u+1) = \Phi(u)+1$ for u outside this interval. (See Figure 20.) Note that $\overline{I}(b)$ is just the set of points on the circle where the associated circle map $\phi_{b,\tau}$ is not locally constant, and that $\phi_{b,\tau}(u) \equiv 2u \pmod{\mathbb{Z}}$ whenever $u \in \overline{I}(b)$.

For any symbol sequence σ which is periodic of period p, we set $\Phi_{\sigma,\tau}$ equal to the p-fold composition $\Phi_{b_{p-1},\tau} \circ \cdots \circ \Phi_{b_0,\tau}$. (Note that $\overline{I}_\tau(\sigma)$ is just the set of all points $t \in \mathbb{R}/\mathbb{Z}$ such that the orbit of t under the associated circle map $\phi_{\sigma,\tau}$ coincides with the orbit of t under $m_2^{\circ p}$.) This composition is also monotone, with $\Phi(t+1) = \Phi(t)+1$, and therefore has a well defined translation number, which we denote by

$$\mathrm{Trans}(b_0, \cdots, b_{p-1};\ \tau) = \mathrm{Trans}(\Phi_{\sigma,\tau}) \in \mathbb{R}.$$

It follows from property (9) that this translation number is invariant under cyclic permutation of the bits $b_0, \dots, b_{p-1}$. Since each $\Phi_{b,\tau}(u)$ increases monotonically

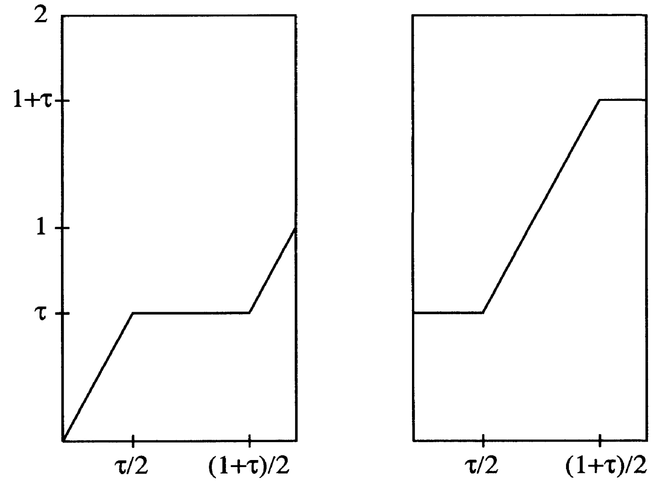


FIGURE 20. Graphs of $\Phi_{0,\tau}$ and $\Phi_{1,\tau}$ (with $\tau = 0.6$).

with τ, with $\Phi_{b,0}(0) = 0$ and $\Phi_{b,1}(0) = b$, it follows easily that $\mathrm{Trans}(\Phi_{\sigma,\tau})$ depends monotonically on τ, increasing from 0 to $b_0 + \cdots + b_{p-1}$ as τ increases from 0 to 1. In other words its image in $\mathbb{R}/\mathbb{Z}$ wraps $b_0 + \cdots + b_{p-1}$ times around the circle as τ varies from 0 to 1. By definition, the rotation number $\mathrm{rot}(b_0, \ldots, b_{p-1}\,;\,\tau)$ of $m_2^{\circ p}$ on the compact set $\overline{I}_\tau(\sigma)$ is equal to the image of the real number $\mathrm{Trans}(\Phi_{\sigma,\tau})$ in the circle $\mathbb{R}/\mathbb{Z}$. □

If a map f_c has critical value angle $t(c) = \tau$, then it is not hard to see that $\mathrm{rot}(b_0, \ldots, b_{p-1}\,;\,\tau)$ coincides with the rotation number as defined in 2.12 for the orbit with periodic symbol sequence $\overline{b_0, \ldots, b_{p-1}} = (b_0, \ldots, b_{p-1}, b_0, \ldots, b_{p-1}, \ldots)$, so long as at least one rational ray lands on this orbit. (Compare [**GM**, Appendix C].)

We will use the notation $\mathcal{S}(q/r)$ for the orbit portrait with orbit period $p = 1$ and rotation number q/r, associated with the q/r-satellite of the main cardioid. (Compare [**G**].) If $\mathcal{P}$ is an arbitrary orbit portrait, then $\mathcal{P} * \mathcal{S}(q/r)$ can be described as its (q/r)-*satellite portrait.* (See 6.4, 8.2.)

To any orbit portrait $\mathcal{P}$ with period $p \geq 1$ and ray period $n = rp \geq p$ we can associate a symbol sequence $\sigma = \sigma(\mathcal{P})$ of period p as follows. Choose any $c \notin M$ in the wake $W_{\mathcal{P}}$, and number the points of the f_c-orbit with portrait $\mathcal{P}$ as $z_0 \mapsto z_1 \mapsto \cdots$, where z_0 is on the boundary of the critical puzzle piece and z_1 is on the boundary of the critical value puzzle piece. Now let $\sigma(\mathcal{P})$ be the symbol sequence for z_0, as described in Appendix A. This is independent of the choice of $c \in W_{\mathcal{P}} \smallsetminus M_{\mathcal{P}}$.

There is an associated *satellite symbol sequence* $\sigma^\star = \sigma^\star(\mathcal{P})$ of period $n = rp$, constructed as follows. (Compare 9.2.) By definition, the k-th bit of $\sigma^\star$ is identical to the k-th bit of σ for $k \not\equiv 0 \pmod{n}$, but is reversed, so that $0 \leftrightarrow 1$, when $k \equiv 0 \pmod{n}$.

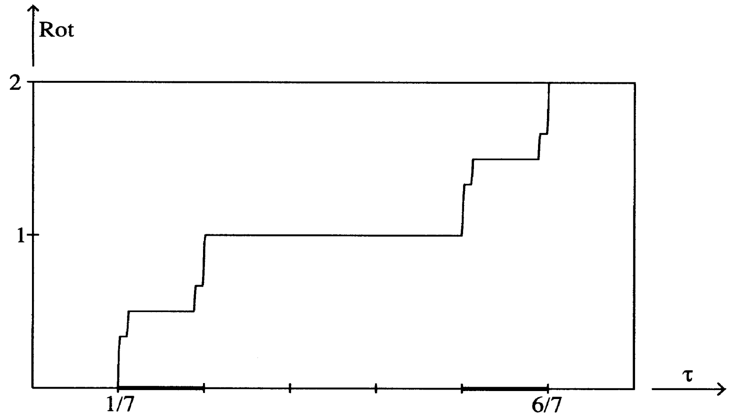


FIGURE 21. The translation number as a function of the critical value exterior angle τ for the period 3 point with symbol sequence $\overline{110} = (1,1,0,1,1,0,\dots)$.

***Lemma B.2* (Satellite Symbol Sequences).** — *For every satellite $\mathcal{P} * \mathcal{S}(q'/r')$ of $\mathcal{P}$, the symbol sequence $\sigma(\mathcal{P} * \mathcal{S}(q'/r'))$ coincides with the satellite sequence $\sigma^\star(\mathcal{P})$. The translation number* $\mathrm{Trans}(\sigma(\mathcal{P}),\tau)$ *is constant for τ in the characteristic arc $\mathcal{I}_\mathcal{P}$, while* $\mathrm{Trans}(\sigma^\star(\mathcal{P}),\tau)$ *increases by $+1$ as τ increases through $\mathcal{I}_\mathcal{P}$, taking the value q'/r'* (mod $\mathbb{Z}$) *on the characteristic arc of $\mathcal{P} * \mathcal{S}(q'/r')$.*

Intuitively, if we tune a map in $H_\mathcal{P}$ by a map in $H_{\mathcal{S}(q'/r')}$ then we must replace the Fatou component containing the critical point for the first map by a small copy of the filled Julia set for a (q'/r')-rabbit. Here the period p point z_0 for $\mathcal{P}$ corresponds to the β-fixed point of this small rabbit, while the period n point z_0 for $\mathcal{P} * \mathcal{S}(q'/r')$ corresponds to the α fixed point for this rabbit. Perturbing out of the connectedness locus M, these two points will be separated by the ray pair terminating at the critical point. Further details will be omitted. □

For example, starting with $\sigma(\{\{0\}\}) = \overline{0}$, where the overline indicates infinite repetition, we find that

$$\sigma(\mathcal{S}(q/r)) = \sigma^\star(\{\{0\}\}) = \overline{1},$$

while

$$\sigma^\star(\mathcal{S}(1/2)) = \overline{01}, \quad \sigma^\star(\mathcal{S}(q/3)) = \overline{011}, \quad \sigma^\star(\mathcal{S}(q/4)) = \overline{0111}, \quad \dots .$$

We can use this discussion to provide a different insight on the counting argument of §5. Since $\mathrm{Trans}(\sigma^\star(\mathcal{P})\,;\,\tau)$ increases by $+1$ on the characteristic arc $\mathcal{I}_\mathcal{P}$, we see that the total number of portraits (or the total number of characteristic arcs) with ray period $rp = n$ is equal to the sum of $b_0 + \cdots + b_{n-1}$ taken over all cyclic equivalence classes of symbol sequences of period exactly n. But the number of such symbol sequences, up to cyclic permutation, is $\nu_2(n)/n$, and the average value of $b_0 + \cdots + b_{n-1}$ is equal to $n/2$, since each symbol sequence with sum different from $n/2$ has an opposite with zero and one interchanged. Therefore, this sum is equal to $\nu_2(n)/2$, as in §5.

Examples. — (Compare Figure 4). Here is a list for all cyclic equivalence classes of symbol sequences of period at most four:

Trans(0; τ) is identically zero.

Trans(1; τ) increases from 0 to 1 for $0 \le \tau \le 1$, taking the value q/r in the characteristic arc for $\mathcal{S}(q/r)$.

Trans(1, 0; τ) increases from 0 to 1 as τ passes through (1/3, 2/3), the characteristic arc for $\mathcal{S}(1/2)$.

Trans(1, 0, 0; τ) increases from 0 to 1 as τ passes through the characteristic arc (3/7, 4/7) for the period 3 portrait with root point $c = -1.75$.

Trans(1, 1, 0; τ) increases by one in the arc (1/7, 2/7) for $\mathcal{S}(1/3)$, and by one more in the arc (5/7, 6/7) for $\mathcal{S}(2/3)$. (Compare Figure 21.)

Trans(1, 0, 0, 0; τ) increases by one in the arc (7/15, 8/15), corresponding to the leftmost period 4 component on the real axis.

Trans(1, 1, 0, 0; τ) increases by one in the arcs (1/5, 4/15) and (11/15, 4/5) associated with the period 4 components on the $1/3^{\text{rd}}$ and $2/3^{\text{rd}}$ limbs. (Figure 12.)

Trans(1, 1, 1, 0; τ) increases by one in the arcs (1/15, 2/15) and (13/15, 14/15) for $\mathcal{S}(1/4)$ and $\mathcal{S}(3/4)$, and also in the arc (2/5, 3/5) for the portrait $\mathcal{S}(1/2) * \mathcal{S}(1/2)$ with root point -1.25.

References

[A] P. Atela, *Bifurcations of dynamic rays in complex polynomials of degree two,* Ergod. Th. Dynam. Sys. **12** (1992) 401–423.

[Be] A. Beardon, "Iteration of Rational Functions", Springer 1991.

[BFH] B. Bielefeld, Y. Fisher and J. Hubbard, *The classification of critically preperiodic polynomials as dynamical systems,* Jour. AMS **5** (1992) 721-762.

[Bou] T. Bousch, *Sur quelques problèmes de la dynamique holomorphique,* Thèse, Université de Paris-Sud (1992).

[Boy1] P. Boyland, *An analog of Sharkovski's theorem for twist maps,* Comtemp. Math. **81** (1988) 119-133.

[Boy2] P. Boyland, *Topological methods in surface dynamics,* Topology and its Applications **58** (1994) 223-298.

[Br] B. Branner, *The Mandelbrot set,* pp. 75-105 of "*Chaos and Fractals, the Mathematics behind the Computer Graphics*", Proc. Symp. Appl. Math **39**, AMS 1989.

[CG] L. Carleson and T. Gamelin, "Complex Dynamics", Springer 1993.

[CM] P. Cvitanović and J. Myrheim, *Complex universality,* Comm. Math. Phys. **121** (1989) 225-254.

[D1] A. Douady, *Systèmes dynamiques holomorphes,* Séminaire Bourbaki, 35^e année 1982-83, n^o **599**; Astérisque **105-106** (1983) 39-63.

[D2] A. Douady, *Julia sets and the Mandelbrot set,* pp. 161-173 of "The Beauty of Fractals", edit. Peitgen and Richter, Springer (1986)

[D3] A. Douady, *Algorithms for computing angles in the Mandelbrot set,* pp. 155-168 of "Chaotic Dynamics and Fractals", ed. Barnsley & Demko, Acad. Press (1986)

[D4] A. Douady, *Chirurgie sur les applications holomorphes,* Proc. Int. Cong. Math. Berkeley, A.M.S. (1987) 724-738.

[D5] A. Douady, *Descriptions of compact sets in* $\mathbb{C}$, pp. 429-465 of "*Topological Methods in Modern Mathematics*", Publish or Perish, Houston (1992)

[DH1] A. Douady and J. Hubbard, *Itération des polynômes quadratiques complexes,* Comptes Rendus Acad. Sci. Paris **294** (1982) 123–126.

[DH2] A. Douady and J. Hubbard, *Etude dynamique des polynômes complexes,* Publications Math. d'Orsay **84-02** (1984) (première partie) and **85-04** (1985) (deuxième partie).

[DH3] A. Douady and J. H. Hubbard, *On the dynamics of polynomial-like mappings,* Ann. Sci. Ec. Norm. Sup. (Paris) **18** (1985) 287-343.

[E] D. Eberlein, *Rational Parameter Rays of Multibrot Sets,* Diploma Thesis, Technische Universität München, in preparation

[G] L. Goldberg, *Fixed points of polynomial maps, I,* Ann. Scient. Ec. Norm. Sup., 4^e série, **25** (1992) 679-685.

[GF] D. Giarrusso and Y. Fisher , *A parametrization of the period* 3 *hyperbolic components of the Mandelbrot set,* Proc. Amer. Math. Soc. **123** (1995) 3731-3737.

[GM] L. Goldberg and J. Milnor, *Fixed points of polynomial maps, II,* Ann. Scient. Ec. Norm. Sup., 4^e série, **26** (1993) 51-98.

[Ha] P. Haïssinsky, *Chirugie parabolique*, Comptes Rendus Acad. Sci. Paris, **327** (1998) 195-198.

[Hu] J. Hubbard, *Local connectivity of Julia sets and bifurcation loci: three theorems of J.-C. Yoccoz,* pp. 467-511 and 375-378 of "*Topological Methods in Modern Mathematics*", Publish or Perish, Houston (1992)

[HS] J. Hubbard and D. Schleicher, *The Spider Algorithm,* "*Complex Dynamics: the Mathematics Behind the Mandelbrot and Julia sets*", pp. 155-180 of Proc. Symposia in Appl. Math. **49** AMS 1994.

[K] J. Kiwi, *Rational Rays and Critical Portraits of Complex Polynomials*, thesis, Stony Brook 1997; Stony Brook IMS Preprint 1997#15.

[LS] E. Lau and D. Schleicher, *Internal Addresses in the Mandelbrot Set and Irreducibility of Polynomials,* Stony Brook IMS Preprint 1994#19.

[La] P. Lavaurs, *Une description combinatoire de l'involution definie par M sur les rationnels à dénominateur impair*, Comptes Rendus Acad. Sci. Paris (4) **303** (1986) 143–146.

[Le] G. Levin, *On the complement of the Mandelbrot set,* Israel J. Math. **88** (1984) 189-212.

[Lo] J. Los, *On the forcing relation for surface homeomorphisms,* preprint, Nice 1994.

[Ly] M. Lyubich, *Dynamics of quadratic polynomials, I, II,* Acta Math. **178** (1997) 185-297.

[Mc] C. McMullen, "Complex Dynamics and Renormalization," Annals of Math. Studies 135, Princeton U. Press 1994.

[M1] J. Milnor, *Self-similarity and hairiness in the Mandelbrot set,* pp. 211-257 of "Computers in Geometry and Topology", edit. Tangora, Lect. Notes Pure Appl. Math. **114**, Dekker 1989 (cf. p. 218 and §5).

[M2] J. Milnor, *Dynamics in One Complex Variable: Introductory Lectures*, Stony Brook IMS Preprint 1990#5.

[M3] J. Milnor, *Local Connectivity of Julia Sets: Expository Lectures,* Stony Brook IMS Preprint 1992#11

[M4] J. Milnor, *On cubic polynomials with periodic critical point*, in preparation.

[PR] C. L. Petersen and G. Ryd, *Convergence of rational rays in parameter space*, preprint, Roskilde Univ. 1998.

[P] A. Poirier, *On postcritically finite polynomials*, parts I and II, Stony Brook IMS Preprint 1993#5 and #7.

[RS] J. Riedl and D. Schleicher, *On the locus of crossed renormalization*, Proceedings 1997 conference on holomorphic dynamics, Kyoto 1998.

[S1] D. Schleicher, *Internal addresses in the Mandelbrot set and irreducibility of polynomials*, Thesis, Cornell 1994.

[S2] D. Schleicher, *Rational external rays of the Mandelbrot set,* to appear in Asterisque; See also Stony Brook IMS Preprint 1997#13.

[S3] D. Schleicher, *On fibers and local connectivity of Mandelbrot and Multibrot sets*, Stony Brook IMS Preprint 1998#13.

[Sh1] M. Shishikura, *On the quasiconformal surgery of rational functions*, Ann. Sci. École Norm. Sup. **20** (1987), 1–29.

[Sh2] M. Shishikura, *The Hausdorff dimension of the boundary of the Mandelbrot set and Julia sets.* Ann. of Math. **147** (1998), 225–267.

[Sø] D. Sørensen, *Complex dynamical systems: rays and non-local connectivity, thesis*, Techn. Univ. Denmark 1994.

[St] N. Steinmetz, "Rational Iteration", de Gruyter 1993.

[Ta] Tan Lei, *Local properties of M at parabolic points*, in preparation.

[Th] W. Thurston, *On the Geometry and Dynamics of Iterated Rational Maps,* Preprint, Princeton University.

J. Milnor, Institute for Mathematical Sciences, State University of New York, Stony Brook, NY 11794-3660 • *E-mail :* `jack@math.sunysb.edu` • *Url :* `http://www.math.sunysb.edu/~jack`

JACOB PALIS

A global view of dynamics and a conjecture on the denseness of finitude of attractors

Astérisque, tome 261 (2000), p. 335-347

<http://www.numdam.org/item?id=AST_2000__261__335_0>

Astérisque
261, 2000, p. 335–347

A GLOBAL VIEW OF DYNAMICS AND A CONJECTURE ON THE DENSENESS OF FINITUDE OF ATTRACTORS

by

Jacob Palis

To Adrien Douady for his lasting contribution to mathematics (July 1995)

Abstract. — A view on dissipative dynamics, i.e. flows, diffeomorphisms, and transformations in general of a compact boundaryless manifold or the interval is presented here, including several recent results, open problems and conjectures. It culminates with a conjecture on the denseness of systems having only finitely many attractors, the attractors being sensitive to initial conditions (chaotic) or just periodic sinks and the union of their basins of attraction having total probability. Moreover, the attractors should be stochastically stable in their basins of attraction. This formulation, dating from early 1995, sets the scenario for the understanding of most nearby systems in parametrized form. It can be considered as a probabilistic version of the once considered possible existence of an open and dense subset of systems with dynamically stable structures, a dream of the sixties that evaporated by the end of that decade. The collapse of such a previous conjecture excluded the case of one dimensional dynamics: it is true at least for real quadratic maps of the interval as shown independently by Swiatek, with the help of Graczyk [GS], and Lyubich [Ly1] a few years ago. Recently, Kozlovski [Ko] announced the same result for C^3 unimodal mappings, in a meeting at IMPA. Actually, for one-dimensional real or complex dynamics, our main conjecture goes even further: for most values of parameters, the corresponding dynamical system displays finitely many attractors which are periodic sinks or carry an absolutely continuous invariant probability measure. Remarkably, Lyubich [Ly2] has just proved this for the family of real quadratic maps of the interval, with the help of Martens and Nowicki [MN].

1. Introduction and Main Conjecture

In the sixties two main theories in dynamics were developed, one of which was designed for conservative systems and called KAM for Kolmogorov-Arnold-Moser. A later important development in this area, in the eighties, was the Aubry-Mather theory

1991 ***Mathematics Subject Classification.*** — 58F10, 58F15.
Key words and phrases. — Attractor, physical measure, homoclinic orbit, stability.

This work has been partially supported by Pronex-Dynamical Systems, Brazil.

for periodic (and Cantori) motions, which has been more recently further improved by Mather. Other outstanding results have been obtained even more recently by Eliasson, Herman, Mañé and others.

The focus of this paper, however, will be the surprising unfolding of the other theory that has been constructed for general systems (nonconservative, dissipative) and called hyperbolic: it deals with systems with hyperbolic limit sets. This means for a diffeomorphism f on a manifold M, that the tangent bundle to M at L, the limit set of f, splits up into df-invariant continuous subbundles $T_L M = E^s \oplus E^u$ such that $df \mid E^s$, $df^{-1} \mid E^u$ are contractions, with respect to some Riemmanian metric. As for most of the concepts in the sequel, a similar definition holds for a non-invertible map g, requiring $dg \mid E^u$ to be invertible. And for a flow X_t, $t \in \mathbb{R}$, we add to the splitting a subbundle in the direction of the vector field that generates the flow

$$T_L M = E^s \oplus E^u \oplus E^0$$

and require for some Riemannian metric and some constants C, $0 < \lambda < 1$, that

$$||dX_t \mid E^s||, \quad ||dX_{-t} \mid E^u|| \leq C\, e^{\lambda t}, \quad t \in \mathbb{R}.$$

See **[PT]**, specially chapter seven, for details and many of the notions presented here.

The concept of hyperbolicity was introduced by Smale, with important contributions to its development as a theory being also given by some of his students at the time, as well as Anosov, Arnold, Sinai and others. Initially, it was created to help pursue the "lost dream" referred to in the abstract: to find an open and dense subset of dynamically (structurally) stable systems; i.e., systems that when slightly perturbed in the C^r-topology, $r \geq 1$, remain with the same dynamics, modulo homeomorphisms of the ambiente space that preserve orbit structures, in the case of flows, or are conjugacies, in the case of transformations. It has actually transcended this objective, loosing through a series of counter-examples its projected character of much universality, i.e. its validity for an open dense subset of systems. But it became the ground basis for a notable evolution that dynamics experienced in the last twenty five years or so. Still, based on previous results, specially by Anosov and by ourselves, Smale and I were able to formulate in 1967 what was called the Stability Conjecture, that would fundamentally tie together hyperbolicity and dynamical stability: a system is C^r-stable if its limit set is hyperbolic and, moreover, stable and unstable manifolds meet transversally at all points. For stability restricted to the limit set, the transversality condition is substituted by the nonexistence of cycles among the transitive (dense orbit), hyperbolic subsets of the limit set.

The theory of hyperbolic systems, i.e. systems with hyperbolic limit sets, was quite developed especially for flows and diffeomorphisms, and it was perhaps even near completion (an exageration!), by the end of the sixties and beginning of the seventies. That included some partial classifications, and an increasing knowledge of their ergodic properties, due to Sinai, Bowen, Ruelle, Anosov, Katok, Pesin, Franks, Williams, Shub, Manning, among several others.

More or less at the same time, the proof of one side of the Stability Conjecture was completed through the work of Robbin, Robinson and de Melo. However, the outstanding part of this basic question from the 60's was proved to be true only in the middle 80's, in a remarkable work of Mañé **[M2]** for diffeomorphisms, followed ten years later by an again remarkable paper of Hayashi **[Ha]** for flows: C^1 dynamically stable systems must be hyperbolic. Before, by 1980, Mañé had proved the two-dimensional version of the result, but independent and simultaneous proofs were also provided by Liao and Sannami. Other partial contributions should be credited to Pliss, Doering, Hu and Wen. A high point in Hayashi's work is his connecting lemma creating homoclinic orbits by C^1 small perturbations of a flow or diffeomorphism: an unstable manifold accumulating on some stable one can be C^1 perturbed to make it intersect one another (the creation of homoclinic or heteroclinic orbits). This fact has been at this very moment sparkling some advance of dynamics in the lines proposed here, as it will be pointed out later.

While the ergodic theory of dynamical systems, as suggested by Kolmogorov and more concretely by Sinai, was being successfully developed, the hope of proposing some global structure for dynamics in general grew dimmer and dimmer in the seventies. This was due to new intricate dynamical phenomena that were presented or suggested all along the decade. First, Newhouse **[N]** extended considerably his previous results, showing that infinitely many sinks occur for a residual subset of an open set of C^2 surface diffeomorphisms near one exhibiting a homoclinic tangency. Perhaps equally or even more striking at the time, was the appearance of attractors having sensitivity with respect to initial conditions in their basin.

Although there are several possible definitions of an attractor, here we will just require it to be invariant, transitive (dense orbit) and attracting all nearby future orbits or at least a Lebesgue positive measure set in the ambient manifold. If A is an attractor for f with basin $B(A)$, we say that it is sensitive to initial conditions, or chaotic if there is $\varepsilon > 0$ such that with total probability on $B(A) \times B(A)$, for each pair of points (x, y) there is an integer $n > 0$ so that $f^n(x)$ and $f^n(y)$ are more than ε apart, where the distances are considered with respect to some Riemannian metric. The definition for flows is entirely similar. Chaotic attractors became also known as strange. The first one, beyond the hyperbolic attractors which are not just sinks, is due to Lorenz **[L]**. Proposed numerically by Lorenz in 1963, it's a rather striking fact that only in the middle seventies most of us became acquainted with Lorenz-like attractors through the examples of Guckenheimer and Williams, which we now call geometric ones. It is still an open and interesting question if the original Lorenz's equations

$$\dot{x} = -10x + 10y$$
$$\dot{y} = 28x - y - xz$$
$$\dot{z} = -(8/3)z + xy$$

in fact correspond to a flow displaying a strange attractor.

Subsequently, again based on numerical experiments, Hénon **[He]** asked about the possible existence of a strange attractor, but now in two dimensions, for certain quadratic diffeomorphisms of the plane

$$f_{a,b}(x,y) = (1 - ax^2 + y, bx)$$

for $a \approx 1.4$ and $b \approx 0.3$. Finally, by the end of the decade, Feigenbaum **[F]** and independently Coullet-Tresser **[CT]** suggested another kind of attractor, now for quadratic maps of the interval and related to a limiting map of a sequence of transformations exhibiting period-doubling bifurcations of periodic orbits. Almost immediately after that, Jacobson **[J]** exhibited strange attractors in the same setting. All this, together with the unsuccessful atempts of the sixties, led to a common belief that perhaps no such a global scenario for dynamics was possible.

However, a series of important results on strange attractors for maps, concerning their persistence, i.e. their existence for a positive Lebesgue measure set in parameter space, and the fact that they carry physical or SRB (Sinai-Ruelle-Bowen) invariant measures, provided by Jacobson **[J]** for the interval, Benedicks-Carleson **[BC]**, Mora-Viana **[MV]**, Benedicks-Young **[BY1]**, **[BY2]** and Diaz-Rocha-Viana **[DRV]** for Hénon-like maps (small C^r perturbations of Hénon maps, $r \geq 1$), were about to take place in the next fifteeen years or so. Perhaps even more striking is the recent proof that they are stochastically stable, a recent remarkable result of Benedicks-Viana **[BV1]**, **[BV2]**. In proving this fact, Benedicks and Viana first showed for Hénon-like attractors, that there are "no holes" in the basin of attraction with respect to the SRB measure, a question I heard from Ruelle and Sinai more than a decade ago: a.e. in the attractor with respect to the SRB measure, there are stable manifolds and their union covers Lebesgue a.a. points in the basin of attraction (and, thus, the union of the stable manifolds is dense in the basin of attraction) **[BV1]**. The concepts of SRB measure and stochastic stability will be presented below. Almost simultaneously, after previous pioneering work of Arnold and Herman (see **[Ar]**, **[H]**), the theory of one-dimensional dynamics experienced a great advance, due to Yoccoz, Sullivan, McMullen, Lyubich, Douady, Hubbard, Swiatek and an impressive number of other mathematicians (see **[dMS]** and **[dM]**).

Such developments, as well as my own work with Takens and Yoccoz, **[PT1]**, **[PT2]**, **[PY1]**, and many inspiring conversations with colleagues, former and present students, among them de Melo, Pujals, Takens, Yoccoz and above all Viana, made me progressively acquire a new global view of dynamics, emphasizing a much more probabilistic approach to the question. I was then able to formulate, just prior to the meeting at the Université Paris-Sud, Orsay in honour of A. Douady, in 1995, the following conjecture:

Global conjecture on the finitude of attractors and their metric stability

(I) *Denseness of finitude of attractors – there is a C^r $(r \geq 1)$ dense set D of dynamics such that any element of D has finitely many attractors whose union of basins of attraction has total probability;*

(II) *Existence of physical (SRB) measures – the attractors of the elements in D support a physical measure;*

(III) *Metric stability of basins of attraction – for any element in D and any of its attractors, for almost all small C^r perturbations in generic k-parameter families of dynamics, $k \in \mathbb{N}$, there are finitely many attractors whose union of basins is nearly (Lebesgue) equal to the basin of the initial attractor; such perturbed attractors support a physical measure;*

(IV) *Stochastic stability of attractors – the attractors of elements in D are stochastically stable in their basins of attraction;*

(V) *For generic families of one-dimensional dynamics, with total probability in parameter space, the attractors are either periodic sinks or carry an absolutely continuous invariant measure.*

As mentioned in the abstract, Lyubich **[Ly2]** solved the last item of the conjecture for families of quadratic maps of the interval, setting the stage for its full solution for one-dimensional real dynamics.

We close this section by briefly recalling the notions of physical or SRB measures and sthocastic stability. Let A be an attractor for a C^r, $r \geq 1$, map and let $B(A)$ be its basin of attraction. We call an invariant probability measure μ with support in A, a physical or SRB measure if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} g(f^i(x)) = \int g \, d\mu$$

for every continuous function g and for a positive Lebesgue measure set of x in $B(A)$. Notice that if the attractor is just a periodic sink, then we take the Dirac measure equally distributed at its orbit as SRB measure. To define stochastic stability, we first assume that μ is an SRB measure as above, with the property holding for a.e. $x \in B(A)$ (like is the case of Lorenz and Hénon-like attractors as commented above). Let now $\{f_n\}$ be independent, identically distributed random variables in the space of C^r maps, $r \geq 1$, and $f_n \in B_\varepsilon(f)$ with (a convenient) probability distribution θ_ε, where $B_\varepsilon(f)$ denotes the ε-neighbourhood of f. For a point $z_0 \in B(A)$, let $z_j = f_j \circ \cdots \circ f_1(z_0)$, $j \geq 1$. We say that (f, A, μ) is stochastically stable on the basin of attraction $B(A)$ if for every neighbourhood V of μ in the weak* topology, the weak* limit of $\frac{1}{n}\sum_{j=0}^{n-1}\delta_{z_j}$ is in V, for every small enough $\varepsilon > 0$ and $(m \times \theta_\varepsilon^{\mathbb{N}})$-almost all $(z_0, f_1, f_2, \dots)$. Here, δ_{z_j} stands for the Dirac measure at z_j and m for Lebesgue measure. We observe that when we consider parametrized families of maps with finitely many parameters, the distribution θ_ε is to be interpreted as random (Lebesgue) choice of parameters. And this is precisely the meaning we want to give to stochastic stability in the conjecture above.

There is another, formally not equivalent, definition of stochastic stability: instead of different maps $f_1, \dots, f_j$ above, one considers only the initial one f but lets the images of an initial point have small random fluctuations. That is, trajectories $(x_0, x_1, x_2, \dots)$ of the perturbed system are obtained by letting each x_{i+1} be chosen at random in a small neighbourhood of $f(x_i)$. However, in all known stable cases, like the hyperbolic systems treated by Kifer and Young (see **[K]**), as well as the Hénon-like diffeomorphisms, dealt with by Benedicks-Viana, both definitions can be applied.

In the case of flows, the definition is basically the same, the random fluctuations taking place at "infinitesimal" intervals of time dt. More formally, one considers a stochastic differential equation like (see **[K]**)

$$dx = X_0(x)\,dt + dY(x),$$

where X_0 is the initial (unperturbed) vector field and dY is a stochastic differential, for instance $dY = \varepsilon dW$ where dW corresponds to the standard Brownian motion. Then, for ε small, the weak* limit of $\frac{1}{T}\int_0^T \delta_{\xi_t}$ should be close to the SRB measure of the vector field X_0 for almost all trajectories of the stochastic flow ξ_t associated to this equation.

At this point, the following remark is in order concerning the definition of attractor. Essentially in all known cases, the basin of attraction is a neighborhood of the set. Still we don't know that such a property is common to a dense subset of all dynamics. For possible relevant situations where the basin contains a set of positive probability but not a full neighborhood of the attractor, random perturbations might cause orbits to escape from the basin of the attractor. In fact this is the case for the random perturbations of flows through Brownian motion, even when the basin is a neighborhood. A related possible situation corresponds to having more than one SRB measure with supports contained in a same (transitive) attractor, so that random orbits cross several of the basins of these measures. In these cases, stochastic stability should mean that the weak* limit of $\frac{1}{n}\sum_{j=0}^{n-1}\delta_{z_j}$ is close to the convex hull of those SRB measures if ε is small (see above and **[V2]**).

2. Other conjectures and some recent results

In his famous essay on the stability of the solar system, written around 1890, Poincaré introduced the notion of homoclinic orbits: in the past and future they converge to the same periodic orbit or are in the intersection of the stable and unstable manifolds (sets) of such a periodic orbit. Subsequently, not only he expressed amazement with the inherent dynamical complexity, but stressed its importance:

"Rien n'est plus propre à nous donner une idée de la complication du problème de trois corps et en général de tous les problèmes de dynamique..."

Indeed, on one hand we have that a transversal homoclinic orbit for a diffeomorphism implies the presence of the dynamics of the horseshoe, as proved by Smale in the early sixties (see **[PT]**). On the other, a generic unfolding of a homoclinic tangency of a dissipative (determinant of the Jacobian smaller than one) surface diffeomorphism yields a rather striking number of rich, intricate dynamical phenomena:

(a) residual subsets of intervals in the parameter line whose corresponding diffeomorphisms display infinitely many coexisting sinks
(b) cascade of period-doubling bifurcations of periodic points (sinks)
(c) positive Lebesgue measure sets in the parameter line whose corresponding diffeomorphism display Hénon-like attractors.

These facts were proved in the late seventies and the eighties by Newhouse, Yorke-Alligood and Mora-Viana extending the fundamental work of Benedicks-Carleson (see **[PT]**). They are also valid in higher dimensions, when the dimension of the unstable manifold of the periodic point is one (codimension-one case) and the product of any two eigenvalues of the derivative of the map at this point has norm less than one (sectionally dissipative). See **[PV]** for the existence of infinitely many coexisting sinks in such a case. It is to be noticed that there are also plenty of values of the parameter such that the corresponding diffeomorphisms have infinitely many coexisting Hénon-like attractors **[C]**.

The results above show how Poincaré once and again had a great mathematical insight. In accordance with his view, I have proposed some time ago the following conjectures:

Conjecture I. — *C^r near any surface diffeomorphism exhibiting one of the above bifurcating phenomena (a), (b), (c), there exists a diffeomorphism displaying a homoclinic tangency, $r \geq 1$. Thus, C^r near any of such bifurcating phenomena we may find all the others. The same in higher dimensions, the homoclinic tangency being now of codimension-one and sectionally dissipative.*

Conjecture II. — *In any dimension, the diffeomorphisms exhibiting either a homoclinic tangency or a (finite) cycle of hyperbolic periodic orbits with different stable dimensions (heterodimensional cycles) are C^r dense in the complement of the closure of the hyperbolic ones, $r \geq 1$.*

Notice that for surface diffeomorphisms, it is not possible to have an heterodimensional cycle. So, we have conjectured that diffeomorphisms exhibiting Hénon-like attractors or repellers (or infinitely many sinks or sources) are dense in the complement of the hyperbolic ones. The same question may be posed for non-invertible maps. Also for flows, but in this case one has to add, to homoclinic tangencies and heterodimensional cycles, flows exhibiting singular attractors, like the Lorenz-like ones, and singular cycles, i.e. cycles involving periodic orbits and singularities, studied in **[LP]**. This conjecture for flows will be mentioned again at the end of the paper, when discussing new Lorenz-like attractors. In general, we have presented in **[PT]**, page 134, a notion of a homoclinic bifurcating dynamical system and formulate a less explicit but similar conjecture in terms of a such a notion.

Progress has been made on conjecture I, by Ures **[U]**, who has showed that for all known cases of diffeomorphisms exhibiting Hénon-like attractors, they may be C^r approximated by one with homoclinic tangencies. Similarly for the limiting map of cascades of period-doubling bifurcations in many interesting cases **[CE1, CE2]**. On the other hand, Pujals and Sambarino seem to be making good progress in showing in any dimension, that a diffeomorphism with infinitely many sinks can be C^1 approximated by one displaying a homoclinic tangency, which is codimension one and sectionally dissipative.

More concretely, Pujals and Sambarino **[PS]** have provided a positive solution to conjecture II for surface diffeomorphisms and $r = 1$. From the comments above, this implies that conjecture I would also be true in the same setting. Remarkably, they also obtain from their method a similar result for surface diffeomorphisms whose topological entropy changes under small C^1 perturbations.

Related to conjecture II, we wish to pose the following weaker version of it:

Conjecture III. — *The subset of dynamical systems that either have their limit set consisting of finitely many hyperbolic periodic orbits or else they have transversal homoclinic orbits, are C^r dense in the set of all dynamical systems, $r \geq 1$.*

Again, Pujals and Sambarino, made some progress towards a positive answer to this question in dimensions higher than two and $r = 1$.

Somewhat related to our main conjecture, stated in the previous section, and initially motivated by **[TY]**, we formulate yet another conjecture.

Conjecture IV. — *When unfolding a homoclinic tangency for a codimension-one sectionally dissipative diffeomorphism, the set of parameter values corresponding to diffeomorphisms with infinitely many sinks or infinitely many Hénon-like attractors has (Lebesgue) measure zero. The same for generic k-parameter families of dynamical systems, $k \in \mathbb{N}$.*

In all the previous assertions concerning homoclinic tangencies for surface diffeomorphisms, the corresponding periodic point may be part of a larger hyperbolic (basic) set. The problem is then translated to the understanding of the arithmetic difference of Cantor sets in the line, obtained from the intersections of the stable and unstable foliations of the hyperbolic set with a line transverse to the leaves at the homoclinic tangency. Such Cantor sets are regular (bounded geometry) if the diffeomorphism is of class C^2 [**PT**]. One can also consider a topology for such Cantor sets. Trying to go beyond the notion of thickness of Cantor sets used by Newhouse (see [**PT**]), I asked if for a residual subset of pairs C_1, C_2 of regular Cantor sets, either $C_1 - C_2$ has measure zero or else it contains an interval. The first case would correspond to $HD(C_1 \times C_2) < 1$ and the second to $HD(C_1 \times C_2) > 1$, where HD stands for Hausdorff dimension. Recently, Moreira and Yoccoz [**Mor**], [**MorY1**], [**MorY2**] have proved this fact, even more strongly, for an open and dense subset of pairs of Cantor sets, extending partial previous results in [**PT1**], [**PT2**] and [**PY1**]. The question for affine Cantor sets is still open:

Conjecture V. — *If C_1, C_2 are affine Cantor sets and $HD(C_1 \times C_2) > 1$, then $C_1 - C_2$ contains an interval. In particular, the question can be posed for most or an open and dense subset of pairs of affine Cantor sets.*

I believe that the result of Moreira-Yoccoz is a major contribution to a more profound understanding of the unfolding of homoclinic bifurcations, which we consider to be central to dynamics. In the previous works, it has been shown that if $HD(C_1 \times C_2) < 1$ then the set of parameter values corresponding to hyperbolicity has full (Lebesgue) density at the value corresponding to the homoclinic tangency [**PT1**], [**PT2**]. Conversely, this is not so if $HD(C_1 \times C_2) > 1$ as proved in [**PY1**]. The results of Moreira-Yoccoz enriches the picture: the set of parameter values corresponding to hyperbolicity or generalized homoclinic tangencies (tangencies between stable and unstable leaves of the foliations) has full density at that same parameter value. In another development, it is shown in [**PY2**] that at the homoclinic bifurcation parameter value, attractors have Lebesgue density equal to zero. This is another indication that attractors perhaps are not so abundant, as Newhouse's result may at first glance suggest.

Concerning attractors in particular carrying an SRB measure, a focusing point of the main conjecture in this article, there has been an explosion of new relevant results, adding to the important development on strange attractors mentioned before. This raises the hope of more progress in the questions posed here as well as in some partial classification of attractors with the properties mentioned in that conjecture.

First of all, Alves [**A**] has shown the existence of SRB measures for Viana's attractors with multidimensional positive Lyapunov exponents [**V1**]. Together with Bonatti, they are announcing a series of results aiming at the following one: a positive Lyapunov exponent on (Lebesgue) many points suffices for the existence of SRB measures [**ABV**].

These results have to do with an important class of diffeomorphisms called partially hyperbolic. This means that there is a continuous invariant decomposition

$$TM = E^1 \oplus E^2$$

where E^1 is uniformly expanding or uniformly contracting and it dominates E^2 [**PT**, page 161]. Shub [**S**], Mañé [**M1**] and Bonatti-Diaz [**BD**] have constructed partially hyperbolic but not hyperbolic diffeomorphisms which are robustly transitive: every C^1 small perturbation is also transitive. Very recently, Diaz-Pujals-Ures [**DPU**] have shown that partial hyperbolicity is in fact a necessary condition for C^1 robust transitiveness in three dimensions. Some time earlier, Pugh and Shub had shown that for volume preserving diffeomorphisms, partial hyperbolicity is a main ingredient for robust ergodicity; see [**PSh**] and references therein.

Concerning flows, Rovella [**R**] has constructed a few years ago an important variation of Lorenz geometric attractor, which is not robust but only persists for perturbations of the flow corresponding to a positive Lebesgue measure set of parameter values. The global spiralling attractors constructed in [**PRV1**], [**PRV2**], whose reduced one-dimensional map exhibits infinitely many critical points, and the critical geometric Lorenz attractors in [**LV1**], [**LV2**] also display measure theoretical persistence. In this setting, de Melo and Martens [**dMM**] have built up families of Lorenz-like maps that are full in the sense that exhibit all combinatorial types. An oustanding question that remained open since the seventies is whether there exist robust transitive attractors for flows containing a singularity with more than one expanding eigenvalue. This has been positively answered very recently by Bonatti-Pumariño-Viana [**BPV**].

The theory of singular or Lorenz-like attractors, specially in dimension three, attained sharp progress in recent years through a series of papers by Morales, Pacifico and Pujals. They have constructed new relevant types of singular attractors in [**MPa**], [**MPu1**], [**MPu2**], [**MPP1**], [**MPP2**] and [**MPP3**]. In some cases, these attractors are across the boundary of hyperbolic flows. They also proved the following partial characterization [**MPP4**]: robustly transitive singular sets, i.e. invariant sets containing singularities, of three dimensional flows are necessarily attractors or repellers and they are singularly hyperbolic, i.e. there exists a continuous invariant decomposition $TM = E^s \oplus E^{cu}$ such that E^s is uniformly contracting, E^{cu} is volume expanding and E^s dominates E^{cu}. Moreover, the derivatives of the flow at the singularities contained in the attractor have eigenvalues with the same relative distribution as in the geometric Lorenz attractors: the unstable (positive) eigenvalue has norm bigger than that of the weak stable (least negative) one. Here again, as well as in [**DPU**] above, Hayashi's connecting lemma is used. Such developments, I believe, should be very inspiring in tackling Conjecture II for flows, at least in three dimensions.

Let me finish by briefly remarking that the scenario we have proposed for most dynamical systems in finite dimension, may apply to some infinite dimensional systems whose solutions are often only defined for positive time (semiflow), like dissipative evolution equations (Euler, Navier-Stokes). Putting the view presented here together with several conjectures/results in this setting, we would have the following conjecture: For most such systems in parametrized form (finitely many parameters) and most initial conditions, the solutions are global and converge to finitely many finite-dimensional attractors with nice ergodic properties, as above.

As a sign of much activity in the area, many of the references below are to appear. They are, however, available at the Internet, for instance at the authors web page at IMPA.

References

[ABV] J.F. Alves, C. Bonatti and M. Viana, *SRB measures for partially hyperbolic systems: the expanding case*, to appear in Invent. Math.

[A] J.F. Alves, *SRB measures for nonhyperbolic systems with multidimensional expansion*, thesis IMPA, 1997, to appear in Ann. Sci. ENS.

[Ar] V.I. Arnold, *On the mapping of the circle into itself*, Izvestia Akad. Nauk., USSR Math. Series bf 25 (1961), 21–86.

[BC] M. Benedicks and L. Carleson, *The dynamics of the Hénon maps*, Ann. of Math. 133 (1991), 73–169.

[BV1] M. Benedicks and M. Viana, *Solution of the basin problem for Hénon-like attractors*, preprint 1999.

[BV2] M. Benedicks and M. Viana, *Stochastic stability and other statistical properties for Hénon-like maps*, in preparation.

[BY1] M. Benedicks and L.-S. Young, *Absolutely continuous invariant measures and random perturbations for certain one-dimensional maps*, Ergod. Th. & Dynam. Sys. 12 (1992), 13–37.

[BY2] M. Benedicks and L.-S. Young, *SBR-measures for certain Hénon maps*, Invent. Math. 112 (1993), 541–576.

[BD] C. Bonatti and L.J. Díaz, *Nonhyperbolic transitive diffeomorphisms*, Ann. of Math. 143 (1996), 357–396.

[BPV] C. Bonatti, A. Pumariño, and M. Viana, *Lorenz attractors with arbitrary expanding dimension*, C.R. Acad. Sci. Paris 325 (1997), 883-888.

[C] E. Colli, *Infinitely many coexisting strange attractors*, to appear, Annales de l'Inst. Henri Poincaré, Analyse Non linéaire 15 (1998), 539-579.

[CE1] E. Catsigeras and H. Enrich, *Homoclinic tangencies near cascades of period doubling bifurcations*, to appear, Ann. de l'Inst. Henri Poincaré Analyse Nonlinéaire.

[CE2] E. Catsigeras and H. Enrich, *Persistence of the Feigenbaum attractor in one-parameter families*, preprint 1997.

[CT] P. Coullet and C. Tresser, *Iterations d'endomorphismes et group de renormalisation*, CRAS Paris 287 (1978), 577–580.

[dMM] W. de Melo and M. Martens, *Universal models for Lorenz maps*, preprint 1997, to appear, Ergodic Th. and Din. Syst.

[dM] W. de Melo, *Rigidity and renormalization in one-dimensional dynamics*, Procs. International Congress of Mathematicians, ICM98-Berlin.

[dMS] W. de Melo and S. van Strien, *One-dimensional dynamics*, Springer-Verlag, 1993.

[DRV] L.J. Díaz, J. Rocha, and M. Viana, *Strange attractors in saddle-node cycles: prevalence and globality*, Invent. Math. 125 (1996), 37– 74.

[DPU] L.J. Díaz, E. Pujals, and R. Ures, *Normal hyperbolicity and robust transitivity*, preprint 1997, to appear in Acta Math.

[F] M. Feigenbaum, *Qualitative universality for a class of nonlinear transformations*, J. Stat. Phys. 19 (1978), 25–52.

[GS] J. Graczyk and G. Swiatek, *Generic hyperbolicity in the logistic family*, Annals of Math. 146 (1997), 1–52.

[Ha] S. Hayashi, *Connecting invariant manifolds and the solution of the C^1 stability and Ω-stability conjectures for flows*, Annals of Math. 145 (1997), 81–137.

[He] M. Hénon, *A two dimensional mapping with a strange attractor*, Comm. Math. Phys. 50 (1976), 69–77.

[H] M. Herman. *Sur la conjugaison différentiable des difféomorphims du cercle à des rotations*, Publ. Math. 49 (1979), 5–233.

[J] M. Jacobson, *Absolutely continuous invariant measures for one-parameter families of one-dimensional maps*, Comm. Math. Phys. 81 (1981), 39–88.

[K] Y. Kifer, *Random perturbations of dynamical systems*, Birkhäuser, 1988.

[Ko] O. Kozlovski, *Structural stability in one-dimensional dynamics*, PhD thesis, Univ. Amsterdam, 1997.

[L] E.N. Lorenz, *Deterministic nonperiodic flow*, J. Atmosph. Sci. 20 (1963), 130–141.

[LP] R. Labarca and M.J. Pacifico, *Stability of singular horseshoes*, Topology 25 (1986), 337–352.

[LV1] S. Luzzatto and M. Viana, *Positive Lyapunov exponents for Lorenz-like maps with criticalities*, these proceedings.

[LV2] S. Luzzatto and M. Viana, *Lorenz-like attractors without invariant foliations*, in preparation.

[Ly1] M. Lyubich, *Dynamics of quadratic polynomials I-II*, Acta Math., 178 (1997), 185–297.

[Ly2] M. Lyubich, *Almost every real quadratic map is either regular or stochastic*, preprint 1997.

[M1] R. Mañé, *Contribution to the stability conjecture*, Topology 17 (1978), 383–396.

[M2] R. Mañé, *A proof of the C^1 stability conjecture*, Publ. Math. I.H.E.S. 66 (1988), 161–210.

[MN] M. Martens and T. Nowicki, *Invariant measures for Lebesgue typical quadratic maps*, these proceedings.

[Mor] C. Moreira, *Stable intersections of Cantor sets and homoclinic bifurcations*, Annales de l'Institut Henri Poincaré, Analyse Non Linéaire, vol. 13, nº 6, (1996), 741–781.

[MorY1] C. Moreira and J.-C. Yoccoz, *Stable intersections of regular Cantor sets with large Hausdorff dimensions*, preprint 1998.

[MorY2] C. Moreira and J.-C. Yoccoz, *Stable homoclinic tangencies for hyperbolic sets of large Hausdorff dimension*, in preparation.

[MV] L. Mora and M. Viana, *Abundance of strange attractors*, Acta Math. 171 (1993), 1–71.

[MPa] C. Morales and M.J. Pacifico, *New singular strange attractors arising from hyperbolic systems*, preprint 1997.

[MPu1] C. Morales and E. Pujals, *Singular strange attractors on the boundary of Morse-Smale systems*, Ann. Sci. E.N.S. 30 (1997), 693–718.

[MPu2] C. Morales and E. Pujals, *Strange attractors containing a singularity with two positive multipliers*, Comm. Math. Phys. 196 (1998), 671-679.

[MPP1] C. Morales, M.J. Pacifico, and E. Pujals, *New strange attractors on the boundary of hyperbolic systems*, preprint 1997.

[MPP2] C. Morales, M.J. Pacifico, and E. Pujals, *Global attractors from the explosion of singular cycles*, C.R. Acad. Sci. Paris 325 (1997), 1317-1322.

[MPP3] C. Morales, M.J. Pacifico, and E. Pujals, *Singular hyperbolic sets*, Proc. Amer. Math. Soc. 127 (1999), 3343-3400.

[MPP4] C. Morales, M.J. Pacifico, and E. Pujals, *On C^1 robust singular transitive sets for three-dimensional flows*, C.R. Acad. Sci. Paris 326 (1998), 81-86.

[N] S. Newhouse, *The abundance of wild hyperbolic sets and nonsmooth stable sets for diffeomorphisms*, Publ. Math. I.H.E.S. 50 (1979), 101–151.

[PRV1] M.J. Pacifico, A. Rovella, and M. Viana, *Infinite-modal maps with global chaotic behaviour*, Annals of Math. 148 (1998), 1-44.

[PRV2] M.J. Pacifico, A. Rovella, and M. Viana, *Persistence of global spiraling attractors*, in preparation.

[PS] E. Pujals and M. Sambarino, *Homoclinic tangencies and hyperbolicity for surface diffeomorphisms: A conjecture of Palis*, preprint 1998, to appear in Annals of Math.

[PSh] C. Pugh and M. Shub, *Stable ergodicity and julienne quasi-conformality*, preprint 1997.

[PT] J. Palis and F. Takens, *Hyperbolicity and sensitive-chaotic dynamics at homoclinic bifurcations*, Cambridge University Press, 1993.

[PT1] J. Palis and F. Takens, *Cycles and measure of bifurcation sets for two-dimensional diffeomorphisms*, Inventiones Math. 82 (1985), 397–422.

[PT2] J. Palis and F. Takens, *Hyperbolicity and the creation of homoclinic orbits*, Annals of Math. 125 (1987), 337–374.

[PV] J. Palis and M. Viana, *High dimension diffeomorphisms displaying infinitely many periodic attractors*, Annals of Math. 140 (1994), 207–250.

[PY1] J. Palis and J.-C. Yoccoz, *Homoclinic tangencies for hyperbolic sets of large Hausdorff dimension*, Acta Mathematica, vol. 172 (1994), 91–136.

[PY2] J. Palis and J.-C. Yoccoz, *Non-uniform hyperbolic horseshoes arising from homoclinic bifurcations*, to appear in C.R. Acad. Sci. Paris.

[R] A. Rovella, *The dynamics of perturbations of the contracting Lorenz attractor*, Bull. Braz. Math. Soc. vol. 24, nº 2 (1993), 233–259.

[S] M. Shub, *Topological transitive diffeomorphisms on T^4*, Lecture Notes in Math. 206 (1971), 39, Springer-Verlag.

[TY] L. Tedeschini-Lalli and J.A. Yorke, *How often do simple dynamical processes have infinitely many coexisting sinks*? Comm. Math. Phys. 106 (1986), 635–657.

[U] R. Ures, *On the approximation of Hénon-like attractors by homoclinic tangencies*, Ergodic Theory and Dynamical Systems, vol. 15 (1995), 1223–1230.

[V1] M. Viana, *Multidimensional nonhyperbolic attractors*, Publ. Math. IHES (1998).

[V2] M. Viana, *Dynamics: a probabilistic and geometric perspective*, Procs. International Congress of Mathematicians, ICM98-Berlin.

J. Palis, Instituto de Matemática Pura e Aplicada - IMPA, Est. D. Castorina 110, 22460-320 Rio de Janeiro, Brazil • *E-mail :* jpalis@impa.br

KEVIN PILGRIM
TAN LEI

Rational maps with disconnected Julia set

Astérisque, tome 261 (2000), p. 349-384
<http://www.numdam.org/item?id=AST_2000__261__349_0>

Astérisque
261, 2000, p. 349–384

RATIONAL MAPS WITH DISCONNECTED JULIA SET

by

Kevin Pilgrim & Tan Lei

Abstract. — We show that if f is a hyperbolic rational map with disconnected Julia set $\mathcal{J}$, then with the possible exception of finitely many periodic components of $\mathcal{J}$ and their countable collection of preimages, every connected component of $\mathcal{J}$ is a point or a Jordan curve. As a corollary, every component of $\mathcal{J}$ is locally connected. We also discuss when a Jordan curve Julia component is a quasicircle and give an explicit example of a hyperbolic rational map with a Jordan curve Julia component which is not a quasicircle.

1. Introduction

For a rational map f of the Riemann sphere $\overline{\mathbf{C}}$ to itself with disconnected Julia set $\mathcal{J}$, we investigate the topological and geometric possibilities for a connected component of $\mathcal{J}$. If $\mathcal{J}$ is disconnected, then f maps components of $\mathcal{J}$ onto components of $\mathcal{J}$, and there are uncountably many such components (*cf.* **[Mi]** and **[Be]**). The *postcritical set* of f

$$\mathcal{P} := \overline{\bigcup_{\substack{n>0 \\ f'(c)=0}} f^{\circ n}(c)}$$

plays a crucial role in our study. We say that f is *hyperbolic* if $\mathcal{P} \cap \mathcal{J} = \varnothing$, *geometrically finite* if $\mathcal{P} \cap \mathcal{J}$ is finite, and *nice* if $\mathcal{P} \cap \mathcal{J}$ is contained in finitely many connected components of $\mathcal{J}$.

Theorem 1.1. — *Let f be a polynomial with disconnected filled Julia set $\mathcal{K}$. Assume that only finitely many connected components of $\mathcal{K}$ intersect $\mathcal{P}$. Then, with the possible exception of finitely many periodic components and their countable collection of preimages, every connected component of $\mathcal{K}$ is a point.*

1991 ***Mathematics Subject Classification.*** — Primary: 58F23; Secondary: 30D05.
Key words and phrases. — Iteration of rational maps, Julia set, connected components, Jordan curve.

Using a variant of this theorem, we establish

Theorem 1.2. — *Let f be a hyperbolic rational map with disconnected Julia set $\mathcal{J}$. Then, with the possible exception of finitely many periodic components and their countable collection of preimages, every connected component of $\mathcal{J}$ is either a point or a Jordan curve.*

The same result for geometrically finite maps can be proved by similar methods as well. We will only sketch the necessary modifications at the end of the paper.

We establish also weaker results for nice maps. The precise statements are given in Propositions [Case 2], [Case 3], [Case 4] and Theorem 9.2.

It is known that $\mathcal{J}$ can be a Cantor set ([**Be**], §1.8), or homeomorphic to the product of a Cantor set with a quasicircle, where each component is a K-quasicircle for some fixed K independent of the component ([**Mc1**]). In §8 we give an explicit example of a hyperbolic rational map which has a Jordan curve Julia component which is not a quasicircle.

Results from plane topology imply that at most countably many Julia components contain an embedding of the letter "Y". Our theorems make precise which ones they are. It is also interesting to see that at most countably many Julia components can be a segment, which a priori is not a restriction from plane topology alone.

By a theorem of McMullen ([**Mc1**], Corollary 3.5), there are at most countably many periodic components of $\mathcal{J}$. Since periodic points of f are dense in $\mathcal{J}$, if $\mathcal{J}$ is disconnected there must be exactly countably many periodic Julia components. Since the degree of f is finite there must be exactly countably many preperiodic Julia components. Hence there are uncountably many wandering Julia components. Our theorems show that under the stated assumptions, no wandering Julia component can be a segment or contain an embedding of the letter "Y", since they must either be points or Jordan curves.

Combining the above theorem with a result of Tan-Yin ([**TY**]), which shows that every preperiodic Julia component for a hyperbolic rational map is locally connected, we answer in the affirmative a question of McMullen [**Bi**]:

Corollary 1.3. — *For a hyperbolic rational map, each Julia component is locally connected.*

This corollary completes another entry in the growing dictionary between the theories of rational maps and Kleinian groups. The analogs of a hyperbolic rational map f and its Julia set $\mathcal{J}$ are a convex compact (or expanding) Kleinian group Γ and its limit set Λ. It is known that each component of Λ is locally connected; the proof depends on the fact that a "wandering" component of Λ, *i.e.* a component with trivial stabilizer, is necessarily a point. See [**AM**] and [**Mc3**] (Theorem 4.18).

The main ideas in our proof are a canonical decomposition $\overline{\mathbf{C}} = E \sqcup \mathcal{U}$, where E is a finite collection of Julia components such that $f(E) \subset E$, and the fact that a hyperbolic rational map f is uniformly expanding on a neighborhood of its Julia set with respect to the Poincaré metric on $\overline{\mathbf{C}} - \mathcal{P}$. Our goal is to show that any Julia component which does not land in E is either a point or a Jordan curve. To this end, we further decompose the sphere into several canonical pieces and measure the itinerary of the orbit of a Julia component J_0 under f with respect to these pieces. We use a combinatorial analysis combined with the lemma below to show that components with certain kinds of itineraries are points. A separate argument treats the case of Jordan curves.

We say that $K \subset \overline{\mathbf{C}}$ is a *full continuum* if K is compact, connected, and $\overline{\mathbf{C}} - K$ is connected. The following Lemma was essentially known to Fatou (see **[Br]**, Thm. 6.2)

***Lemma 1.4* (Fatou).** — *Let f be a rational map, $Q = \bigsqcup_{i=1}^{k} Q_i$ be the union of finitely many disjoint full continua, such that $Q \cap \mathcal{P} = \varnothing$. Then any connected set $J \subset \mathcal{J}$ satisfying $f^n(J) \subset Q$ for infinitely many n is a point.*

Contents. — In §2 we give some motivating examples, define the above mentioned decomposition and state four basic lemmas for nice maps, and give a more precise statement (Theorem 1.2') of Theorem 1.2. We then reduce the proof of Theorem 1.2' to three cases, Cases 2, 3 and 4. Related results for nice maps are stated as well. In §3 we analyze the topology and dynamics of the decomposition and prove the four basic lemmas. §4 contains analytic preliminaries for use in §§5 and 6. In §§5, 6, and 7 we prove the Propositions in Cases 2, 3, and 4, respectively; §7 contains also the proof of Theorem 1.1. §8 lists related results and discusses when a Jordan curve Julia component is a quasicircle. §9 contains sketches of proofs–a generalization our results to the geometrically finite case and some further results for nice maps. §10 is an appendix of technical topological results used in our proofs.

Acknowledgments. — Recently G. Cui, Y. Jiang, and D. Sullivan **[CJS]** have also proven Theorem 1.2 for geometrically finite maps in a different context. Their methods are in some respects similar, but they do not make use of a canonical decomposition. The authors would like to thank Cui for providing a copy of their manuscript, A. Douady, D. Epstein, M. Lyubich, C. McMullen, B. Sevennec and M. Shishikura for many useful discussions, and MSRI for financial support.

2. The decomposition and the reduction to three cases

Let f be a rational map with Julia set $\mathcal{J} = \mathcal{J}(f)$ and postcritical set $\mathcal{P} = \mathcal{P}(f)$.

For J' a continuum (*i.e.* a compact, connected set) in $\overline{\mathbf{C}}$, and P a compact set disjoint from J', we say that J' *separates* P if either $J' \cap P \neq \varnothing$, or $J' \cap P = \varnothing$ and there are at least two components of $\overline{\mathbf{C}} - J'$ intersecting P. We say that J' *separates*

P into exactly q parts if $J' \cap P = \varnothing$ and $\overline{\mathbf{C}} - J'$ has exactly q components intersecting P.

Let J' be a component of $\mathcal{J}$. We say that J' is *critically separating* if J' separates $\mathcal{P}$. We say that two distinct components J' and J'' of $\mathcal{J}$ are *parallel*, if they are both critically separating and the unique annulus component in $\overline{\mathbf{C}} - (J' \cup J'')$ (see Lemma 3.1) does not intersect $\mathcal{P}$.

***Definition* (decomposition $\overline{\mathbf{C}} = E \sqcup \mathcal{U}$, first step).** — Let E be the union of Julia components J' such that either

(1) $J' \cap \mathcal{P} \neq \varnothing$, or
(2) $J' \cap \mathcal{P} = \varnothing$ and J' separates $\mathcal{P}$ into three or more parts, or
(3) J' separates $\mathcal{P}$ into exactly two parts and J' separates no two Julia components which are parallel to J', *i.e.* all Julia components J'' parallel to J' are contained in the same component of $\overline{\mathbf{C}} - J'$.

We think of J' as an *extremal* Julia component. Set $\mathcal{U} = \overline{\mathbf{C}} - E$. The set E may be empty, *e.g.* if $\mathcal{J}$ is a Cantor set.

Examples. — Let

$$f_0(z) = z^2 + 10^{-9} z^{-3} \text{ (McMullen)} \quad \text{and} \quad f_1(z) = \frac{1}{z} \circ (z^2 - 1) \circ \frac{1}{z} + 10^{-11} z^{-3}.$$

The Julia set of f_1, in $\log(z)$-coordinates, is shown in Figure 1. The Julia set of f_0 is homeomorphic to product of a Cantor set with a quasicircle ([**Mc1**]).

For f_1, the point at infinity and -1 form the unique attracting cycle. There are five critical points in the annular Fatou component near the center of the picture, which maps to the Fatou component containing zero (at left) by degree five. $\mathcal{P}$ is contained in the union of the disc Fatou component containing zero (at left) and the immediate basins of infinity (at far right) and -1 (the prominent disc at right). The set E consists of a homeomorphic copy J^+ of the Julia set of $z^2 - 1$, at right, and its preimage J^- (at left) which is a threefold cover of J^+. J^+ and J^- are parallel, J^- maps to J^+, and J^+ is fixed.

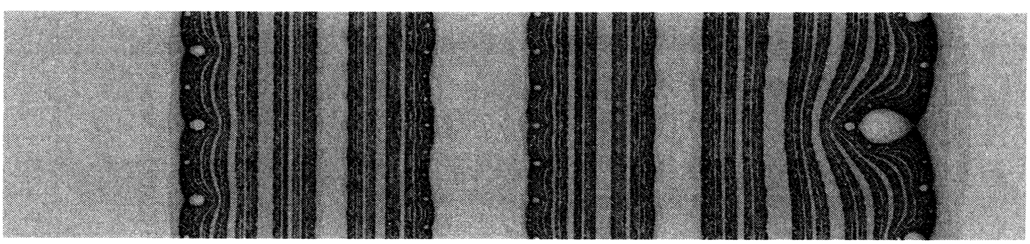


FIGURE 1. Julia set of $\frac{1}{z} \circ (z^2 - 1) \circ \frac{1}{z} + 10^{-11} z^{-3}$

Lemma 2.1. — *If f is nice, the set E consists of at most finitely many Julia components, $f(E) \subset E$, $f^{-1}\mathcal{U} \subset \mathcal{U}$, and every component of E is either periodic or preperiodic.*

This lemma will be proved in the next section. One can also show that E is closed and forward-invariant for any rational map f; we omit the proof.

***Definition* (decomposition of $\mathcal{U}$, second step).** — Define

- $\mathcal{A} = \bigcup$ {annular components of $\mathcal{U}$ disjoint from $\mathcal{P}$};
- $\mathcal{D} = \bigcup$ {disc components of $\mathcal{U}$ disjoint from $\mathcal{P}$};
- $\mathcal{L} = \bigcup$ {components of $\mathcal{U}$ not in $\mathcal{A}$ or $\mathcal{D}$}
  $= \bigcup$ {non-disc non-annular components of $\mathcal{U}$,
  or components of $\mathcal{U}$ intersecting $\mathcal{P}$}.
- $\mathcal{D}' = \bigcup$ {disc components of $\mathcal{U}$ disjoint from $\mathcal{P}$ but intersect $f^{-1}E$};
- $\mathcal{D}'' = \bigcup$ {disc components of $\mathcal{U}$ disjoint from $\mathcal{P} \cup f^{-1}E$}.

Note that $\mathcal{U} = \mathcal{A} \sqcup \mathcal{D} \sqcup \mathcal{L} = \mathcal{A} \sqcup \mathcal{D}' \sqcup \mathcal{D}'' \sqcup \mathcal{L}$.

Example. — Let $f = f_1$. Then $\mathcal{U} = \overline{\mathbf{C}} - E$ is decomposed into:

- $\mathcal{A}$ = a single annulus, bounded by J^- and J^+ ($\overline{\mathcal{A}}$ is not a closed annulus).
- $\mathcal{L}$ = three disc components, each intersecting $\mathcal{P}$. Two are disc components of $\mathcal{U}$ containing the attractor at infinity and -1 with boundaries contained in J^+, and the other is the component of $\mathcal{U}$ containing zero with boundary contained in J^-.
- $\mathcal{D} = \mathcal{D}''$ = the countable set of remaining components of $\mathcal{U}$, all discs with boundaries in E.
- $\mathcal{D}' = \varnothing$.

The set $f^{-1}E$ consists of E plus two other components contained in $\mathcal{A}$ and parallel to components in E.

***Definition* (decomposition of $f^{-1}\mathcal{A}$, third and final step).** — We denote by

- $\mathcal{A}^S$, the union of components A' of $f^{-1}\mathcal{A}$ such that $A' \subset \mathcal{A}$ and $A' \hookrightarrow \mathcal{A}$ is not homotopic to a constant map, and
- $\mathcal{A}^O$, the union of components A' of $f^{-1}\mathcal{A}$ such that $A' \subset \mathcal{A}$ and $A' \hookrightarrow \mathcal{A}$ is homotopic to a constant map.

For f_1, the set $\mathcal{A}^S$ consists of two essential subannuli of $\mathcal{A}$, and the set $\mathcal{A}^O$ is empty. Here is a more precise statement of Theorem 1.2.

Theorem 1.2'. — *Let f be a hyperbolic rational map and $\overline{\mathbf{C}} = E \sqcup \mathcal{U}$ be the decomposition above. Let J_0 be a Julia component. Set $J_n = f^n J_0$. Then exactly one of the following occurs:*

(1) *$J_n \subset E$ for $n \geq n_0$, in which case J_0 is preperiodic, or*

(2) $J_n \subset \mathcal{A}^S$ *for* $n \geq n_0$, *in which case* J_0 *is a Jordan curve, or*
(3) *there is a sequence* $n_k \to \infty$ *such that* $J_{n_k} \subset \mathcal{U} - \mathcal{A}^S$, *in which case* J_0 *is a point.*

If $E = \varnothing$, *the Julia set* $\mathcal{J}$ *is totally disconnected.*

This decomposition is canonical and natural with respect to conjugation by Möbius transformations. While the first decomposition $\overline{\mathbb{C}} = \mathcal{U} \sqcup E$ is the same for any iterate of f, the further decomposition can change.

The following lemmas will be proved in the next section:

Lemma 2.2. — *If* f *is nice, the sets* $\mathcal{L}$, $\mathcal{A}$, $\mathcal{D}'$, $\mathcal{A}^S$ *and* $\mathcal{A}^O$ *all have finitely many components.*

Lemma 2.3. — *Let* f *be a nice map. Then each component* L *of* $\mathcal{L}$ *contains a unique Fatou component* W *such that* $\partial W \supset \partial L$ *and* $W \cap \mathcal{P} = L \cap \mathcal{P}$.

Lemma 2.4. — *Let* f *be a nice map. Then every Julia component* J_0 *is in one of the following four cases: let* $J_n = f^n(J_0)$,
Case 1. *There is* n_0 *such that* $J_n \subset E$ *for* $n \geq n_0$.
Case 2. *There is* n_0 *such that* $J_n \subset \mathcal{A}^S$ *for* $n \geq n_0$.
Case 3. $J_n \subset \mathcal{A}^O \cup \mathcal{D}'$ *for infinitely many* n.
Case 4. $J_n \subset \mathcal{L}$ *for infinitely many* n.
While these cases cover all possibilities, the last two are not mutually disjoint.

This result tegether with Lemma 2.2 means that some finite part of the decomposition encodes a significant portion of the orbit of each Julia components. We are going to prove:

Proposition **(Case 2).** — *In Case 2,* J_0 *is a Jordan curve if* f *is hyperbolic, or* $\overline{\mathbb{C}} - J_0$ *has exactly two components if* f *is nice.*

Proposition **(Case 3).** — *In Case 3,* J_0 *is a point if* f *is hyperbolic, or* $\overline{\mathbb{C}} - J_0$ *is connected if* f *is nice.*

Proposition **(Case 4).** — *In Case 4,* J_0 *is a point if* f *is nice (in particular if* f *is hyperbolic).*

Our cases are also distinguished by our methods of proof. In Case 2, we extract a dynamical system consisting of a finite collection of annuli and covering maps and analyze this restricted system. Case 3 is similar to Case 2. Case 4 is more delicate. We actually prove a stronger result, Theorem 7.1, from which both Proposition [Case 4] and Theorem 1.1 follow as corollaries.

Theorem 1.2 and Theorem 1.2' are direct consequences of the above Propositions.

3. Topology and dynamics of the decomposition

Throughout this section, f denotes a nice rational map. We first prove Lemmas 2.1 and 2.2, and then analyze the topological and dynamical possibilities for the sets in our decomposition in order to prove Lemmas 2.3 and 2.4. We will frequently use the following result from plane topology (see [Ne] for a proof):

Lemma 3.1. — *For a nonempty set $\mathcal{J}$ of disjoint continua J' in S^2, every component of $S^2 - J'$ is a disc (simply connected). Given J' and J'' two disjoint continua, the set $S^2 - (J' \cup J'')$ has a unique annulus component $A(J', J'')$, the component J'' is contained in a component U' of $S^2 - J'$, and*

$$U' = A(J', J'') \sqcup J'' \sqcup \bigcup\{V \mid V \textit{ is a component of } S^2 - J'' \textit{ and } V \cap J' = \varnothing\}.$$

If J^0 is a continuum disjoint from $J' \cup J''$ but separating J' and J'', then

$$A(J', J'') = A(J', J^0) \sqcup A(J^0, J'') \sqcup J^0 \sqcup \bigcup\{V \mid V \textit{ is a component of } S^2 - J^0 \textit{ and } V \cap (J' \cup J'') = \varnothing\}.$$

Proof of Lemma 2.1. — Since f is nice, there exists a compact set $B \subset \overline{\mathbf{C}}$ such that

(1) B has finitely many connected components,
(2) $B \supset \mathcal{P}$, and each connected component of B intersects $\mathcal{P}$,
(3) B contains every Julia component intersecting $\mathcal{P}$ and no other Julia components.

B may be taken to be the union of Julia components intersecting $\mathcal{P}$ together with the suitable preimages of the following: closed, forward-invariant neighborhoods of attracting and superattracting basins; closed, forward-invariant attracting parabolic petals, invariant closed sub-discs of Siegel discs containing points of $\mathcal{P}$, and invariant sub-rings of Herman rings containing points of $\mathcal{P}$. Then a Julia component J is critically separating if and only if it is either contained in B, or is disjoint from $\mathcal{P}$ and separates components of B.

An easy induction argument shows that the number of Julia components J' which separate B into three or more pieces is finite and bounded by $k - 2$ if B has k components. Such Julia components are in E by definition.

Now we deal with the Julia components in E that separate B into exactly two parts. By a method similar to the above, one can prove that if each continuum of the set is disjoint from B and separates B into exactly two parts, and no two continua are parallel (relative to B), then this set of continua is finite. Now assume $J^0 \subset E$ and J^0 separates $\mathcal{P}$ into two parts. We will see that among the Julia components parallel to J^0 at most one of them is contained in E. Let J^1, J^2 be two distinct parallels of J^0. Since J^0 does not separates its parallels, J^1 and J^2 are contained in the same component of $\overline{\mathbf{C}} - J^0$. There are two possible cases:

(a) The component J^1, say, separates J^0 from J^2. Then, according to Lemma 3.1, J^1 is contained in $A(J^0, J^2)$ and $J^1 \hookrightarrow A(J^0, J^2)$ is homotopically non trivial. Since $A(J^0, J^2) \cap \mathcal{P} = \varnothing$, the component J^1 is also parallel to J^2. So J^1 separates $\mathcal{P}$ into exactly two parts, and separates parallels of J^1. Thus J^1 is not contained in E.

(b) We have $J^1 \subset A(J^0, J^2)$ but J^1 does not separate $J^0 \cup J^2$ (therefore $J^2 \subset A(J^0, J^1)$ but J^2 does not separate $J^0 \cup J^1$). This is impossible for the following reason: The annulus $A(J^0, J^2)$ contains J^1 and all but one (disc)-components of $\overline{\mathbf{C}} - J^1$. Since J^1 is critically separating, at least one of these components contains points of $\mathcal{P}$. So $A(J^0, J^2) \cap \mathcal{P} \neq \varnothing$. This is a contradiction to the assumption that J^0 and J^2 are parallel.

So E contains at most finitely many components that separate B into two or three parts, hence E has at most finitely many components.

The following lemma implies that $f(E) \subset E$; since $\mathcal{U} = \overline{\mathbf{C}} - E$, $f^{-1}\mathcal{U} \subset \mathcal{U}$. Therefore each component of E is either periodic or preperiodic. If E is not empty, it consists of finitely many periodic cycles of Julia components, and some (finitely many) of their preimages. □

Proof of Lemma 2.2. — If $E = \varnothing$, we have $\mathcal{L} = \overline{\mathbf{C}}$ and $\mathcal{A} = \mathcal{D}' = \varnothing$. There is nothing to prove.

Assume now $E \neq \varnothing$. Since $f^{-1}E$ has finitely many components, so is $\mathcal{D}'$. It remains to prove that $\mathcal{A} \cup \mathcal{L}$ has finitely many components. Using the notation in the proof of Lemma 2.1, the set $\mathcal{P}$ is contained in B, which consists of finitely many connected components and contains finitely many Julia components. Let U be a disc-component of $\mathcal{L}$ intersecting $\mathcal{P}$. Then either U contains a component of B, or U contains a preimage of a closed parabolic attracting petal containing points of $\mathcal{P}$ used in the construction of B. Since the number of components of B and the number of such petals is finite, the number of disc-components of $\mathcal{L}$ intersecting $\mathcal{P}$ is finite. The other components of $\mathcal{A} \cup \mathcal{L}$ are precisely the non simply connected components of $\overline{\mathbf{C}} - E$. Since E consists of finitely many components, only finitely many components of $\overline{\mathbf{C}} - E$ can be non simply connected. □

Lemma 3.2. — *If a Julia component J' is critically separating then $f(J')$ is also critically separating. If $f(J'')$ of a Julia component J'' separates $\mathcal{P}$ into two parts and separates the parallels of $f(J'')$, then either J'' does not separate $\mathcal{P}$ or J'' separates the parallels of J'' and separates $\mathcal{P}$ into two parts.*

Proof. — We prove that if $f(J')$ is not critically separating then neither is J'. If $J' \cap \mathcal{P} \neq \varnothing$ then $f(J') \cap \mathcal{P} \neq \varnothing$. Hence we may assume that J' and $f(J')$ are disjoint from $\mathcal{P}$. Assume that $\mathcal{P}$ is contained in one (disc)-component of $\overline{\mathbf{C}} - f(J')$. There is a Jordan curve γ separating $f(J')$ and $\mathcal{P}$. Let C be the disc-component of $\overline{\mathbf{C}} - \gamma$ containing $f(J')$. Since $C \cap \mathcal{P} = \varnothing$, every component of $f^{-1}(C)$ is again a disc, and again disjoint from $\mathcal{P}$ since $f^{-1}\mathcal{P} \supset \mathcal{P}$. One of them, say bounded by γ', contains J'. Thus J' does not separate $\mathcal{P}$.

Now we prove the second statement of the lemma. By assumption $\overline{\mathbf{C}} - f(J'')$ has exactly two components U^1 and U^2 intersecting $\mathcal{P}$, and there are Julia components $J^1 \subset U^1$ and $J^2 \subset U^2$ parallel to $f(J'')$. Then $A(J^1, J^2) \cap \mathcal{P} = \varnothing$, since, according to Lemma 3.1, $A(J^1, J^2)$ is the union of $A(J^1, f(J''))$, $A(f(J''), J^2)$, $f(J'')$ and the components of $\overline{\mathbf{C}} - f(J'')$ distinct from U^1 and U^2, and none of these sets intersects $\mathcal{P}$.

Now each component of $f^{-1}(A(J^1, J^2))$ is again an annulus, and again disjoint from $\mathcal{P}$. One of them, say A'', contains J'', and the components of $\partial A''$ are contained in two distinct Julia components J_1' and J_2' which are separated by J''.

Since all but two components of $\overline{\mathbf{C}} - J''$ are contained in A'', and $A'' \cap \mathcal{P} = \varnothing$, the component J'' separates $\mathcal{P}$ into at most two parts. If it does separate $\mathcal{P}$, so do J_1' and J_2'. Hence J_1' and J_2' are parallel to J''. □

Lemma 3.3. — *Suppose $f : S^2 \to S^2$ is a branched covering, $U, V \subset S^2$ are finitely connected open subsets with $f(U) = V$ and $f|U : U \to V$ proper. Then $f(\partial U) = \partial V$ and f maps connected components of ∂U onto connected components of ∂V.*

Remark. — A subtlety is that the map $f : \partial U \to \partial V$ need not be open in the subspace topology, in other words, a component of ∂U may be a proper subset of a component of $f^{-1}(\partial V)$.

Proof. — That $f(\partial U) \subset \partial V$ follows by properness. Since $f(U) = V$, $f(\partial U) \subset \partial V$, and $f(\overline{U})$ is a closed subset containing V, we have $f(\partial U) = \partial V$. Finally, let K' be a boundary component of U and let K be the component of ∂V containing $f(K')$. Since f is a branched covering and V has finitely many boundary components, there are open annuli $A' \subset U$, $A \subset V$ such that K', K are boundary components of A', A respectively and $f : A' \to A$ is a covering map. Then $f|A' : A' \to A$ satisfies the hypotheses of the first conclusion, so $f(\partial A') = \partial A$. It follows easily that $f(K') = K$. □

Lemma 3.4. — *Given any integer n and any component V of $f^{-n}\mathcal{U}$, any Julia component is either contained in V or disjoint from V. The boundary ∂V has finitely many components, and each component is contained in a different Julia component which is disjoint from V.*

Proof. — $f^n(V)$ is a component of $\mathcal{U}$ and the mapping $f^n : V \to f^n(V)$ is proper and finite-to-one. Hence V has finitely many boundary components since $f^n V \in \mathcal{U}$ has finitely many boundary components. Since f^n preserves the Julia components, a Julia component intersecting both V and ∂V would be mapped to a Julia component intersecting both $f^n(V)$ and $\partial f^n(V)$, which is impossible. □

Lemma 3.5. — *Any Julia component J^0 not in E separating $E \cup \mathcal{P}$ is disjoint from $\mathcal{P}$ and critically separating.*

Proof. — Since J^0 is not in E, if it separates E it must separate components of E. Assume that J^1 and J^2 are two Julia components in E, separated by J^0. Then there is a component U^i of $\overline{\mathbf{C}} - J^0$ containing J^i, $i = 1, 2$, and $U^1 \cap U^2 = \varnothing$. By Lemma 3.1 the set U^1 contains all but one component of $\overline{\mathbf{C}} - J^1$. Since J^1 is critically separating, $U^1 \cap \mathcal{P} \neq \varnothing$. Similarly $U^2 \cap \mathcal{P} \neq \varnothing$. Thus J^0 separates $\mathcal{P}$.

Assume now J^1 is a Julia component in E and $x \in \mathcal{P}$ such that J^1 and x are separated by J^0. One can show similarly that J^0 is also critically separating. □

Lemma 3.6. — *Critically separating Julia components are contained in $E \sqcup \mathcal{A}E \sqcup \mathcal{A}^S$.*

Proof. — Here we denote by UV the set of $z \in U$ for which $f(z) \in V$.

We show at first that those Julia components are contained in $E \sqcup \mathcal{A}$. If $J' \cap \mathcal{P} \neq \varnothing$ then $J' \in E$ by definition. Thus we may assume that $J' \cap \mathcal{P} = \varnothing$ and that J' is critically separating and is not in E. Then there are exactly two components U^1 and U^2 of $\overline{\mathbf{C}} - J'$ meeting $\mathcal{P}$, and each U^i contains parallels of J'. Set $P^i = U^i \cap \mathcal{P}$. We will construct a component of $\mathcal{A}$ containing J'.

For $i = 1$ and 2, let

$$\mathcal{W}_i = \{U \mid U \text{ is a component of the complement of some Julia component and } U \cap \mathcal{P} = P^i\}.$$

Set $W^i = \bigcup_{U \in \mathcal{W}_i} U$.

Here we apply the topological result Lemma A.1 to conclude that W^i is an open disc which is also an element of $\mathcal{W}^i$ and is in fact the unique maximal element. Because each U^i contains parallels of J', we have $\partial W^1 \cap \partial W^2 = \varnothing$ and $W^1 \cup W^2 = \overline{\mathbf{C}}$. Thus $W^1 \cap W^2$ is an annulus. To show that this annulus is a component of $\mathcal{A}$, we just need to show $\partial(W^1 \cap W^2) \subset E$ and $W^1 \cap W^2 \cap E = \varnothing$.

Denote by J^i the Julia component containing ∂W^i, $i = 1, 2$.

Assume by contradiction that J^1, say, is not contained in E. Since W^1, as a component of $\overline{\mathbf{C}} - J^1$, meets only part of $\mathcal{P}$, the Julia component J^1 is critically separating. By definition of E, the only possibility for J^1 not being in E is that there is another component W of $\overline{\mathbf{C}} - J^1$ such that $P^2 = \mathcal{P} - P^1$ is contained in W, and there is a Julia component $J^0 \subset W$ parallel to J^1. Thus $\overline{\mathbf{C}} - J^0$ has a component U containing $J^1 \cup W^1$. Furthermore $U \cap \mathcal{P} = W^1 \cap \mathcal{P} = P^1$ (Lemma 3.1). In other words, U is also an element of $\mathcal{W}^1$. This contradicts the fact that W^1 is the maximal element of $\mathcal{W}^1$.

Thus $\partial(W^1 \cap W^2) \subset E$. Now any critically separating Julia component in $W^1 \cap W^2$ would also separate J^1 and J^2, therefore separate $\mathcal{P}$ into two parts and be parallel to both J^1 and J^2. So $W^1 \cap W^2$ contains no component of E. As a consequence,

$W^1 \cap W^2$ is an annulus component of $\overline{\mathbf{C}} - E$ disjoint from $\mathcal{P}$. By definition, $W^1 \cap W^2$ is a component of $\mathcal{A}$.

So $J' \subset \mathcal{A}$. Thus every critically separating Julia component is contained in $E \sqcup \mathcal{A}$.

Now $\mathcal{A}$ is decomposed into $\mathcal{A}E \sqcup \mathcal{A}^S \sqcup \mathcal{A}^O \sqcup \mathcal{A}\mathcal{L} \sqcup \mathcal{A}\mathcal{D}$. For any Julia component J' in $\mathcal{A}\mathcal{L} \sqcup \mathcal{A}\mathcal{D}$, we have $f(J') \subset \mathcal{L} \sqcup \mathcal{D}$. Hence $f(J')$ is not critically separating. By Lemma 3.2, the component J' is not critically separating either. Since the inclusion map of each component of $\mathcal{A}^O$ into $\mathcal{A}$ is homotopic to a constant map, and $\mathcal{A}$ is disjoint from $\mathcal{P}$, no continuum in $\mathcal{A}^O$ can be critically separating.

Thus all critically separating Julia components are contained in $E \sqcup \mathcal{A}E \sqcup \mathcal{A}^S$. □

Lemma 3.7. — *For U a component of $\mathcal{L} \sqcup \mathcal{D}$, there is a unique component U^R of $f^{-1}\mathcal{U}$ which we call a* reduced component *with the following properties:*

(1) *$U^R \subset U$, $\partial U \subset \partial U^R$, and each component of ∂U is a component of ∂U^R;*
(2) *If $U \cap f^{-1}E = \varnothing$ then $U^R = U$. Otherwise U^R is the complement in U of the union of finitely many disjoint full continua, each of which is contained in U;*
(3) *$(U - U^R) \cap \mathcal{P} = \varnothing$;*
(4) *If U is a component of $\mathcal{L}$ then $f(U^R)$ is also a component of $\mathcal{L}$. In particular, $f(\partial\mathcal{L}) \subset \partial\mathcal{L}$.*
(5) *$f(U^R)$ is a component V of $\mathcal{U}$, $f(\partial U^R) = \partial V$, and f maps connected components of ∂U^R onto connected components of ∂V.*
(6) *There are finitely many components U in $\mathcal{D}$ such that $U \cap f^{-1}(\mathcal{P} \cup E) \neq \varnothing$. For any such U, $f(U^R)$ is a component of $\mathcal{L} \cup \mathcal{A}$.*

Proof

(1) and (2) Note that $f^{-1}\mathcal{U} = \overline{\mathbf{C}} - f^{-1}E$. If $U \cap f^{-1}E = \varnothing$, the set U is also a component of $\overline{\mathbf{C}} - f^{-1}E$ (since $f^{-1}E \supset E$), and we set $U^R = U$. Otherwise, let $C_1, \dots, C_k$ denote the components of $f^{-1}E$ which are contained in U. Lemmas 3.5 and 3.6 imply that no C_i separates components of ∂U. Hence for each i, there is a unique component V_i of $\overline{\mathbf{C}} - C_i$ containing ∂U. Let $K_i = \overline{\mathbf{C}} - V_i$ and $K = \cup_i K_i$. Lemma 3.1 implies that either $K_i \cap K_j = \varnothing$ or $K_i \subset K_j$ or $K_j \subset K_i$. Each K_i is full since V_i is connected. Then $U^R := U - K$ has the first two properties in the lemma.

(3) Now we show that no component of ∂U^R separates $(\mathcal{P} \cap U) \cup \partial U$. This is trivial for components of ∂U^R which are also components of ∂U. For the other components of ∂U^R, if this does not hold, there would be a Julia component J' in U separating $\mathcal{P} \cup E$, and thus J' would be critically separating (Lemma 3.5). This is impossible by Lemma 3.6. Therefore $U^R \cap \mathcal{P} = U \cap \mathcal{P}$ and $(U - U^R) \cap \mathcal{P} = \varnothing$.

(4) Assume now that U is a component of $\mathcal{L}$. $f(U^R)$ is a component of $\mathcal{U}$. By definition of $\mathcal{L}$, either $U \cap \mathcal{P} \neq \varnothing$ or U is neither a disc nor an annulus. In the first case

$$f(U^R) \cap \mathcal{P} \supset f(U^R \cap \mathcal{P}) = f(U \cap \mathcal{P}) \neq \varnothing,$$

so $f(U^R)$ is again a component of $\mathcal{L}$. In the second case, either $U^R \cap f^{-1}\mathcal{P} \neq \varnothing$ (in which case $f(U^R) \cap \mathcal{P} \neq \varnothing$ and hence $f(U^R)$ is in $\mathcal{L}$), or $f : U^R \to f(U^R)$ is an unbranched covering. Since U is not a disc or annulus neither is U^R. Any covering over a disc (resp. an annulus) is again a disc (resp. an annulus), so $f(U^R)$ is neither a disc nor an annulus and hence $f(U^R)$ is a component of $\mathcal{L}$.

(5) This follows immediately from the properness of $f : U^R \to V$ and Lemma 3.3.

(6) Suppose $U \cap f^{-1}(\mathcal{P} \cup E) \neq \varnothing$. If $U \cap f^{-1}E \neq \varnothing$, then U is a component of $\mathcal{D}'$. Otherwise U contains a Fatou component intersecting $\mathcal{P}$. By the No Wandering Domains theorem, the number of such Fatou components is finite, hence the number of components U of the latter type is finite. Combining with the fact that $\mathcal{D}'$ has only finitely many components (Lemma 2.2), we get the finiteness. Now if $f(U^R)$ was an element of $\mathcal{D}$, then $f(U^R)$ would be an open disc disjoint from $\mathcal{P}$ and E, hence U would be an open disc disjoint from $f^{-1}(\mathcal{P} \cup E)$. Hence $f(U^R)$ is a component of $\mathcal{A} \cup \mathcal{L}$. □

Lemma 3.8. — *Let A be a component of $\mathcal{A}$ and δ^+ be a component of ∂A. Then there is a unique component A^+ of $f^{-1}\mathcal{U}$ with the following properties:*

(1) *$A^+ \subset A$ and $\delta^+ \subset \partial A^+$;*
(2) *Either A^+ is a component of $\mathcal{A}^S$ or A^+ is a component of $\mathcal{AL}$,* i.e. *$f(A^+)$ is a component of $\mathcal{A}$ or $\mathcal{L}$;*
(3) *$f(\partial A^+) = \partial f(A^+)$ and f maps connected components of ∂A^+ onto connected components of $\partial f(A^+)$. Thus f maps boundary components of components of $\mathcal{A}^S$ onto boundary components of components of $\mathcal{A}$ or $\mathcal{L}$.*

The proof is similar to the one above.

Lemma 3.9. — *Assume $E \neq \varnothing$. For U a component of $\mathcal{U}$, the set $f(U)$ is again a component of $\mathcal{U}$ if and only if $U \cap f^{-1}E = \varnothing$. Either there is a minimal integer $k \geq 0$ such that $f^kU \cap f^{-1}E \neq \varnothing$, or some iterate V of U is a periodic Fatou component. In the latter case, V is finitely connected, is itself a component of $\mathcal{U}$, and either*

(1) *$V \cap \mathcal{P} \neq \varnothing$, and V is a component of $\mathcal{L}$ which is either a simply-connected attracting or parabolic basin, or a Siegel disc or Herman ring intersecting $\mathcal{P}$, or*
(2) *$V \cap \mathcal{P} = \varnothing$, and V is either a Siegel disc and a component of $\mathcal{D}$, or a Herman ring and a component of $\mathcal{A}$.*

Proof. — If $U \cap f^{-1}E \neq \varnothing$ then $f(U) \cap E \neq \varnothing$, and so $f(U)$ can not be a component of $\mathcal{U} = \overline{\mathbf{C}} - E$. Otherwise $U \cap f^{-1}E = \varnothing$ and so U is a component of $f^{-1}\mathcal{U}$ which maps properly under f onto a component of $\mathcal{U}$.

Assume now that for every $n \geq 0$, $f^n(U) \cap f^{-1}E = \varnothing$. Then f^nU is a component of $\mathcal{U}$ for every $n \geq 0$. By Montel's theorem the family $\{f^n|_U\}_n$ is then normal (since E is uncountable if it is nonempty), so U coincides with a Fatou component (since

$\partial U \subset \mathcal{J}$). By the No Wandering Domains Theorem, some iterate V of U is a periodic Fatou component, and so V is a component of $\mathcal{L}$. Since components of $\mathcal{U}$ are finitely connected, V is finitely connected. The Lemma then follows from the classification of periodic Fatou components and the fact that attracting or parabolic basins are either simply connected or infinitely connected ([**Be**], §7.5). □

The following lemma is a more precise version of Lemma 2.3:

Lemma 3.10. — *Assume $E \neq \varnothing$. Let $L_0, L_1, \ldots, L_m$ be the (finitely many) components of $\mathcal{L}$. Then each L_i contains a unique Fatou component W_i such that $\partial W_i \supset \partial L_i$. Moreover the components of ∂L_i are precisely the components of ∂W_i separating $\mathcal{P}$.*

We say $f_*(L_i) = L_j$ if $f(L_i^R) = L_j$. In this case every L_i is preperiodic under f_* and $f(W_i) = W_j$. Furthermore, if $\{L_0, \cdots, L_{p-1}\}$ is a periodic cycle of f_*, then either

(1) *$L_i^R = L_i$ for all $0 \leq i \leq p-1$, in which case $W_i = L_i$ and either*
   (a) *$W_i \cap \mathcal{P} \neq \varnothing$ for all i, in which case W_i is a simply connected attracting or parabolic basin, a Siegel disc or a Herman ring intersecting $\mathcal{P}$, or*
   (b) *$W_i \cap \mathcal{P} = \varnothing$ for all i, in which case W_i is a Siegel disc or Herman ring disjoint from $\mathcal{P}$; or*

(2) *$L_i^R \neq L_i$ for some i, in which case W_i is an infinitely connected attracting or parabolic basin for each $0 \leq i \leq p-1$.*

Proof. — By Lemma 2.2 the set $\mathcal{L}$ has only finitely many components. By Lemma 3.6 no Julia component separates ∂L_i (resp. ∂L_i^R) or is critically separating. Corollary A.5 implies that there is a unique Fatou component W_i (resp. W_i^R) such that $W_i \subset L_i$ and $\partial W_i \supset \partial L_i$ (resp. $W_i^R \subset L_i^R$ and $\partial W_i^R \supset \partial L_i^R$).

By Lemma 3.7, we have $\partial L_i^R \supset \partial L_i$, thus by uniqueness with respect to the property of containing ∂L_i, we have $W_i^R = W_i$.

Note that every connected component of ∂W_i is either contained in ∂L_i or is contained in L_i. Since no Julia component in L_i is critically separating, and every component of ∂L_i is critically separating, the components of ∂W_i separating $\mathcal{P}$ are precisely those in ∂L_i.

If $f_*(L_i) = L_j$, by uniqueness of W_j, we have $f(W_i) = f(W_i^R) = W_j$.

By Lemma 3.7, for any i, $f(L_i^R)$ is again a component of $\mathcal{L}$, thus coincides with some L_j. So $f_*(L_i)$ is well defined for each i. Since there are only finitely many components in $\mathcal{L}$, each of them is eventually periodic under f_*.

Let $\{L_0, \ldots, L_{p-1}\}$ be a periodic cycle of f_*. Then $\{W_0, \ldots, W_{p-1}\}$ forms a periodic cycle of Fatou components.

If $L_i^R = L_i, 0 \leq i \leq p-1$, then the conclusion (1) follows by Lemma 3.9. Otherwise, L_i^R has at least two boundary components, hence W_i has at least two boundary components. W_i cannot be a Siegel disc or Herman ring. For in these cases, $\partial W_i \cap \mathcal{P} \neq$

$\varnothing$ and so $L_i = W_i$ is itself a component of $\mathcal{U}$, contradicting $L_i^R \neq L_i$ for some i. Hence W_i is either an attracting or parabolic basin with at least two boundary components, hence is infinitely connected. $\square$

For our example $f = f_1$ above, $\mathcal{L}$ has a periodic cycle of period 2 formed by the Fatou components containing infinity and -1.

Proof of Lemma 2.4. — Since $f(E) \subset E$, if $J_{n_0} \subset E$ for some n_0, then $J_n \subset E$ for all $n \geq n_0$. This is our case 1.

Assume now $J_n \subset \mathcal{U}$ for all n. Assume furthermore that J_0 is not in Case 2, that is, $J_n \cap \mathcal{A}^S = \varnothing$ for infinitely many n. We are going to show $J_n \subset \mathcal{A}^O \cup \mathcal{D}' \cup \mathcal{L}$ for infinitely many n (so J_0 is in Case 3 or 4 or both).

We show at first that $J_n \subset \mathcal{A}^O \cup \mathcal{D} \cup \mathcal{L} = \mathcal{A}^O \cup \mathcal{D}' \cup \mathcal{D}'' \cup \mathcal{L}$ for infinitely many n. Denote by AB the set of $z \in A$ for which $f(z) \in B$. We have $J_n \cap \mathcal{A}E = \varnothing$ for all n. If $J_n \subset \mathcal{A}\mathcal{D} \cup \mathcal{A}\mathcal{L}$ for some n, then $J_{n+1} \subset \mathcal{D} \cup \mathcal{L}$. Since

$$\mathcal{U} = \mathcal{A}E \cup \mathcal{A}^S \cup \mathcal{A}^O \cup \mathcal{A}\mathcal{D} \cup \mathcal{A}\mathcal{L} \cup \mathcal{D} \cup \mathcal{L},$$

we are done.

We now show that if $J_n \subset \mathcal{D}''$ for infinitely many n, then $J_n \subset \mathcal{L} \cup \mathcal{D}'$ for infinitely many n. Assume $J_{n_1} \subset D$ for D a component of $\mathcal{D}''$. By definition, $D \cap f^{-1}E = \varnothing$. By Lemma 3.7, either $f(D)$ is a component of $\mathcal{D}' \cup \mathcal{D}''$ (this corresponds to the case $D \cap f^{-1}\mathcal{P} = \varnothing$), or $f(D)$ is a component of $\mathcal{L}$. As a consequence of Sullivan's non-wandering domain theorem, there is an integer $0 < k < \infty$ such that D, $f(D)$, ..., $f^{k-1}(D)$ are components of $\mathcal{D}''$ and $f^k(D)$ is a component of $\mathcal{D}' \cup \mathcal{L}$. Therefore $J_{n_1+k} \subset \mathcal{D}' \cup \mathcal{L}$. $\square$

4. Analytic preliminaries

We now restrict to the case when f is hyperbolic. The results generalize to geometrically finite maps; the Poincaré metric ρ is replaced by a more complicated metric for which the map is still expanding (*cf.* **[TY]** and §9).

Recall that f is hyperbolic if and only if $\mathcal{J} \cap \mathcal{P} = \varnothing$. If $|\mathcal{P}| = 2$ then f is conjugate to z^n and $\mathcal{J}$ is connected. Moreover, $\overline{\mathbf{C}} - \mathcal{P}$ is connected. Let $\rho|dz|$ denote the Poincaré Riemannian metric on $\overline{\mathbf{C}} - \mathcal{P}$, $d_\rho(x, y)$ the corresponding distance, and $l_\rho(\gamma)$ the length of a curve with respect to ρ. Then $f : \overline{\mathbf{C}} - f^{-1}(\mathcal{P}) \to \overline{\mathbf{C}} - \mathcal{P}$ is expanding with respect to ρ. If B is the subset given in the proof of Lemma 2.1, then since f is hyperbolic we have $\mathcal{P} \subset B \subset \mathcal{L} \cap (\overline{\mathbf{C}} - \mathcal{J})$, and $f : \overline{\mathbf{C}} - f^{-1}(\mathrm{int}(B)) \to \overline{\mathbf{C}} - \mathrm{int}(B)$ expands ρ uniformly by some definite factor $\lambda > 1$. The inverse of f is then uniformly contracting, in the following sense: if $\gamma : [0, 1] \to \overline{\mathbf{C}} - \mathrm{int}(B)$, then $l_\rho(\widetilde{\gamma}) < (1/\lambda^n) l_\rho(\gamma)$ for any lift $\widetilde{\gamma}$ of γ under f^n. This observation will be the main tool in our proofs of Propositions [Case 2] and [Case 3].

If U is a path-connected subset of $\overline{\mathbf{C}} - \mathcal{P}$, we define the path metric $\mathrm{dpath}_U(x, y)$ on U with respect to ρ by

$$\mathrm{dpath}_U(x, y) = \inf_{\gamma}\{l_\rho(\gamma) \mid \rho : [0, 1] \to U,\ \gamma(0) = x,\ \gamma(1) = y\}.$$

Lemma 4.1. — *Let J' be a periodic or preperiodic Julia component and U a component of $\overline{\mathbf{C}} - J'$. Then there is a C^1 Jordan curve $\gamma : S^1 \to U - B$, a continuous surjective map $h : S^1 \to \partial U$ and a constant L depending on U, such that for each $t \in S^1$, one can find a path $\eta_t : [0, 1] \to \overline{U}$ with ρ-length at most L, such that $\eta_t(]0, 1]) \subset U$, $\eta_t(0) = h(t)$ and $\eta_t(1) = \gamma(t)$.*

Recall that we have a partition $E \sqcup \mathcal{U}$ of $\overline{\mathbf{C}}$, and $\mathcal{A} \sqcup \mathcal{D}$ is the union of disc and annulus components of $\mathcal{U}$ disjoint from $\mathcal{P}$. Moreover $\partial(\mathcal{A} \sqcup \mathcal{D}) \subset E$, and E consists of finitely many preperiodic Julia components.

Corollary 4.2. — *Each component U of $\mathcal{A}, \mathcal{D}, \mathcal{A}^O$, and $\mathcal{A}^S$ has finite path-diameter relative to ρ, and each has locally connected boundary.*

Proof. — We first assume that J' is fixed and that the ideal boundary (*cf.* [**Mc1**]) of U is also fixed by f. In other words, there is a component U' of $f^{-1}U$ such that $U' \subset U$, but $\partial U \subset \partial U'$ and $f(\partial U) = \partial U$. We have $f(U') = U$. Choose $\gamma = \gamma_0 : S^1 \to U$ such that the annulus A between ∂U and $\gamma(S^1)$ contains no points of $f^{-1}B$. Denote by A' the component of $f^{-1}A$ such that $A' \subset U'$ and $\partial U \subset \partial A'$. We may adjust γ so that $A' \subset A$, see [**Mc1**].

Choose $x' \in A'$ such that $f(x') = \gamma(0)$. The degree $d = \deg(f : A' \to A)$ is a positive integer. We define $\gamma_1(t)$ so that $\gamma_1(0) = x'$ and $f(\gamma_1(t)) = \gamma(d \cdot t)$. Let $H_0 : [0, 1] \times S^1 \to A$ be a C^1 map such that $H_0(0, \cdot) = \gamma_1$ and $H_0(1, \cdot) = \gamma$.

Since $f : A' \to A$ is a covering, one can lift H_0 to get $H_1 : [0, 1] \times S^1 \to A'$ such that $H_1(1, \cdot) = H_0(0, \cdot)$. Define $\gamma_2(\cdot) = H_1(0, \cdot)$. One can then define H_{n-1} and γ_n by induction.

To control the convergence, we proceed as follows. For each $t_0 \in S^1$, the ρ-length of the curve $\{H_0(s, t_0),\ s \in [0, 1]\}$ is finite, depending continuously on $t_0 \in S^1$. So it has a finite maximum, say C'. For $t \in S^1$, $d_\rho(\gamma_n(t), \gamma_{n-1}(t))$ is smaller than or equal to the length of the curve $\{H_n(s, t),\ s \in [0, 1]\}$ which is smaller than or equal to C'/λ^{n-1}. So $\{\gamma_n\}$ forms a Cauchy sequence.

Therefore $\gamma_n(t)$ converges uniformly to a limit map, h, and the path distance between $\gamma(t)$ and $\gamma_n(t)$ is uniformly bounded by $C'\lambda/(\lambda - 1)$.

To define η_t for each $t \in S^1$, note that the set $h(t) \cup (\bigcup_{n \geq 0} \bigcup_{s \in [0,1]} H_n(s, t))$ is an embedded closed arc with finite length. Reparametrizing it we get $\eta(t)$.

For periodic J' or periodic ideal boundary, we consider an iterate of f. For preperiodic cases, we pull back the curves given by the result for the periodic cases. □

5. Proof of Proposition [Case 2]

This is the case where a Julia component J_0 satisfies $J_n = f^n(J_0) \subset \mathcal{A}^S$ for $n \geq n_0$. We may assume $n_0 = 0$.

Part I, f is hyperbolic. — We will show that J_0 is a Jordan curve.

Recall that $\mathcal{A}^S$ consists of components of $f^{-1}\mathcal{A}$ parallel to some components of $\mathcal{A}$.

Denote by $A_1, A_2, \ldots, A_p$ the components of $\mathcal{A}^S$, and by A_{ij} the set of points z such that $z \in A_i$, $f(z) \in A_j$.

If $A_{ij} \neq \varnothing$, it is an essential subannulus of A_i, and $f : A_{ij} \to A_j$ is a covering. For $f : A_i \to f(A_i)$ is a covering, $f(A_i)$ is a component of $\mathcal{A}$ and A_j is a subannulus of $f(A_i)$ parallel to $f(A_i)$.

For each A_j, choose $\gamma_j : S^1 \to A_j$ an injective homotopically non-trivial C^1 curve. For each i such that $A_{ij} \neq \varnothing$, choose $x_{ij} \in A_{ij}$ such that $f(x_{ij}) = \gamma_j(0)$. Then the homotopy classes of γ_i and γ_j determine uniquely generators for $\pi_1(A_{ij})$ and $\pi_1(A_j)$. The degree $d_{ij} = \deg(f : A_{ij} \to A_j)$ is then a positive or negative integer. Define a lift $\gamma_{ij}(t)$ of γ_j so that $\gamma_{ij}(0) = x_{ij}$ and $f(\gamma_{ij}(t)) = \gamma_j(d_{ij} \cdot t)$. Let $H_{ij} : [0,1] \times S^1 \to A_i$ be a C^1 map such that $H_{ij}(0, \cdot) = \gamma_i$ and $H_{ij}(1, \cdot) = \gamma_{ij}$.

Let $\alpha = (a_0 a_1 \cdots)$ be any infinite sequence such that for all $n \geq 0$ we have $1 \leq a_n \leq p$ and $A_{a_n a_{n+1}} \neq \varnothing$ and call such a sequence *admissible.*

For $n \geq 0$, denote by $T_n = T_n(\alpha)$ the set of points z such that $f^k(z) \in A_{a_k}$ for $0 \leq k \leq n$. Then T_{n+1} is an essential subannulus of T_n for $n \geq 0$ and $f^n : T_n \to A_{a_n}$ is a covering. The curve γ_{a_0} determines a generator for $\pi_1(T_n)$ and we let $d_n = \deg(f^n : T_n \to A_{a_n})$; it can be a positive or negative integer.

Set $J' = J'(\alpha) = \bigcap_n \overline{T}_n$. Since $\overline{T}_n$ forms a nested sequence of compact connected sets which are critically separating, J' is also compact connected and critically separating. We will show that either $J' \subset \partial T_N$ for some N, or J' is a Jordan curve, and a Julia component. The proof is split into several lemmas.

Lemma 0. — *For each $n \geq 0$, there is a (parametrized) C^1 curve $\zeta_n(t)$, and for $n \geq 1$, a homotopy $G_n : [0,1] \times S^1 \to T_{n-1}$ such that $\zeta_n(S^1) \subset T_n \subset A_{a_0}$, $G_n(0,t) = \zeta_{n-1}(t)$, $G_n(1,t) = \zeta_n(t)$. Moreover $f^n(\zeta_n(t)) = \gamma_{a_n}(d_n \cdot t)$.*

Proof. — Set $\zeta_0 = \gamma_{a_0}$. Assume we have constructed ζ_{n-1} and G_{n-1}. Since the map $f^{n-1} : T_{n-1} \to A_{a_{n-1}}$ is a covering, mapping T_n onto A_{a_{n-1}, a_n}, one can lift the homotopy H_{a_{n-1}, a_n} to a map $G_n : [0,1] \times S^1 \to T_{n-1}$ with $G_n(0,t) = \zeta_{n-1}(t)$. Set $\zeta_n(t) = G_n(1,t)$. These are the maps required by the lemma. Finally, by our choice of generators of fundamental groups, we have $f^n(\zeta_n(t)) = \gamma_{a_n}(d_n \cdot t)$. □

Next, by our choice of B (in particular $B \cap \mathcal{J} = \varnothing$) we have $\bigcup \overline{A}_i \subset \overline{\mathbf{C}} - B$ and $\overline{A}_{ij} \subset \overline{\mathbf{C}} - f^{-1}(B)$ for all possible pairs (i,j). By Corollary 4.2, there is a positive number M such that for $i = 1, \ldots, p$, the path diameter of $A_i \in \mathcal{A}^S$ with respect to ρ

is at most M. Moreover, on $\overline{\mathcal{A}^S} \subset \overline{\mathbf{C}} - f^{-1}(\mathrm{int}(B))$, f expands ρ by a definite factor $\lambda > 1$.

Lemma 1. — *The curves $\zeta_n(t)$ defined in Lemma 0 converge uniformly to a continuous map $\zeta : S^1 \to \overline{A}_{a_0}$.*

Proof. — For each possible pair (i,j), the ρ-length of the curve $\{H_{ij}(s,t_0),\ s \in [0,1]\}$ is finite, depending continuously on $t_0 \in S^1$. So it has a finite maximum. Let C' be the maximum of the ρ-length of the curves among all possible couples (i,j) and all $t_0 \in S^1$; it is again finite. For $t \in S^1$, $d_\rho(\zeta_n(t), \zeta_{n-1}(t))$ is less than or equal to the length of the curve $\{G_n(s,t),\ s \in [0,1]\}$ which is smaller than or equal to C'/λ^{n-1}. So $\{\zeta_n\}$ forms a Cauchy sequence. □

Choose a base point $x^\pm$ in each component of ∂A_{a_0}. Denote by δ_n^+ (resp. δ_n^-) the component of ∂T_n which either contains x^+ (resp. x^-) or which separates x^+ (resp. x^-) and T_n. Denote by D_H the Hausdorff distance on compact subsets of $\overline{\mathbf{C}} - \mathrm{int}(B)$ with respect to the metric d_ρ. By definition

$$D_H(F,G) = \max\left(\ \max_{x\in F}\min_{y\in G} d_\rho(x,y),\ \max_{y\in G}\min_{x\in F} d_\rho(x,y)\ \right).$$

Lemma 2. — *For every $\varepsilon > 0$, there exists an N independent of α such that for every $n \geq N$, $D_H(J', \delta_n^+) < \varepsilon$, $D_H(J', \delta_n^-) < \varepsilon$ and $D_H(J', \overline{T}_n) < \varepsilon$.*

Lemma 3. — $D_H(J', \zeta_n(S^1)) \to 0$.

Proof of Lemmas 2 and 3. — Fix $y \in \delta_n^+$. We first show

$$\min_{x\in J'} d_\rho(x,y) < M/\lambda^n.$$

Let $y_n = f^n(y), n \geq 0$, and for each n choose $x_n \in f^n(J')$. Then for each n, there is a path $\eta_n : [0,1] \to \overline{A}_{a_n}$ such that $\eta_n(0) = y_n$, $\eta_n(1) = x_n$, $\eta_n(]0,1[) \subset A_{a_n}$, and $l_\rho(\eta_n) \leq M$. For any n, $f^{-n}(f^n(J')) \cap \overline{T}_n = J'$, since $f^n(\bigcap_k \overline{T}_k) = \bigcap_k f^n\overline{T}_k$. Hence there is a lift $\widetilde{\eta}_n : [0,1] \to \overline{T}_n$ of η_n under f^n joining y to some point $x'_n \in J'$. Hence by expansion

$$\min_{x\in J'} d_\rho(x,y) \leq d_\rho(x'_n, y) \leq l_\rho(\widetilde{\eta}_n) \leq M/\lambda^n.$$

Hence

$$\max_{y\in\delta_n^+}\min_{x\in J'} d_\rho(x,y) \leq M/\lambda^n.$$

A similar argument bounds $\max_{x\in J'}\min_{y\in\delta_n^+} d_\rho(x,y)$ by the same quantity. The remainder of the two lemmas are proved similarly. □

Lemmas 1 and 3 imply that $\zeta(S^1) = J'$. As a consequence, J' is locally connected.

Lemma 4. — *Either J' coincides with one boundary component of T_N for some N, or $J' \subset T_n$ for all n. In the second case, J' is a Jordan curve, and a Julia component.*

Proof. — We first show $J' \subset \mathcal{J}$. First, $\partial J' \subset \mathcal{J}$ since $\partial T_n \subset \mathcal{J}$ for all n and $\mathcal{J}$ is closed. Second, $J' = \partial J'$. For otherwise there is a nonempty component W of $\mathrm{int}(J')$. Then $W \subset T_n$ for all n, and as a consequence $f^n(W) \subset \mathcal{A}$ for all n. Thus $\{f^n\}$ is normal on W. So W coincides with a Fatou component. But for a hyperbolic map every Fatou component is eventually attracting, and meets eventually $\mathcal{P}$, contradicting $\mathcal{A} \cap \mathcal{P} = \varnothing$.

There are thus two possibilities: either $J' \subset \partial T_N$ for some N, or $J' \subset T_n$ for all n. In the first case J' coincides with one boundary component of T_k for all $k \geq N$ (Lemma 2). In the second case, J' must be a Julia component. For otherwise, J' is a nonempty proper closed subset of some Julia component J''; if $x \in J'' - J'$ then $D_H(x, J') > 0$, and hence Lemma 2 implies that for some n, ∂T_n either separates x and J' or ∂T_n contains x. But this implies that J'' intersects ∂T_n, hence J' is contained in a boundary component of T_n, violating our assumption.

Moreover, Lemma 2 implies that $\overline{\mathbf{C}} - J'$ has exactly two components U_1, U_2, and $\partial U_1 = \partial U_2 = J'$. The lemma below (pointed out to us by M. Lyubich) allows us to conclude that J' is a Jordan curve. □

Consider now our Julia component J_0 such that $J_n \subset \mathcal{A}^S$ for all n. It determines an admissible sequence $\alpha = (a_0 a_1 \cdots)$ by setting $a_n = m$ if $J_n \subset A_m$. Then $J_0 = J'(\alpha)$, and it is a Jordan curve.

Remark. — The proof actually shows much more; see §8.

Part II, f is nice. — Let J_0 be a Julia component such that $J_n \subset \mathcal{A}^S$ for all n. Define T_n to be the component of $f^{-n}\mathcal{U}$ containing J_0 (it is in fact the same T_n as in Part I). Then each T_n is an open annulus, contained essentially in T_{n-1}. With the help of Sullivan's non-wandering domain theorem, one can show easily that $J_0 = \bigcap_n T_n$. On the other hand, since J_0 is disjoint from $\partial T_n \subset f^{-n}(E)$ for all n, there is a sequence $n_k \to \infty$ such that $\overline{T_{n_k}} \subset T_{n_{k-1}}$. Therefore $\overline{\mathbf{C}} - J_0$ has exactly two components.

Lemma 5.1. — *Assume that K is a closed subset of $\overline{\mathbf{C}}$ satisfying either conditions a) and b) or condition c):*

a) K is the common boundary of two disjoint open connected sets U_1 and U_2.

b) K is locally connected.

c) $\overline{\mathbf{C}} - K$ has exactly two components V_1 and V_2 and each point of K is accessible from both V_1 and V_2.

Then K is a Jordan curve. There are counter examples if one of the above conditions is not satisfied.

Proof. — Condition a) shows that U_1 and U_2 are simply connected and K is compact connected. By b) and Carathéodory's theorem, a Riemann map $\phi : \Delta \to U_1$ extends continuously to the boundary, and the extension is locally non constant. But the extension is also injective, for otherwise the image by ϕ of a pair of distinct radial segments of Δ would form a Jordan curve ν and each component of $\overline{\mathbf{C}} - \nu$ would intersect U_2. The contradicts the fact that $\nu \cap U_2 = \varnothing$ and U_2 is connected.

For a proof in condition c) case, see Newman ([**Ne**]), Theorem 16.1. □

6. Proof of Proposition [Case 3]

This is the case where a Julia component J_0 satisfies $J_n = f^n(J_0) \subset \mathcal{A}^O \cup \mathcal{D}'$ for infinitely many n.

Part I. f is hyperbolic. — We will need the following variant of Lemma 1.4; the difference is that we do not assume $\overline{Q}_i$ is full.

Lemma 6.1. — *Let f be a hyperbolic rational map and $Q = \sqcup_{i=1}^k Q_i$ be a finite family of disjoint path-connected subsets of $\overline{\mathbf{C}}$ such that*

(1) *$\overline{Q}_i \cap \mathcal{P} = \varnothing$ for each i,*
(2) *the inclusion maps $\iota_i : Q_i \to \overline{\mathbf{C}} - \mathcal{P}$ are homotopic to constant maps, and*
(3) *the path-diameter of each Q_i with respect to dpath_{Q_i} is finite.*

Then any connected set J satisfying $f^n(J) \subset Q$ for infinitely many n is a point.

Proof. — We may assume $\overline{Q}_i \cap B = \varnothing$, where B is a closed neighborhood of $\mathcal{P}$ constructed previously. Then since f is hyperbolic, f expands ρ uniformly by some factor $\lambda > 1$ on $\overline{\mathbf{C}} - f^{-1}(\mathrm{int}(B))$. Let $J_n = f^n(J)$ and let Q^n denote the component of Q containing J_n. Choose $x_0, y_0 \in J_0$ and let $x_n = f^n(x_0), y_n = f^n(y_0)$. Let M be an upper bound on the path-diameters of the Q_i with respect to ρ. Then for each n, there is a path $\eta_n : [0,1] \to Q^n$, $\eta_n(0) = x_n$, $\eta_n(1) = y_n$ for which $l_\rho(\eta_n) \le M$. Since $Q^n \hookrightarrow \overline{\mathbf{C}} - \mathcal{P}$ is homotopic to a constant map, for each n there is a lift $\widetilde{\eta}_n$ of η_n under f^n joining x_0 to y_0. By expansion,

$$d_\rho(x_0, y_0) \le l_\rho(\widetilde{\eta}_n) \le l_\rho(\eta_n)/\lambda^n \le M/\lambda^n \to 0.$$

Hence $x_0 = y_0$ and J is a point. □

Now assume that J_0 is a Julia component for f, and $J_n \subset \mathcal{A}^O \cup \mathcal{D}'$ for infinitely many n. By Lemma 2.2 the set $\mathcal{A}^O \cup \mathcal{D}'$ has finitely many components, each of finite path diameter by Corollary 4.2. The above Lemma then applies and hence J_0 is a point.

Part II. f is nice. — For each component U of $\mathcal{A}^O \subset \mathcal{A}$, take a simple closed curve γ which is a generator of the fundamental group of U. Then $\overline{\mathbf{C}} - \gamma$ has exactly one disc component V contained in $\mathcal{A}$. The set $\widehat{U} = U \cup V$ is an open disc contained in $\mathcal{A}$ (so is disjoint from $\mathcal{P}$. Therefore the enlarged set $\widehat{\mathcal{A}^O} \cup \mathcal{D}'$ consists of finitely many open discs disjoint from $\mathcal{P}$. The rest of the proof is very similar to Part II in the proof of Proposition [Case 2]. We omit the details here.

7. Proof of Proposition [Case 4] and Theorem 1.1

In this section we derive Theorem 1.1 and Proposition [Case 4] from

Theorem 7.1. — *Let f be a rational map (not necessarily hyperbolic or nice). Let $\mathcal{W}$ be the union of finitely many Fatou components $W_0, \dots, W_m$ such that*

(1) $f(\mathcal{W}) \subset \mathcal{W}$;
(2) *each W_i eventually lands on an attracting or parabolic periodic basin under iteration of f, and $W_i \cap \mathcal{P} \neq \varnothing$ for all i;*
(3) *for each W_i, only finitely many components $K_{1,i}, \dots, K_{m_i,i}$ of $\overline{\mathbf{C}} - W_i$ intersect $\mathcal{P}$.*

Then there is a finite union Q of disjoint full continua in $\overline{\mathbf{C}} - \mathcal{P}$ such that any Julia component J_0 satisfying $f^n(J_0) \subset \bigcup_i \left(\overline{\mathbf{C}} - \bigcup_{j=1}^{m_i} K_{j,i}\right)$ for infinitely many n passes infinitely often through Q, and is a point.

Proof of Theorem 1.1. — Let W_0 be the basin of infinity of the polynomial f. It is a fixed attracting component, and infinitely connected. Since only finitely many components of $\mathcal{K} = \overline{\mathbf{C}} - W_0$ intersect $\mathcal{P}$, we may apply the above theorem to prove that every Julia component of f passing infinitely many times through U_0 is a point, where

$$U_0 = W_0 \cup \{K \mid K \text{ is a } \mathcal{K}\text{-component disjoint from } \mathcal{P}\}.$$

On the other hand, every Julia component is the boundary of a $\mathcal{K}$-component. So every $\mathcal{K}$-component passing infinitely many times through U_0 is a point. But in this particular case, we know also that the orbit of a $\mathcal{K}$-component either stays entirely in U_0 or lands eventually on a $\mathcal{K}$-component intersecting $\mathcal{P}$, which is preperiodic, since there are only finitely many of such $\mathcal{K}$-components and the union of them is forward invariant. □

Proof of Proposition [Case 4]. — Let f be a nice map and J_0 be a Julia component. Assume that $J_n = f^n(J_0) \subset \mathcal{L}$ for infinitely many n. For each component L_i of $\mathcal{L}$, Lemma 3.10 provides a unique Fatou component W_i such that $W_i \subset L_i$ and $\partial L_i \subset \partial W_i$. The union of these W_i's satisfies the conditions of Theorem 7.1. Moreover, in the notation of the statement and the proof of the theorem, we have $L_i = U_i = \overline{\mathbf{C}} - \bigcup_{j=1}^{m_i} K_{j,i}$. Therefore we can apply Theorem 7.1 to conclude that J_0 is a point. □

We now start the proof of Theorem 7.1. — We will use the following notation. Let S be a closed subset of S^2 and W be an open connected subset of S^2. Define

$$U(W, S) = W \cup \bigcup \{K \mid K \text{ a component of } S^2 - W \text{ such that } K \cap S = \varnothing\}.$$

Then $U(W, S)$ is an open set.

For $i = 0, \dots, m$, set $U_i = U(W_i, \mathcal{P}) = \overline{\mathbf{C}} - \bigcup_{j=1}^{m_i} K_{j,i}$. Our aim is to find a compact set Q satisfying the properties in the theorem. We then apply Lemma 1.4 to obtain the result.

The proposition below is proved in the Appendix.

Proposition 7.2. — *Let W be a Fatou component of a rational map f. Suppose that $\mathcal{P} \cap (\overline{\mathbf{C}} - W)$ is contained in the union $K_1 \sqcup \cdots \sqcup K_r$ of finitely many connected components of $S^2 - W$. Let W' be a connected component of $f^{-1}(W)$ and let $U = U(W, \mathcal{P})$ and $U' = U(W', f^{-1}(\mathcal{P}))$. Then*

(1) *U and U' are connected open subsets with finitely many boundary components;*
(2) *$f : U' \to U$ is proper, surjective and a branched covering;*
(3) *$f : \partial U' \to \partial U$ is surjective, and given any connected component K'_i of $S^2 - U'$, f maps $\partial K'_i$ surjectively onto ∂K_j for some unique component K_j of $S^2 - U$.*

Since each component W_i contains points of $\mathcal{P}$, $U_i \cap U_j = \varnothing$, $i \neq j$. For each i, set $U'_i = U(W_i, f^{-1}\mathcal{P})$. Then $U'_i \subset U_i$ since $f^{-1}(\mathcal{P}) \supset \mathcal{P}$. By Proposition 7.2, U'_i has finitely many boundary components, hence $Q_i := U_i - U'_i$ consists of finitely many full continua disjoint from $f^{-1}(\mathcal{P})$, hence disjoint from $\mathcal{P}$. This will be one piece of our set Q.

Proposition 7.2 also implies that $f : U'_i \to U_j$ is proper if $f(W_i) = W_j$. We analyze the dynamical system $f : \sqcup_i U'_i \to \sqcup_i U_i$. Define $f_*(U_i) = U_j$ if $f(U'_i) = U_j$ (equivalently, if $f(W_i) = W_j$).

Assume that a Julia component J_0 satisfies that $f^n(J_0) \subset \bigcup_i U_i$ for infinitely many n. Then either $f^n(J_0) \subset \bigcup_i Q_i$ for infinitely many n or $f^n(J_0) \subset \bigcup_i U'_i$ for all $n \geq n_1$. In the former case $f^n(J_0) \subset Q$ for infinitely many n too since Q is going to be defined as a set containing $\bigcup_i Q_i$. In the latter case the orbit of J_0 lands eventually into a periodic cycle of f_*. We now analyze this second case.

Let $U_0, \ldots, U_{p-1}$ be a periodic cycle of f_*. Set $g = f^p$. Then $\mathcal{J}(g) = \mathcal{J}(f)$ and $U_i = U(W_i, \mathcal{P}(f)) = U(W_i, \mathcal{P}(g))$.

For any Julia component J_0 such that $J_0 \subset U'_0$ and $f^n(J_0) \subset \bigcup_i U'_i$ for all $n \geq 0$, we have $f(J_0) \subset U'_1, \ldots, f^{p-1}(J_0) \subset U'_{p-1}$ and $f^p(J_0) \subset U'_0$ and so on. Therefore $g^n(J_0) = f^{np}(J_0) \subset U'_0 \subset U_0$ for all n.

Set $W = W_0$ and $U = U_0 = U(W, \mathcal{P})$. Then W is a fixed attracting or parabolic Fatou component of g.

We are going to find a compact set Q'_0 which is the union of finitely many disjoint full continua such that, if the g-orbit of a Julia component J_0 passes infinitely many times through U, then it passes infinitely many times through Q'_0 as well.

Denote by X the empty set in case W is an attracting basin, or the set of one single element which is the fixed parabolic point of W, in case W is a parabolic basin. Set $X_n = \bigcup_{0 \leq j \leq n} g^{-j} X$.

One can find a disc $V \subset W$ with Jordan curve boundary such that $\overline{V} \subset W \cup X$ and $g(\overline{V}) \subset V \cup X$. We may choose V such that $(\partial V - X) \cap \mathcal{P} = \varnothing$. For any n, let V_n be the unique component of $g^{-n}(V)$ containing $V_0 = V$. Then $\overline{V}_n \subset V_{n+1} \cup X_n$ for any n. There is an integer N guaranteed by Lemma 7.3, such that every component of $\overline{\mathbf{C}} - V_N$ contains at most one of the finitely many components of $\overline{\mathbf{C}} - W$ which intersect $\mathcal{P}$, and $\mathcal{P} \cap W \subset V_N$.

The set $\overline{\mathbf{C}} - V_N$ is a disjoint union of finitely many full continua with

$$\mathcal{J} \subset \operatorname{int}(\overline{\mathbf{C}} - V_N) \cup X_N.$$

Denote the components of $\overline{\mathbf{C}} - V_N$ by $\overline{D_1}, \ldots, \overline{D_l}, \ldots, \overline{D_n}$ such that, for $j > l$, $\overline{D_j} \cap \mathcal{P} = \varnothing$; and for $1 \leq j \leq l$, $\overline{D_j}$ contains a unique component K_j of $\overline{\mathbf{C}} - W$ such that $K_j \cap \mathcal{P} \neq \varnothing$.

For $1 \leq j \leq l$, set $A_j = \operatorname{int}(\overline{D_j}) - K_j$. Then A_j is an open annulus with possibly finitely many pinched points at points in X_N. Moreover A_j contains no critical value (by the choice of N).

Now we look at the level $N+1$. For $U = U(W, \mathcal{P})$ and $U' = U(W, g^{-1}\mathcal{P})$, the map $g : U' \to U$ is a branched covering (Proposition 7.2). Since $V_N \cup (\bigcup_{j=1}^{l} A_j \cup \bigcup_{j>l} \overline{D_j}) = U$, we conclude that $U' - g^{-1}(\bigcup_{j=1}^{l} \overline{A_j} \cup \bigcup_{j>l} \overline{D_j})$ is connected, and coincides with $g^{-1}V_N \cap U'$. So $g^{-1}V_N$ has a unique component in U', which is V_{N+1}.

Denote the components of $\overline{\mathbf{C}} - V_{N+1}$ by $\overline{D'}_1, \ldots, \overline{D'}_l, \ldots, \overline{D'}_k$ such that $K_j \subset \overline{D'}_j \subset \overline{D}_j$ for $1 \leq j \leq l$, and $\overline{D'}_j \cap \mathcal{P} = \varnothing$ for $j > l$.

We set $g_*(K_i) = K_j$ if $g(\partial K_i) = \partial K_j$, which is well-defined by Proposition 7.2.

Set $A'_j = \operatorname{int}(\overline{D'_j}) - K_j$, $j = 1, \ldots, l$. If $g_*(K_i) = K_j$, then $g(A'_i) = A_j$ and $g : A'_i \to A_j$ is a covering map.

Note that $A_j - A'_j = \operatorname{int}(\overline{D_j}) - \operatorname{int}(\overline{D'_j})$. Moreover $\mathcal{J} \cap (\operatorname{int}(\overline{D_j}) - \overline{D'}_j)$ is contained in $\bigcup_{s>l} \overline{D'}_s$.

We claim then every Julia point x in $\bigcup_{j=0}^{l} A'_j$ must have some iterate in

$$\bigcup_{j=0}^{l} (A_j - A'_j).$$

If not, there is $n_0 > 0$ such that for all $n > n_0$,

$$g^n(x) \in \bigcup \{A'_j \mid K_j \text{ is periodic for } g_*\}.$$

On the other hand, for $K_0, \ldots, K_{q-1}$ a periodic cycle of g_*, and for any $y \in A'_0$, there is a minimal integer $s > 0$ such that $g^s(y) \notin \bigcup_{j=0}^{q-1} A'_j$, for Lemma 7.3 below implies that each K_j is the nested intersection of sets of the form $\cap_n B_n$, where B_n is a component of $\overline{\mathbf{C}} - V_n$, $n \geq 0$.

Therefore if a Julia component J_0 satisfies that $g^n(J_0) \subset \bigcup_{j=0}^{l} A'_j \cup \bigcup_{s>l} \overline{D'_s}$ for infinitely many n, then $g^n(J_0) \subset \bigcup_{s>l} \overline{D'_s}$ for infinitely many n.

On the other hand, for our set $U = U(W, \mathcal{P})$, we have

$$\mathcal{J} \cap U \subset \bigcup_{j=0}^{l} A'_j \cup \bigcup_{s>l} \overline{D'_s} \cup X_{N+1}.$$

Thus saying $g^n(J_0) \subset U$ for infinitely many n is the same as saying

$$g^n(J_0) \subset \bigcup_{j=0}^{l} A'_j \cup \bigcup_{s>l} \overline{D'_s}$$

for infinitely many n.

Return to our original map f now. The components of $Q_0 = U_0 - U'_0$ are disjoint full continua contained in $\bigcup_{j=0}^{l} A'_j \cup \bigcup_{s>l} \overline{D'_s}$. Therefore $Q'_0 = Q_0 \cup \bigcup_{s>l} \overline{D'_s}$ is again a union of finitely many disjoint full continua.

For each periodic cycle of f_*, choose one representative U_i in the cycle, and define Q'_i in the same way, for the other i in the cycle, define $Q'_i = Q_i$. We write $\{0, 1, \ldots, m\} = I' \sqcup I$, where $j \in I'$ if U_j is f_*-periodic and $j \in I$ otherwise.

Set $Q = \bigcup_{i \in I'} Q'_i \cup \bigcup_{i \in I} Q_i$. This is again a union of finitely many disjoint full continua, disjoint from $\mathcal{P}$. Now let J_0 be a Julia component passing infinitely many times through $\bigcup_i U_i$. Then either it passes infinitely many times through $\bigcup_i Q_i \subset Q$, or there is some $i \in I'$, such that J_0 passes infinitely many times through $Q'_i \subset Q$. In both cases J_0 must pass infinitely many times through Q.

Finally we apply Lemma 1.4 to conclude that such Julia components are points.

We mention here two particular cases.

1. The component W is simply connected. In this case $U = W$ and no Julia component passes through U.

2. The set $U(W, \mathcal{P})$ coincides with $\overline{\mathbb{C}}$, that is $\mathcal{P} \subset W$. In this case $l = 0$ and each Julia component is a point. That is $\mathcal{J}$ is totally disconnected. □

A variant of the following lemma can be found in **[St]**, page 63 and 117.

Lemma 7.3. — *Let W be an fixed attracting or parabolic basin of a rational map f. Let V be either a disc neighborhood of the attracting fixed point with Jordan curve boundary or a Fatou petal in W of the parabolic fixed point. Then for V_n the component of $f^{-n}(V)$ containing V, we have $W = \bigcup_n V_n$, and each component of $\overline{\mathbb{C}} - V_n$ is closed disc with possibly finitely many pinching points. Furthermore, given any two components K_1 and K_2 of $\overline{\mathbb{C}} - W$, there is an integer N such that V_N separates K_1 and K_2.*

8. Further results

8.1. Diameter of Julia components. — Let f be a hyperbolic rational map.

Corollary 8.1. — *A Julia component J_0 of $\mathcal{J}$ is not a point if and only if there is an integer N such that $J_n = f^n(J_0) \subset \mathcal{A}^S \sqcup f^{-1}E$ for any $n \geq N$.*

If J_0 is a point, we may regard N as $+\infty$.

One can show that the set $f^{-1}(\mathcal{A}^S \sqcup f^{-1}E) - (\mathcal{A}^S \sqcup f^{-1}E)$ is contained in finitely many discs. Thus the same technique as in the above sections can prove also the following results: to each integer N, there is a positive number $\varepsilon(N)$, with $\varepsilon(N) \to 0$ as $N \to \infty$, such that for J_0 a Julia component, and N the minimal integer such that $f^n(J_0) \subset \mathcal{A}^S \sqcup f^{-1}E$ for $n \geq N$, the spherical diameter of J_0 is at most $\varepsilon(N)$.

8.2. Symbolic dynamics of Julia components in $\mathcal{A}^S$. — Let f be a hyperbolic rational map. The arguments given in §5 actually show much more. The space Σ of admissible sequences $\{a_0 a_1 \dots\}$ equipped with the product topology and the one-sided shift map σ is a subshift of finite type. Let $\mathrm{Cl}(\overline{\mathbf{C}})$ denote the space of all closed subsets of $\overline{\mathbf{C}}$ in the Hausdorff topology. The map f induces a map from this space to itself which we again denote by f. Given $\alpha \in \Sigma$, we have defined in §5 a nested sequence of annuli $T_n(\alpha)$ and a continuum $J'(\alpha) = \cap_n \overline{T_n(\alpha)}$. The set $J'(\alpha)$ may or may not be a Julia component.

Theorem 8.2. — *The map*

$$\Phi : \alpha \mapsto J'(\alpha);\ \Sigma \to \mathrm{Cl}(\overline{\mathbf{C}})$$

defines a uniformly continuous injective map conjugating σ to $f|\Phi(\Sigma)$. Moreover,

(1) *there is a (at most) countable subset Σ_e and a finite subset $\Sigma_{e,0} \subset \Sigma$ such that*
   (a) *$\alpha \in \Sigma_{e,0}$ if and only if $J'(\alpha)$ is a boundary component of A_{a_0} for some $A_{a_0} \in \mathcal{A}^S$,* i.e. *is a boundary component of $T_0(\alpha)$.*
   (b) *$\alpha \in \Sigma_e$ if and only if $J'(\alpha)$ is a boundary component of $T_n(\alpha)$ for some n.*
   (c) *$\sigma(\Sigma_{e,0}) \subset \Sigma_{e,0}$.*
   (d) *$\Sigma_e = \cup_{n\geq 0}\sigma^{-n}(\Sigma_{e,0})$.*

(2) *$\alpha \notin \Sigma_e$ if and only if $J'(\alpha)$ is a Jordan curve Julia component satisfying $f^n(J'(\alpha)) \subset \mathcal{A}^s$ for all $n \geq 0$.*

(3) *Let $J_{e,0} = \Phi(\Sigma_{e,0})$. Then as a subset of $\overline{\mathbf{C}}$,*

$$\bigcup_{\alpha\in\Sigma} \Phi(\alpha) = \{z \mid f^n(z) \in \mathcal{A}^s \sqcup (\cup_{\delta \in J_{e,0}} \delta)\ \forall n \geq 0\}.$$

The elements of the set of continua $J_e := \Phi(\Sigma_e)$ may be thought of as "exposed" boundary components, in the sense that they are boundary components of T_n.

Corollary 8.3. — *The following are equivalent:*

(1) *Σ is uncountable.*
(2) *There is a wandering component of the Julia set which is a Jordan curve.*
(3) *There are uncountably many wandering components of the Julia set which are Jordan curves.*
(4) *There are infinitely many periodic Jordan curve Julia components.*
(5) *There exists a component C' of $\mathcal{A}$, disjoint essential subannuli $A, B \subset C'$, and integers $m, n > 0$ such that $f^m : A \to C', f^n : B \to C'$ are covering maps.*

Proof. — For any subshift of finite type (Σ, σ), the following are equivalent:

- the space of admissible sequences Σ is uncountable;
- there is a wandering sequence;
- there are uncountably many wandering sequences;
- there are infinitely many periodic sequences;
- there are two finite-length sequences $\alpha = (a_0, \ldots, a_m)$, $\beta = (b_0, \ldots, b_n)$ with $c := a_0 = a_m = b_0 = b_n$ and $a_i \neq b_i$ for some $0 \leq i \leq \min m, n$.

By the above Theorem and the preceding facts, the first three conditions are thus equivalent, and (1) implies (4). The set E contains finitely many Julia components, and each periodic Jordan curve Julia component which is not in E must be contained in $\mathcal{A}^S$. Each such component is $J'(\alpha)$ for some periodic α, by the above Theorem, hence Σ has infinitely many periodic sequences and so (4) implies (1). We now show the equivalence of (4) and (5). First, note that (5) is equivalent, by pulling back, to the same condition with C', a component of $\mathcal{A}$, replaced by C, a component of $\mathcal{A}^S$. The equivalence of (4) and (5) then follows immediately from the above Theorem and the preceding paragraph. □

Proof of Theorem 8.2. — The continuity of Φ follows immediately from Lemma 2 of §5. That Φ is a semiconjugacy follows from the fact that $f : T_n(\alpha) \to T_{n-1}(\sigma(\alpha))$ is a covering map, hence $f : \overline{T_n(\alpha)} \to \overline{T_{n-1}(\sigma(\alpha))}$ is surjective, and so

$$f(\Phi(\alpha)) = f(J'(\alpha)) = f(\cap_n \overline{T_n(\alpha)}) = \cap_n \overline{T_{n-1}(\sigma(\alpha))} = J'(\sigma(\alpha)) = \Phi(\sigma(\alpha)).$$

To see that Φ is injective, first note that given a boundary component δ of $A \in \mathcal{A}$, no other component $B \in \mathcal{A}$ has δ as a boundary component. For δ is locally connected (*e.g.* by Corollary 4.2) and hence δ is homeomorphic to S^1 if it is the common boundary of two disjoint open annuli, by Lemma 5.1. But if this occurs then δ is a component of the Julia set separating its parallels, violating the construction of $\mathcal{A}$. More generally, since f sends boundary components of $T_n(\alpha)$ to boundary components of $T_{n-1}(\sigma(\alpha))$, if δ is the common boundary of $T_n(\alpha)$ and $T_n(\beta)$, then $T_n(\alpha) = T_n(\beta)$, *i.e.* if $\alpha = \{a_0 a_1 \ldots\}$, $\beta = \{b_0 b_1 \ldots\}$ then $a_i = b_i$, $0 \leq i \leq n$. If $J'(\alpha) = J'(\beta)$, then either $J'(\alpha) = J'(\beta)$ is a boundary component of $T_n(\alpha)$, some n, or $J'(\alpha) \subset T_n(\alpha)$ for all n. In the former case we have $T_n(\alpha) = T_n(\beta)$ for all n by the above observation while in the latter we must have $T_n(\alpha) = T_n(\beta)$ for all n since if $\alpha \neq \beta$ then for some n, $T_n(\alpha) \cap T_n(\beta) = \varnothing$. Hence $\alpha = \beta$.

(1) Define $\Sigma_{e,0}$ as in 1(a). We first prove that $J_{e,0} = \Phi(\Sigma_{e,0})$ is forward-invariant under f. Let $J_0 \in J_{e,0}$ and let $J_1 = f(J_0)$. Then J_0 is a boundary component of $T_0(\alpha)$ for some α. The proof of Lemma 4 of §5 shows that then J_0 is a boundary component of $T_k(\alpha)$ for all $k \geq 0$, hence J_0 is a boundary component of $T_1(\alpha)$. Since $f : T_1(\alpha) \to T_0(\sigma(\alpha))$ is a covering, $f(J_0) = J_1$ is a boundary component of $T_0(\sigma(\alpha))$ and so $J_1 \in J_{e,0}$. Now 1(a) holds by definition, 1(c) follows from the above result and the fact that Φ is a semiconjugacy. Define Σ_e by 1(d). 1(b) then follows immediately from the fact that Φ is a semiconjugacy and the fact that $f^n : T_n(\alpha) \to T_0(\sigma^n(\alpha))$.

(2) Note that $\alpha \notin \Sigma_e$ if and only if $J'(\alpha) \subset T_n(\alpha)$ for all n, by 1b). The result then follows by Lemma 4 of §5.

(3) First, let $J_0 = J'(\alpha)$. If $\alpha \notin \Sigma_e$, then $J_0 \subset T_n(\alpha)$ for all $n \geq 0$, hence $f^n(J_0) \subset T_0(\sigma^n(\alpha)) \subset \mathcal{A}^S$ for all $n \geq 0$. Otherwise, there is a minimal $N \geq 0$ such that for $0 \leq i < N$, $J_0 \subset T_i(\alpha)$ and for $i = N$, J_0 is a boundary component of $T_i(\alpha)$. Hence for $0 \leq i < N$, $f^i(J_0) \subset T_0(\sigma^i(\alpha)) \subset \mathcal{A}^S$, and for $i = N$, $f^i(J_0)$ is a boundary component of $T_0(\sigma^i(\alpha)) \in J_{e,0}$. Since $J_{e,0}$ is forward-invariant under f the inclusion in this direction is proved.

To prove the other direction, given z an element of the right-hand side we must produce α so that $z \in J'(\alpha)$. Let $z_n = f^n(z)$. If $z_n \in A_{b_n} \in \mathcal{A}^S$ for all $0 \leq n < N$ we set $a_n = b_n$. Hence we may assume $z_n \in \partial\mathcal{A}^S$ for all $n \geq 0$. Choose a component A_{a_0} so that $z_0 \in J_0$, a boundary component of A_{a_0} (there may be more than one such component, since the annuli in $\mathcal{A}^S$ may have closures which intersect). Then $J_0 \in J_{e,0}$. Since $J_{e,0}$ is forward-invariant, $J_n = f^n(J_0)$ is a boundary component of a unique component A_{a_n} of $\mathcal{A}^S$. We now claim that $\alpha = (a_0 a_1 \dots)$ chosen in this fashion is admissible, *i.e.* that $A_{a_i a_{i+1}} \neq \varnothing$ for all i. For on the one hand, $f(A_{a_i}) = A \in \mathcal{A}$, hence $J_{i+1} = f(J_i)$ is a boundary component of A. On the other hand, J_{i+1} is a boundary component of $A_{a_{i+1}} \subset B \in \mathcal{A}$. If $A \neq B$ then we must have that J_{i+1} is the common boundary of A and B, which we have previously noted is impossible. Hence $A = B$ and so $A_{a_i a_{i+1}} \neq \varnothing$. Hence $J_0 = J'(\alpha)$ contains z. □

8.3. Moduli restrictions and description of the shift. — Let f be a nice rational map. Let $\{C_1, \dots, C_k\}$ be the set of annuli in $\mathcal{A}$. Denote by m_j the modulus of C_j. Note that $0 < m_j < \infty$. For each j choose a Jordan curve $\gamma_j \subset C_j$ which is a homotopically non trivial in C_j. Set $\Gamma = \{\gamma_1, \dots, \gamma_k\}$. Let $\Lambda(i,j)$ denote the set of components $C_{i,j,\lambda}$ of $f^{-1}(C_j)$ homotopic to C_i and let $d_{i,j,\lambda}$ be the positive degree of $f : C_{i,j,\lambda} \to C_j$.

Define two linear maps

$$f_\Gamma, f_{\Gamma,\#} : \mathbf{R}^\Gamma \to \mathbf{R}^\Gamma$$

by

$$f_\Gamma(\gamma_j) = \sum_{\gamma_i \in \Gamma} \Big(\sum_{\lambda \in \Lambda(i,j)} \frac{1}{d_{i,j,\lambda}} \Big) \gamma_i$$

and

$$f_{\Gamma,\#}(\gamma_j) = \sum_{\gamma_i \in \Gamma} \Big(\sum_{\lambda \in \Lambda(i,j)} 1 \Big) \gamma_i.$$

If $\Lambda(i,j) = \varnothing$ we take the coefficient of γ_i to be zero. The map f_Γ is the Thurston linear transformation defined by the multicurve Γ. The Grötzsch inequality implies that if $\vec{m} = (m_j) \in \mathbf{R}^\Gamma$ is the vector of moduli m_j, then

$$(f_\Gamma(\vec{m}))_j < m_j, 1 \leq j \leq k.$$

As a consequence, the leading eigenvalue of f_Γ is smaller than 1 (consider the product $vf_\Gamma m$ with v a leading eigenvector of $(f_\Gamma)^t$).

The map $f_{\Gamma,\#}$ is almost the transition matrix for the subshift (Σ,σ). The components of $\mathcal{A}^S$ are in one-to-one correspondence with the collection of elements

$$\{\lambda \mid \lambda \in \Lambda(i,j)\}.$$

If $A_\lambda, A_\mu \in \mathcal{A}^S$ where $\lambda \in \Lambda(i,j) \neq \varnothing$ and $\mu \in \Lambda(k,l) \neq \varnothing$ then

$$A_{\lambda\mu} = \{z \mid z \in A_\lambda,\ f(z) \in A_\mu\} \neq \varnothing \quad \text{if and only if} \quad j = k.$$

An admissible sequence $\alpha \in \Sigma$ is thus given by an infinite sequence λ_n where successive terms satisfy the above condition. Hence by removing any basis elements γ_j for which the jth row of $f_{\Gamma,\#}$ consists only of zeros we obtain a new matrix. Iterating this removal process we obtain a matrix which gives exactly the subshift (Σ,σ).

8.4. Quasicircles and non-quasicircles. — A Jordan curve J in the sphere is a *K-quasicircle* if it is the image of a round circle under a K-quasiconformal map, and it is said to have *K'-bounded turning* if the ratio

$$\operatorname{diam}(L)/d(x,y) \leq K'$$

for all $x, y \in J$, where L is the component of $J - \{x,y\}$ with smallest diameter. Here distance is measured with respect to the spherical metric d. J is said to be a *quasicircle* if it is a K-quasicircle for some K. It is known that J is a quasicircle if and only if it has bounded turning ([**LV**] II, §8).

If a Jordan curve component J of $\mathcal{J}$ is preperiodic and f is hyperbolic, it is a quasicircle by the surgery argument of McMullen [**Mc1**]. However, a wandering Jordan curve component J need not be a quasicircle. We first show

Theorem 8.4. — *Let f be a hyperbolic rational map. If the orbit of α under σ does not accumulate on $\Sigma_{e,0}$ then $\Phi(\alpha)$ is a quasicircle.*

(Here we use the same notation as in §8.2). The proof also shows the following corollary. However, we have no example where the hypotheses are satisfied.

Corollary 8.5. — *If $\Sigma_{e,0}$ is empty (*i.e.* if every boundary component of a component of $\mathcal{A}^S$ maps to a boundary component of $\mathcal{L}$ which is not also a boundary component of any component of $\mathcal{A}^S$), then there exists a K such that for every $\alpha \in \Sigma$, $\Phi(\alpha)$ is a Julia component which is a $K-$quasicircle.*

Proof. — Let $J_0 = \Phi(\alpha)$. If α does not accumulate on $\Sigma_{e,0}$ then there is a positive integer N_0 such that

$$\mathcal{T}_0' = \{T_{N_0}(\sigma^i(\alpha)),\ i \geq 0\}$$

is a disjoint union of open annuli which is compactly contained in $\mathcal{A}^S$, and which contains the entire forward orbit of J_0. Since the orbit of J_0 does not accumulate on a boundary component of $\mathcal{A}^S$, it does not accumulate on boundary components of $\mathcal{T}_0'$. Hence there is an integer $N_1 > N_0$ such that

$$\mathcal{T}_1' = \{T_{N_1}(\sigma^i(\alpha)),\ i \geq 0\}$$

is a disjoint union of open annuli which is compactly contained in $\mathcal{T}_0'$ and which contains the entire forward orbit of J_0. Then $f^{N_1-N_0}(\mathcal{T}_1') = \mathcal{T}_0'$.

Enlarge $\mathcal{T}_0'$ slightly to a disjoint union $\mathcal{T}_0$ of open annuli whose boundaries are real-analytic Jordan curves and which is compactly contained in $\mathcal{A}^S$. Let $\mathcal{T}_1$ be the union of preimages of components of $\mathcal{T}_0$ under $f^{N_1-N_0}$ containing components of $\mathcal{T}_1'$. Then $\mathcal{T}_1$ is compactly contained in $\mathcal{T}_0$, and

$$g = f^{N_1-N_0} : \mathcal{T}_1 \to \mathcal{T}_0$$

is an expanding conformal dynamical system.

On the other hand, we may build another model for this dynamical system of the form

$$h : \mathcal{R}_0 \to \mathcal{R}_1$$

where $\mathcal{R}_i$ consist of round annuli and the map h on each component is $z \mapsto z^d$, some d. We may also find smooth maps $\phi_i : \mathcal{R}_i \to \mathcal{T}_i$ such that $\phi_0 \circ h = g \circ \phi_1$ on $\mathcal{R}_1$ and $\phi_0 \simeq \phi_1$ *rel* $\partial\mathcal{R}_0$. That is, the pair (ϕ_0, ϕ_1) gives a *combinatorial equivalence* between the two dynamical systems. By Theorem A.1 of **[Mc2]**, ϕ_i are isotopic *rel* $\partial\mathcal{R}_0$ to a quasiconformal conjugacy ψ. Since the Julia set of h is the product of a Cantor set with a round circle, J_0 is a component of the Julia set of g and hence J_0 is a quasicircle.

If $\Sigma_{e,0}$ is empty, then $f^{-1}(\mathcal{A}^S) \cap \mathcal{A}^S$ is compactly contained in $\mathcal{A}^S$. We may then apply the argument above with $\mathcal{T}_0' = \mathcal{A}^S$ and $\mathcal{T}_1' = f^{-1}(\mathcal{A}^S) \cap \mathcal{A}^S$ to prove the Corollary. □

Example. — We illustrate this by our example $f = f_1$. In this case $\mathcal{A}^S$ has two components A_0 and A_1. We choose A_0 to be the outermost one, *i.e.* the one with a fixed boundary (the component J^+). There are no components of $\mathcal{J}$ which are points, and indeed $\mathcal{J} = \{z \mid f^n(z) \in \overline{\mathcal{A}^S}\ \forall n\}$. Moreover, the boundary components of components of $\mathcal{A}^S$ are entire Julia components, and $\overline{A}_0 \cap \overline{A}_1 = \varnothing$. It follows that the connected components of $\mathcal{J}$ are precisely the sets of the form $J'(\alpha), \alpha \in \Sigma$. Hence by Theorem 8.2, the dynamics on the space of connected components of $\mathcal{J}$ in the Hausdorff topology is conjugate to the one-sided shift on two symbols 0 and 1. Note that then $\Sigma_{e,0} = \{(000\dots)\}$ and hence that $\Sigma_e = \{(a_0a_1\dots) \mid a_n = 0\ \forall n \geq n_0\}$.

The dynamical system (Σ, σ) is conjugate via ϕ to (C, g), where C is the invariant Cantor set for the interval map $g : [0, 1/3] \sqcup [2/3, 1] \to [0, 1]$ given by $g(c) = 3c$ for $0 \leq c \leq 1/3$ and $g(c) = 3(1-c)$ for $2/3 \leq c \leq 1$. The conjugacy ϕ is defined by $\phi(c) = (a_0a_1a_2\dots)$ where $a_n = 0$ if $g^n(c) \in [0, 1/3]$ and $a_n = 1$ otherwise. Note that the point $0 \in C$ corresponds to $(0000\dots)$ and the point $1 \in C$ to $(1000\dots)$. Order the components of $\mathcal{J}$ so that $J > J'$ if J' separates J and the point at infinity. Then the composition $\Phi \circ \phi$ is order-preserving with respect to the usual ordering on the interval and $>$.

Say $c \in C$ is *exposed* if c is a boundary point of a component of the complement of C and is *buried* otherwise. Similarly, we say a Julia component is *exposed* if it meets the boundary of a Fatou component of f and is *buried* otherwise. Then c is exposed if and only if $\phi(c) \in \Sigma_e$ if and only if $\Phi \circ \phi(c)$ is exposed if and only if $\Phi \circ \phi(c)$ is a covering of J^+ if and only if $\Phi \circ \phi(c)$ is not a Jordan curve.

Finally, we note that Σ_e is dense in Σ.

We now show

Theorem 8.6. — *$\Phi \circ \phi(c)$ is a quasicircle if and only if c does not accumulate at 0 under g,* i.e. *if and only if $\alpha = \phi(c)$ does not contain arbitrarily long strings of consecutive zeros.*

Thus quasi-circle components form a dense subset in the space of Julia components. On the other hand, the non quasi-circle components form a residual set in Baire's category.

Proof. — Let $J_0 = \Phi(c_0)$ and $J_n = f^n(J_0) = \Phi(g^n(c_0))$, $n \geq 0$. By Theorem 8.4 it suffices to show that if zero is a limit point of the orbit of c_0 then J_0 is not a quasicircle. On a compact neighborhood of $\overline{\mathcal{A}^S}$ avoiding the point at infinity the spherical and planar metrics are equivalent, hence J_0 is a quasicircle if and only if it has bounded turning with respect to the planar metric. For simplicity we conjugate the map by $1/z$ so that J^+ is near zero and the J_n separate J^+ from infinity.

By Theorem 8.2, if $g^n(c_0)$ has zero as limit point, then there is a subsequence $J_{n_k} \to J^+$ in the Hausdorff topology. In J^+ we may find a cut point p and small open discs W' and W with $\overline{W} \subset W'$ and such that $W' \cap \mathcal{P} = \varnothing$, $p \in W$, and W contains a connected component L of $J^+ - \{p\}$. Since p is a cut point p is accessible from the component of $\overline{\mathbf{C}} - J^+$ containing infinity (recall we conjugated our map by $1/z$) via two distinct accesses η_x, η_y. Since $J_{n_k} \to J^+$ in the Hausdorff topology and the J_{n_k} are critically separating, for k sufficiently large, there are points $x_k \in \eta_x \cap J_{n_k} \cap W$ and $y_k \in \eta_y \cap J_{n_k} \cap W$ such that $x_k, y_k \to p$ and for which the component L_k of $J_{n_k} - \{x_k, y_k\}$ of smallest diameter is also contained in W and has diameter bounded below by $D = \operatorname{diam}(L)$.

Since $W' \cap \mathcal{P} = \varnothing$, there is a univalent branch h_k of $(f^{n_k})^{-1}$ on W' sending L_k to a subarc of J_0. Let $x_{0k} = h_k(x_k)$, $y_{0k} = h_k(y_k)$, and $L_{0k} = h_k(L_k)$. Then the Koebe distortion theorem implies that since W is compactly contained in W', there is a constant $C \geq 1$ independent of k such that h_k distorts ratios of planar distances between points of W by at most a factor of C. Hence for all k

$$C^{-1}\frac{\operatorname{diam}(L_k)}{|x_k - y_k|} \leq \frac{\operatorname{diam}(L_{0k})}{|x_{0k} - y_{0k}|} \leq C\frac{\operatorname{diam}(L_k)}{|x_k - y_k|}.$$

But $\mathrm{diam}(L_k)$ is bounded from below by D while $|x_k - y_k| \to 0$. Hence J_0 has unbounded turning and so is not a quasicircle. □

8.5. Constructing other examples. — The map f_1 was initially found by using quasiconformal surgery to glue $z^2 - 1$ with $1/z^3$ in the appropriate fashion. This construction and its generalizations are the subject of work in progress. For example, one may apply the same construction by replacing $z^2 - 1$ with any polynomial p whose filled Julia set has interior, and glue it to $1/z^3$.

Question. — If p is a quadratic polynomial with non-locally connected Julia set and a Siegel disc, can one obtain, by this gluing, a map with wandering Julia components which are critically separating but not Jordan curves?

8.6. How to find the set E. — There are finitely many Fatou components $W_0, \ldots, W_m$ either intersecting $\mathcal{P}$ or separating $\mathcal{P}$ into at least three parts (Lemma 2.2). For each W_i take the finitely many boundary components which are critically separating. Saturate them into Julia components. Call the union of them E'. It is finite and forward invariant. Fill in the disc components of $\overline{\mathbf{C}} - E'$ disjoint from $\mathcal{P}$. We get a fattened E'' of E'. We distinguish three types of components of $\overline{\mathbf{C}} - E''$: type I, components containing exactly one W_i (these are also the components intersecting $\bigcup W_i$); type II, annulus components disjoint from $\bigcup W_i$; and type III, non-annulus components disjoint from $\bigcup W_i$. In each of type III component, there are finitely many Julia components separating $\mathcal{P}$ into at least three parts. Adding them to E', we get E. Note that the set $\mathcal{L}$ is precisely the union of components of type I (see also Lemma 3.10).

Let W be a fixed attracting basin for a rational map f. How do we know that only finitely many components of $\overline{\mathbf{C}} - W$ intersect $\mathcal{P}$? And if so how can we find V_N in Lemma 7.3 and in the proof of Theorem 7.1? Here is a constructive answer. Take V_0 a open disc with Jordan curve boundary in W such that $f(\overline{V_0}) \subset V_0$. Let V_n be the component of $f^{-n}(V_0)$ containing V_0. Then only finitely many components of $\overline{\mathbf{C}} - W$ intersect $\mathcal{P}$ if and only if there is a minimal integer N such that each boundary curve of V_{N+1} is either not critically separating, or is parallel to a boundary curve of V_N. And for this V_N, each component of $\overline{\mathbf{C}} - V_N$ contains at most one component of $\overline{\mathbf{C}} - W$ intersecting $\mathcal{P}$.

9. Generalizations

9.1. Geometrically finite maps. — Our techniques allow a generalization of Theorem 1.2 to the case of geometrically finite maps f. In **[TY]** (§1, Step 4 and Prop. 1.3)) (*cf.* also **[DH2]**) a Riemannian metric is constructed for which f is uniformly expanding on a neighborhood of its Julia set. The arguments given above for Propositions [Case 2] and [Case 3] then apply in this more general setting. Proposition

[Case 4] is proved for nice maps, therefore applies automatically to geometrically finite maps.

We may also recover a variant of Theorem 8.2 for geometrically finite maps as well, though we may lose the injectivity of Φ–it may be possible to construct a map f with a periodic Jordan curve Julia component J_0 which intersects $\mathcal{P}$ and which is the common boundary of two components of $\mathcal{A}^S$. Then J_0 may be $J'(\alpha)$ for two distinct admissible sequences α. At worst, however, Φ is two-to-one. We define $\Sigma_{e,0}$ and Σ_e the same as for hyperbolic maps.

Theorem 8.4 holds for geometrically finite maps as well, and we may recover a result of Cui et. al. ([**CJS**], Prop. 6.2):

Proposition 9.1. — *If f is geometrically finite, then there are at most finitely many periodic Jordan curve Julia components J_0 which are not quasicircles.*

Proof. — Let $J_n = f^n(J_0)$. Then either $J_n \subset \mathcal{A}^S$ for all n, or J_n is a boundary component of a component of $\mathcal{A}^S$ for all n, since J_0 is periodic. The latter set of such J_0 is finite, while in the former case $J_0 = J'(\alpha)$ for some unique $\alpha \in \Sigma$. The sequence α cannot accumulate on $\Sigma_{e,0}$ under σ since otherwise J_0 accumulates on a boundary component of $\mathcal{A}^S$, by Theorem 8.2. Hence J_0 is a quasicircle by Theorem 8.4. □

9.2. A further result for nice maps

Theorem 9.2. — *If f is nice and every component of $\mathcal{A}^S$ is also a component of $\mathcal{A}$, then every Julia component J_0 not eventually landing on a component of E is a point.*

Proof. — The hypotheses imply that every component of $\mathcal{A}^S$ is a Herman ring or a preimage of a Herman ring, therefore contains no Julia components. So Case 2 of Lemma 2.4 does not occur.

By Proposition [Case 4], if J_0 is a Julia component such that $J_n = f^n(J_0) \subset \mathcal{L}$ for infinitely many n, then J_0 is a point.

Assume now J_0 is a Julia component that is not in Cases 1, 2 and 4 of Lemma 2.4. Replacing J_0 by a forward iterate of it if necessary, we may assume that

$$J_n \cap (E \cup \mathcal{A}^S \cup \mathcal{L}) = \varnothing \quad \text{for all } n \geq 0.$$

Furthermore, by Lemma 2.4, $J_n \subset \mathcal{A}^O \cup \mathcal{D}'$ for infinitely many n. Therefore

$$J_n \subset \hat{Q} = (\mathcal{A}^O - \mathcal{A}^O\mathcal{A}^S \cup \mathcal{A}^O\mathcal{A}\mathcal{L}) \cup (\mathcal{D}' - \mathcal{D}'\mathcal{A}^S \cup \mathcal{D}'\mathcal{L} \cup \mathcal{D}'\mathcal{A}\mathcal{L})$$

for infinitely many n, where the notation ABC means

$$\{z \mid z \in A,\ f(z) \in B, f^2(z) \in C\}.$$

We claim that $\hat{Q}$ is contained in the disjoint union of finitely many full continua disjoint from $\mathcal{P}$. We can then apply Lemma 1.4 to conclude that J_0 is a point.

Let A' be a component of $\mathcal{A}^O$. Either it coincides with a component of $\mathcal{A}^O\mathcal{A}^S$, or $A = f(A')$ is a component of $\mathcal{A} - \mathcal{A}^S$. Assume we are in the latter case. By

Lemma 3.8, for $\delta^{\pm}$ the boundary components of A, there are components V^+, V^- of $\mathcal{AL}$ (which may coincide) such that $\delta^+ \subset \partial V^+$ and $\delta^- \subset \partial V^-$. Thus $\overline{A - (V^+ \cup V^-)}$ is disjoint from ∂A. Hence $\overline{A' - A'\mathcal{AL}}$ is contained in A'. Since the inclusion map from A' into $\overline{\mathbf{C}} - \mathcal{P}$ is homotopic to a constant map, $\overline{A' - A'\mathcal{AL}}$ is contained in a full continuum disjoint from $\mathcal{P}$.

The case of a component of $\mathcal{D}'$ is similar, using Lemma 3.7. □

Appendix A
Technical results about plane topology

In this appendix we collect technical results used in the course of our proofs.

The following lemma was used in the proof of Lemma 3.6.

Lemma A.1. — *Let P^1 and P^2 be two disjoint non empty closed sets of S^2. Set*

$$\mathcal{W} = \{U \mid U \textit{ is a complement component of some component of } J,\ P^1 \subset U \textit{ and } P^2 \cap U = \varnothing\}.$$

Then $\mathcal{W}$ is either empty or totally ordered with respect to inclusion, and $W = \bigcup_{U \in \mathcal{W}} U$ as the unique maximal element, and $\partial \mathcal{W}$ is connected.

Proof. — That $\mathcal{W}$ is totally ordered with respect to inclusion if it is nonempty follows from Lemma 3.1. Since each $U \in \mathcal{W}$ is a disc and the U's are nested, we have that W is a disc. Hence ∂W is connected.

We now show that W is also an element of $\mathcal{W}$. Let $x \in \partial W$. For any integer n, there is a point x' in W such that the distance between x' and x is less than $1/n$, and there is $U \in \mathcal{W}$ such that $x' \in U$. The segment $[x, x']$ intersects ∂U, which is a subset of J. We conclude then either $x \in \partial U$ or there are points of J arbitrarily close to x. In both cases $x \in J$.

Since ∂W is a connected subset of J, it is contained in some component S of J.

Now W must be a component of $\overline{\mathbf{C}} - S$. This is because $S \cap U = \varnothing$ for any $U \in \mathcal{W}$, therefore $S \cap W = \varnothing$.

Moreover $W \supset P^1$, but $W \cap P^2 = \varnothing$. So W is indeed the maximal element of $\mathcal{W}$. □

The remaining results are essentially ingredients in the proof of Theorem 7.1. The logical dependencies are:

Lemma A.2 $\to$ Corollary A.3 $\to$ Corollary A.4 $\to$ Corollary A.5 $\to$ Lemma 3.10 $\to$ Theorem 7.1.

We will need the following fact from general topology.

Let X be a compact Hausdorff space. Then every component Y of X coincides with the intersection of open and closed subsets of X containing Y. Moreover, for any open neighborhood U of Y, there is a closed and open subset V of X such that

$Y \subset V \subset U$. The component Y is always closed. In case Y is not open in X, there is a sequence of closed and open sets V_n such that $Y \subsetneqq V_n \subsetneqq V_{n-1}$, and $\bigcap_n V_n = Y$.

If X is the Julia set of some map, and is disconnected, no component of X is open in X (see **[Be]**).

In the next three results, let $\mathcal{J} \sqcup \mathcal{F}$ be a decomposition of S^2 with $\mathcal{J}$ closed and disconnected, let d denote spherical distance and d_H the Hausdorff distance on closed sets with respect to d.

The following Lemma is known as Zoretti's Theorem (**[Wh]**, p. 109).

Lemma A.2. — *Given J' and J'' two distinct components of $\mathcal{J}$, and $\varepsilon > 0$, there is a Jordan curve γ in $\mathcal{F}$ separating J' and J'' such that $\sup_{x\in\gamma} d(x, J') < \varepsilon$.*

The next two results are easy consequences of Lemma A.2.

Corollary A.3

(1) *Let U be a connected component of $S^2 - J'$ for some connected component of $\mathcal{J}$. Then there is a sequence of closed discs D_n bounded by Jordan curves γ_n such that $\gamma_n \subset U \cap \mathcal{F}$, $D_n \subset \operatorname{int}(D_{n-1})$, $\partial U \subset \operatorname{int}(D_n)$, and $D_H(\partial U, \gamma_n) \to 0$ (where D_H denotes the Hausdorff distance of compact sets).*

(2) *Given W a component of $\mathcal{F}$ and K a component of $S^2 - W$, there is a sequence of closed disks $\overline{D}_n$ such that $\overline{D}_n \subset \operatorname{int}(D_{n-1})$, $\partial D_n \subset W$ and $\bigcap_n D_n = K$.*

Proof. — Let $\{K_m\}_{m=0}^{\infty}$ be a sequence of closed discs such that $K_{m-1} \subset \operatorname{int}(K_m)$, $\cup_m K_m = U$, and $d_H(\partial U, \partial K_m) \to 0$. Let $\mathcal{J}_0 = \mathcal{J} \cup K_0, F_0 = S^2 - \mathcal{J}_0$. Then $\mathcal{J}_0$ is closed and disconnected and Lemma A.2 implies that there is a Jordan curve $\gamma_0 \subset F_0 \subset F$ separating ∂U from ∂K_0. Let D_0 be the disc bounded by γ_0 and containing ∂U. Inductively define γ_n as follows. There is an $m(n)$ such that $\gamma_{n-1} \subset \operatorname{int}(K_{m(n)})$. Let $\mathcal{J}_n = \mathcal{J} \cup K_{m(n)}, F_n = S^2 - \mathcal{J}_n$. Lemma A.2 implies that there is $\gamma_n \subset F_n \subset F$ separating ∂U from $K_{m(n)}$ and bounding a closed disc D_n containing ∂U. This shows the first part; the second is similar. □

Corollary A.4. — *The following conditions are equivalent:*

(1) *there is a component W of $\mathcal{F}$ contained in U such that $\partial W \cap \partial U \neq \varnothing$;*

(2) *there is a component W of $\mathcal{F}$ contained in U such that ∂U is a component of ∂W;*

(3) *there is an n, such that no component of $\mathcal{J}$ separates ∂U and γ_n, where γ_n is a sequence of Jordan curves as in Corollary A.3, Part 1.*

Proof. — That $2 \Rightarrow 1$ is obvious; $2 \Rightarrow 3$ follows directly from Part 2 of Corollary A.3 and the hypothesis that ∂U is a component of ∂W. To see $3 \Rightarrow 2$, note that $\mathcal{F} \cap (U - D_n)$ must be connected, where the D_n are as in Part 1 above. Let $X = \partial W$. Then ∂U is contained in a nested intersection of sets which are open and closed in X, hence ∂U is a connected component of ∂W. Finally, that $1 \Rightarrow 2$ follows easily from the observation that if C_0 is the connected component of ∂W which intersects ∂U, then C_0 is contained in some component J_0 of $\mathcal{J}$, hence C_0 cannot separate points of U. □

Corollary A.5. — *Let E be a nonempty set of finitely many disjoint components of $\mathcal{J}$ and let U be a connected component of $S^2 - E$.*

(1) *If ∂U is connected, let W_0 be a component of $\mathcal{F}$ contained in U, and assume no component of $\mathcal{J}$ separates W_0 from ∂U.*

 Then ∂U is a connected component of ∂W_0, and W_0 is the unique component of $\mathcal{F}$ contained in U whose boundary intersects ∂U.

(2) *If ∂U is not connected, assume no component of $\mathcal{J}$ separates components of E.*

 Then there exists a component W_0 of $\mathcal{F}$ contained in U such that every connected component of ∂U is a connected component of ∂W_0, and W_0 is the unique component of $\mathcal{F}$ contained in U whose boundary intersects each component of ∂U.

Proof. — Assume first that ∂U is connected. By Corollary A.3, Part 1, there is a sequence $\gamma_n \to \partial U$ of Jordan curves in $\mathcal{F} \cap U$; we may take these curves to separate some point w_0 of W_0 from ∂U. By hypothesis and Corollary A.4, Part 3, there exists a component W of $\mathcal{F}$ contained in U such that ∂U is a connected component of ∂W. If $W \neq W_0$, then since $W_0, W \subset U$ we must have that W and W_0 are separated by a component J' of $\mathcal{J}$ which is contained in U. Then $W \cup \partial U$ is a connected set which is separated from W_0 by J', hence ∂U is separated from W_0 by J', a contradiction. The uniqueness assertion follows from the equivalence of the first two parts in the previous lemma, and the proof of the second case is similar. □

Appendix B
Proof of Proposition 7.2

A subtlety to prove Proposition 7.2 is that the map $f : \partial W' \to \partial W$ need not be open in the subspace topology. We first establish

Claim. — *Let f be a rational map and let W', W be two Fatou components of f such that $f(W') = W$. Let K (resp. K') be a component of $\overline{\mathbf{C}} - W$ (resp. $\overline{\mathbf{C}} - W'$) such that $f(\partial K') \cap \partial K \neq \varnothing$. Then*

(1) *$f(\partial K') = \partial K$.*
(2) *$f(K') \supset K$.*
(3) *$K \cap P = \varnothing$ if and only if $K' \cap f^{-1}P = \varnothing$. In this case $f(K') = K$.*
(4) *For any set $S \supset P$, $K' \cap f^{-1}S \neq \varnothing$ if and only if $K \cap S \neq \varnothing$.*

Proof of Proposition 7.2

(1) By (1) above, if K' is a connected component of $\overline{\mathbf{C}} - W'$, then $f(\partial K') = \partial K$ where K is a connected component of $\overline{\mathbf{C}} - W$. By (1) and (4), there are finitely many such K' for which $K' \cap f^{-1}(\mathcal{P}) \neq \varnothing$.

(2) It suffices to show $f(U') \subset U$ and $f(\partial U') \subset \partial U$. Since U' is the union of W' with components of $\overline{\mathbf{C}} - W'$ disjoint from $f^{-1}\mathcal{P}$, this follows from (1) and (3).

(3) Again this follows directly from the finiteness of the number of boundary components and (1). □

Proof of Claim

(1) Denote by J the Julia component containing ∂K.

Denote by V the component of $\overline{\mathbf{C}} - J$ containing W and by V' the component of $\overline{\mathbf{C}} - f^{-1}J$ containing W'. Then $f : V' \to V$ is proper and f maps each boundary component of V' onto the boundary of V, which is $\partial K'$ (see Lemma 3.3).

Clearly $f(\partial K') \subset \partial K$. So $\partial K' \cap V' = \varnothing$. Since $W' \subset V'$ and $\partial K' \subset \partial W'$, we have $\partial K' \subset \partial V'$. So $\partial K'$ is contained in a component S' of $\partial V'$. But S' is in the Julia set, and the Julia component of $\partial K'$ is contained in K', so $S' \subset K'$. Since $\partial K'$ separates W' from $\mathrm{int}(K')$, no point of $\mathrm{int}(K')$ can be in V'. Therefore $S' \subset \partial K'$. Thus $S' = \partial K'$ and $f(\partial K') = \partial K$.

(2) A rational map f maps connected components of preimages of a compact connected set surjectively onto its image (see **[Be]**, Ch. 5). Let L' be the connected component of the preimage of K intersecting $\partial K'$. Then $f(L') = K$ and $\partial K' \subset L' \subset K'$. Hence $f(K') \supset K$.

(3) Assume at first that $K \cap P = \varnothing$. Since K is full the component L' in point (2) is also full. A simple topological argument then shows $K' = L'$. So $f(K') = K$ and $K' \cap f^{-1}P = \varnothing$.

Assume now that $K' \cap f^{-1}P = \varnothing$. If $K \cap P \neq \varnothing$, since $f(K') \supset K$ we would have $K' \cap f^{-1}P \neq \varnothing$ which is impossible. So $K \cap P = \varnothing$, $K' = L'$ and $f(K') = K$.

(4) Assume at first that $K \cap P = \varnothing$. Since $f(K') = K$, we get 4). Assume now $K \cap P \neq \varnothing$. Then $K \cap S \neq \varnothing$ and $K' \cap f^{-1}S \neq \varnothing$ (for otherwise $K' \cap f^{-1}P = \varnothing$ and $K \cap P = \varnothing$). □

References

[AM] J. Anderson and B. Maskit, On the local connectivity of limit sets of Kleinian groups, *Complex Variables*, **31** (1996), pp. 177-183.

[Be] A. Beardon, *Iteration of Rational Functions*, Springer Verlag Graduate Texts in Mathematics 132, 1991.

[Bi] B. Bielefeld (ed). Conformal dynamics problem list, IMS SUNY/SB preprint 1990/1.

[Br] H. Brolin, Invariant sets under iteration of rational functions, *Arkiv för Matematic* **6**(6) (1965), pp. 103-144.

[CJS] G. Cui, Y. Jiang, and D. Sullivan, Dynamics of geometrically finite rational maps, *Manuscript 1996.*

[DH1] A. Douady and J.H. Hubbard, On the dynamics of polynomial-like mappings, *Ann. Sci. Éc. Norm. Sup.* **18** (1985), pp. 287-344.

[DH2] A. Douady and J.H. Hubbard, *Étude dynamique des polynômes complexes, II*, Pub. Math. d'Orsay, 1984.

[LV] O. Lehto and K.J. Virtanen, *Quasiconformal mappings in the plane*, Springer, Berlin-Heidelberg-New York, 1973.

[Mc1] C. McMullen, Automorphisms of rational maps, In *Holomorphic Functions and Moduli I*, pp. 31-60. Springer Verlag, 1988.

[Mc2] C. McMullen, Self-similarity of Siegel disks and Hausdorff dimension of Julia sets, *Preprint 1995.*

[Mc3] C. McMullen, The classification of conformal dynamical systems, *Current Developments in Mathematics 1995*, International Press, 1995, pp. 323-360.

[Mi] J. Milnor, Dynamics in One Complex Variable: Introductory Lectures, *preprint 1990/5* of SUNY at Stony Brook.

[Ne] M.H.A. Newman, *Elements of the topology of planesets of points*, Cambridge University Press, 1961.

[St] N. Steinmetz, *Rational Iteration: Complex Analytic Dynamical Systems*, De Gruyter Studies in Mathematics 16, 1993.

[TY] Tan Lei and Yin Yongcheng, Local connectivity of the Julia set for geometrically finite rational maps, Science in China, Series A (1) **39** (1996), pp. 39-47.

[Wh] G. T. Whyburn, *Analytic Topology*, AMS Colloquium Publications vol. 28 (1942).

K. PILGRIM, Department of Mathematics and Statistics, University of Missouri, Rolla Building, 1870 Minor Circle, Rolla, MO 65409-0020, U.S.A. • *E-mail :* `pilgrim@umr.edu`
Url : `http://www.umr.edu/~mathstat`

TAN LEI, Labo. de mathématiques, UPRESA CNRS 8088, Université de Cergy-Pontoise, 2, avenue Adolphe Chauvin, 95302 Cergy-Pontoise cedex, France • *E-mail :* `tanlei@math.u-cergy.fr`
Url : `http://www.u-cergy.fr/Recherche/Laboratoires/07/index.html`

FELIKS PRZYTYCKI

Hölder implies Collet-Eckmann

Astérisque, tome 261 (2000), p. 385-403
<http://www.numdam.org/item?id=AST_2000__261__385_0>

Astérisque
261, 2000, p. 385–403

HÖLDER IMPLIES COLLET-ECKMANN

by

Feliks Przytycki

Abstract. — We prove that for every polynomial f if its basin of attraction to ∞ is Hölder and Julia set contains only one critical point c then f is Collet-Eckmann, namely there exists $\lambda > 1$, $C > 0$ such that, for every $n \geq 0$, $|(f^n)'(f(c))| \geq C\lambda^n$. We introduce also *topological Collet-Eckmann rational maps* and *repellers*.

0. Introduction

J. Graczyk and S. Smirnov proved in **[GS]** that if a rational map is Collet-Eckmann (abbr. CE), then every component of the complement of Julia set J is Hölder. Another proof was provided later in **[PR1]**. The question whether a converse fact holds remained unanswered. Moreover it has been proved in **[PR2]** (using an example from **[CJY]**) that if there are at least two critical points in J, then the converse may occur false, even for polynomials. Namely if the forward trajectory of a critical point c at some times approaches very closely another critical point, but all critical points in J are nonreccurrent, then A_∞ the basin of infinity is John even, but $|(f^n)'(f(c))|$ does not grow exponentially fast.

Here (in Sec.3) we prove that A_∞ Hölder implies CE for polynomials if there is only one critical point in J. In fact we prove this in a more general setting of rational functions. We prove this by using Graczyk and Smirnov's "reversed telescope" idea.

In Section 4 we introduce for rational maps the property *topological Collet-Eckmann* (abbr. TCE). This property means roughly a possibility of going from many small scales around each point to large scale round discs with uniformly bounded criticality

1991 ***Mathematics Subject Classification.*** — 58F23.
Key words and phrases. — Collet-Eckmann holomorphic maps, repellers, Hölder basin of attraction, non-uniformly hyperbolic, telescope, rational maps, iteration.

Written on the author's stay at the Institute of Advanced Studies of the Hebrew University of Jerusalem in Fall 1996 and Summer 1997. Supported also by Polish KBN Grant 2P301 01307.

under the action of the iterates of f. This property is topological (*i.e.* it is preserved under topological conjugacies) and we prove in Section 4 that it implies CE, provided there is only one critical point in J (or more than one, but none in the ω-limit set of the others). Since by **[PR1]** CE implies TCE we obtain a new elementary proof that CE is a topological property. The first proof was provided in **[PR2]**: For f being CE, and g topologically conjugate to it, it was proved that the conjugacy can be improved to a quasiconformal one on a neighbourhood of $J(f)$. This implied CE for g, by a method not much different from presented here (but simpler technically).

In the unimodal maps of the interval case the fact CE is a topological property was proved in **[NP]** via the same TCE property called there *finite criticality.* The intermediate property used there was *uniform hyperbolicity on periodic orbits* (abbr. UHPer). Here this idea also appears implicitly, though we cannot prove UHPer implies CE (the fact proved for unimodal maps of interval with negative Schwarzian derivative by T. Nowicki and D. Sands in **[NS]**.)

Finally, in Section 5, we introduce and study *holomorphic TCE invariant sets* in particular *repellers* and prove that if a repeller is the boundary of an open connected domain in $\overline{\mathbb{C}}$, then it is TCE iff the domain is Hölder. In consequence, for each domain with repelling boundary, to be Hölder is a topological property. We prove also the analogous rigidity result for Hölder immediate basins of attraction to attracting fixed points.

1. Preliminaries on Hölder basins

Definition 1.1. — Let $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ be a rational map of the Riemann sphere. We call an f-critical point c (*i.e.* such that $f'(c) = 0$) *exposed* if its forward f-trajectory does not meet other critical points.

The map f is called Collet-Eckmann if its every exposed critical point c that belongs to the Julia set $J = J(f)$, or its forward orbit converges to J, satisfies the following Collet-Eckmann condition:

There exists $\lambda > 1$ such that for every $n \geq 0$

$$(CE) \qquad |(f^n)'(c_1)| \geq \mathrm{Const}\, \lambda^n.$$

Notation. — By Const we denote various positive constants which can change from one formula to another. We use the notation $x_n = f^n(x)$.

The definition of holomorphic Collet-Eckmann map was introduced in **[P2]** with (CE) assumed only for critical points in J. This allowed parabolic periodic points. Here we modify the definition, in accordance with **[GS**, Def 1.2**]**.

One calls a simply-connected open hyperbolic domain A *Hölder* if there exists $\alpha > 0$ such that any Riemann mapping from the unit disc D onto A is Hölder continuous. This can be generalized to non-simply connected domains, see **[Po]** or **[GS**, Def 5.1**]**.

We shall not rewrite here this definition in absence of dynamics because we do not need this. However if A is an immediate basin of attraction to a sink, for a rational mapping f, $f(A) = A$, Graczyk and Smirnov provided an equivalent definition [**GS**, Def.1.4, Sec.5 Prop.3] which will be of use for us. Denote $\mathrm{Crit}^+ := \bigcup_{j=1}^{\infty} f^j(\mathrm{Crit})$, where Crit= Crit($f$) means the set of all critical points for f.

Definition 1.2. — We call A Hölder if there exists $\lambda_{\mathrm{Ho}} > 1$ such that for every $z \in A \setminus \mathrm{cl}\,\mathrm{Crit}^+$ there exists $C_1 > 0$ such that for every $y \in f^{-n}(\{z\}) \cap A$

$$|(f^n)'(y)| \geq C_1 \lambda_{\mathrm{Ho}}^n. \tag{1.1}$$

We extend this definition to periodic A, $f^k(A) = A$, by replacing f by f^k above. This replacement allows in proofs to assume $f(A) = A$.

We need also the following

***Notation (cf.* [PUZ]).** — Suppose $f(A) = A$. Let $z^1, \dots, z^d$ be all the pre-images of z in A. Consider smooth curves $\gamma^j : [0,1] \to A \setminus \mathrm{cl}\,\mathrm{Crit}^+$, $j = 1, \dots, d$, joining z to z^j respectively (*i.e.* $\gamma^j(0) = z$, $\gamma^j(1) = z^j$).

Let $\Sigma^d := \{1, \dots, d\}^{\mathbb{Z}^+}$ denote the one-sided shift space and σ the shift to the left, *i.e.* $\sigma((\alpha_n)) = (\alpha_{n+1})$. For every sequence $\alpha = (\alpha_n)_{n=0}^{\infty} \in \Sigma^d$ we define $\gamma_0(\alpha) := \gamma^{\alpha_0}$. Suppose that for some $n \geq 0$, for every $0 \leq m \leq n$, and all $\alpha \in \Sigma^d$, the curves $\gamma_m(\alpha)$ are already defined. Write $z_n(\alpha) := \gamma_n(\alpha)(1)$.

For each $\alpha \in \Sigma^d$ define the curve $\gamma_{n+1}(\alpha)$ as the lift (image) by $f^{-(n+1)}$ of $\gamma^{\alpha_{n+1}}$ starting at $z_n(\alpha)$.

The graph $\mathcal{T} = \mathcal{T}(z, \gamma^1, \dots, \gamma^d)$ with the vertices z and $z_n(\alpha)$ and edges $\gamma_n(\alpha)$ is called a *geometric coding tree* with the root at z. For every $\alpha \in \Sigma^d$ the subgraph composed of $z, z_n(\alpha)$ and $\gamma_n(\alpha)$ for all $n \geq 0$ is called a *geometric branch* and denoted by $b(\alpha)$. Denote by $b_n(\alpha)$ for $n \geq 0$ the subgraph composed of $z_j(\alpha)$ and $\gamma_{j+1}(\alpha)$ for all $j \geq n$.

The branch $b(\alpha)$ is called *convergent* to $x \in \partial A$ if $z_n(\alpha) \to x$.

For an arbitrary basin of attraction A we define the *coding map* $z_\infty : \mathcal{D}(z_\infty) \to \mathrm{cl}\,U$ by $z_\infty(\alpha) := \lim_{n\to\infty} z_n(\alpha)$ on the domain $\mathcal{D} = \mathcal{D}(z_\infty)$ of all such α's for which $b(\alpha)$ is convergent. By Lemma 1.3 below, for A Hölder, $\mathcal{D} = \Sigma^d$ and z_∞ is Hölder.

Finally let $U^1, \dots, U^d$ be open topological discs with closures in $A \setminus \mathrm{cl}\,\mathrm{Crit}^+$, containing $\gamma^1, \dots, \gamma^d$ respectively. For each α and $n \geq 0$ denote by $U_n(\alpha)$ the component of $f^{-n}(U^{\alpha_n})$ containing $\gamma_n(\alpha)$.

In the subsequent Lemmas A is a Hölder immediate basin of attraction to a periodic sink for a rational function f.

Lemma 1.3. — *There exists $C_2 > 0$ such that for every $\alpha \in \Sigma^d$ and every positive integer m*

$$\mathrm{diam}\, U_m(\alpha) \leq C_2 \lambda_{\mathrm{Ho}}^{-m} \tag{1.2}$$

and

$$U_m(\alpha) \subset B(z_\infty(\alpha), C_2\lambda_{\mathrm{Ho}}^{-m}/(1-\lambda_{\mathrm{Ho}}^{-1})). \tag{1.3}$$

Proof. — This follows from (1.1) and uniformly bounded distortion for all the branches of f^{-n}, $n \geq m$ on U^j involved. □

Lemma 1.4. — *For every $x \in \partial A$ there exists $\alpha \in \Sigma^d$ such that $b(\alpha)$ is convergent to x.*

Proof. — Notice that $x = \lim z_{n_k}(\widehat{\alpha}^k)$ for a sequence $\widehat{\alpha}^k \in \Sigma^d$ and a sequence of integers n_k, see [**PZ**, the proof of (9)]. Now any α a limit of a convergent subsequence of $\widehat{\alpha}^k$ satisfies the assertion of the Lemma. The convergence of $b(\alpha)$ is even exponential. This follows from Lemma 1.3 □

Lemma 1.5. — *Let A be a Hölder immediate basin of attraction to a sink for a rational map f. Then for every $\lambda : 1 < \lambda < \lambda_{\mathrm{Ho}}$ there exist $\delta > 0$ and $n_0 > 0$ such that for every $n \geq n_0$ and every $x \in \partial A$, if for every $j = 0, \dots, n-1$*

$$\operatorname{dist}(x_j, \mathrm{Crit}) > \exp -\delta n, \tag{1.4}$$

then $|(f^n)'(x)| > \lambda^n$.

Proof. — Consider $\alpha \in \Sigma^d$ such that $b(\alpha)$ converges to x_0. Then for

$$s = [C_3 + n\delta/(\log \lambda_{\mathrm{Ho}})] + 1$$

(the square brackets stand for the integer part), where

$$C_3 = \frac{(\log C_2/\varepsilon(1-\lambda_{\mathrm{Ho}}))}{\log \lambda_{\mathrm{Ho}}},$$

for an arbitrary $\varepsilon : 0 < \varepsilon < 1$, one obtains by (1.3)

$$z_s(\sigma^n(\alpha)) \subset B(x_n, C_2\lambda_{\mathrm{Ho}}^{-s}/(1-\lambda_{\mathrm{Ho}}^{-1}) = B(x_n, \varepsilon \exp -\delta n).$$

Moreover for every $0 \leq j \leq n$

$$z_{s+j}(\sigma^{n-j}(\alpha)) \in B(x_{n-j}, \varepsilon \exp -\delta n). \tag{1.5}$$

For $y := z_{s+n}(\alpha)$ we have

$$|(f^n)'(y)| = |(f^{n+s})'(y)| \cdot |(f^s)'(f^n(y))|^{-1} \geq C_1\lambda_{\mathrm{Ho}}^{n+s}L^{-s}$$

for $L := \sup|f'|$.

Using the definition of s we see that for δ small enough and n large, the latter expression is larger than $\tilde{\lambda}^n$ for an arbitrary $\tilde{\lambda} : 1 < \tilde{\lambda} < \lambda_{\mathrm{Ho}}$, so

$$|(f^n)'(y)| > \tilde{\lambda}^n. \tag{1.6}$$

For ε small enough, in view of (1.4) and (1.5), we can replace y by x in (1.5), changing $\tilde{\lambda}$ by a factor arbitrarily close to 1. □

Definition 1.6. — Let X be an f-forward invariant set. We say that f on X satisfies the property *exponential shrinking of components* if there exist $\xi : 0 < \xi < 1$ and $r > 0$ such that for every $x \in X$ and positive integer n the component of $f^{-n}(B(f^n(x), r))$ containing x has diameter bounded by ξ^n.

Lemma 1.7. — *The property A is Hölder implies the property:* exponential shrinking of components, *for f on ∂A, with ξ arbitrarily close to $\lambda_{\mathrm{Ho}}^{-1}$.*

In the proof we shall use the following fact, a variant of the *telescope lemma* [**P1**, Lemma 5]:

Lemma 1.8. — *Let X be a compact set in $\overline{\mathbb{C}}$ and $f : U \to \overline{\mathbb{C}}$ be a holomorphic map on a neighbourhood of X such that $f(X) = X$. Then $(\exists\, C > 0)(\forall\, \mu > 1)(\exists\, \eta > 0)$ such that for every $x \in X$ and positive integer $n > 0$, for every $r > 0$, the disc $B := B(x_n, r)$ and every compact connected set $Y \subset B$ the following holds:*

Denote $W_j := \mathrm{Comp}_{x_{n-j}}(f^{-j}(B))$ for $j = 0, \ldots, n$. Let Y_n be an arbitrary component of $f^{-n}(Y)$ in W_n. Assume finally that $\operatorname{diam} W_j \le \eta$ *for every $j = 0, \ldots, n-1$. Then*

$$\frac{\operatorname{diam} W_n}{\operatorname{diam} B} \le C\mu^n \frac{\operatorname{diam} Y_n}{\operatorname{diam} Y}.$$

Proof of Lemma 1.8. — See [**P1**]. The idea of the proof is that if W_j is far from Crit then, denoting $Y_j = f^{n-j}(Y_n)$,

$$\frac{\operatorname{diam} W_{j+1}}{\operatorname{diam} W_j} \approx \frac{\operatorname{diam} Y_{j+1}}{\operatorname{diam} Y_j}.$$

If W_{j+1} is close to a critical point of multiplicity ν then, instead of $\approx$, the inequality $\le$ with a constant depending on ν appears on the right hand side. These cases however happen rarely as long as $\operatorname{diam} W_j$ are small. □

Proof of Lemma 1.7. — Fix[(1)] an arbitrary $n > 0$ and $x \in \partial A$. By Lemma 1.4 we can find $\alpha \in \Sigma^d$ such that $b(\alpha) \to x$. By the continuity of f for every $0 \le j \le n$ we have $b(\sigma^j(\alpha)) \to x_j$. Let $m(r)$ be the largest integer such that $\gamma_{m(r)}(\sigma^n(\alpha))$ intersects ∂B for $B := B(x_n, r)$. Denote by b' the curve in $b_{m(r)-1}(\sigma^n(\alpha))$ contained in B and joining ∂B to x_n. Denote by W_j the component of $f^{-j}(B)$ containing x_{n-j}. Denote finally by b'_j the component of $f^{-j}(b')$ contained in $b_{m(r)-1+j}(\sigma^{n-j}(\alpha))$. By Lemma 1.3 we have

$$\operatorname{diam} b'_j \le \mathrm{Const}\, \lambda_{\mathrm{Ho}}^{-(m(r)+j)}.$$

So, using Lemma 1.8 and due to $\operatorname{diam} b'$ comparable with $\operatorname{diam} B$, we obtain

$$\operatorname{diam} W_j \le \mu^j\, \mathrm{Const}\, \lambda_{\mathrm{Ho}}^{-(m(r)+j)}. \tag{1.7}$$

(1) A different proof, for polynomials – using *puzzles*, was obtained jointly by the author and Jacques Carette.

for $\mu > 1$ arbitrarily close to 1, as long as all $\operatorname{diam} W_i$ for all $i < j$ are small. Observe however that if $r \to 0$ then $m(r) \to \infty$ (more precisely $m(r) \geq (\log(1/r)/\log L) -$ Const). So if r is small enough that $\lambda_{\mathrm{Ho}}^{-m(r)}$ compensates Const, $\operatorname{diam} W_i$ are small and (1.7) holds by induction. □

Lemma 1.9. — *Let A be a Hölder domain. Then for every $\vartheta : 0 < \vartheta < \log \lambda_{\mathrm{Ho}}/\log L$ there exists $r(\vartheta) > 0$ such that for every $x \in \partial A$ and $a \leq r(\vartheta)$ such that for $W =$* $\operatorname{Comp}_x f^{-n}(B(x_n, a))$

$$\operatorname{diam} W \leq a^{\vartheta}.$$

Proof. — Set $s = [\log(r/a)/\log L]$. Let a be small enough that $s > 0$. We have chosen r and ξ according to Definition 1.6 and Lemma 1.7.

As L is a Lipschitz constant for f, we obtain for $B' := \operatorname{Comp}_{x_n} f^{-s}(B(x_{n+s}, r))$

$$B' \supset B(x_n, a).$$

By Lemma 1.7 $\operatorname{diam} \operatorname{Comp}_x f^{-n}(B') \leq \xi^{n+s}$, hence

$$\operatorname{diam} W \leq \xi^{n+s} \leq \xi^s.$$

By $s > \dfrac{\log(r/a)}{\log L} - 1$ we obtain

$$\operatorname{diam} W \leq \left(\frac{r}{L}\right)^{-\vartheta} a^{\vartheta} \quad \text{for } \vartheta = \frac{\log(1/\xi)}{\log L}.$$

□

2. A technical lemma

Lemma 2.1. — *For every $\nu \geq 2$ there exist $\varepsilon_1 : 0 < \varepsilon_1 < 1/2$ such that the following holds:*

Write $g(z) = g_u(z) = z^{\nu} + u$ for an arbitrary u with $|u| < 1$. Consider any $\Phi : g^{-1}(\mathbb{D}) \to \overline{\mathbb{C}}$ univalent and such that in the spherical metric $\operatorname{diam} \Phi(g^{-1}(\mathbb{D})) \leq \operatorname{diam} \frac{1}{2}\overline{\mathbb{C}}$. *Here $\mathbb{D}$ is the unit disc in $\mathbb{C}$, considered later with the euclidean metric. Write $F := \Phi^{-1}$ with the domain $\Phi(g^{-1}(\mathbb{D}))$. Assume*

$$u \in \Phi(g^{-1}(\mathbb{D})). \tag{2.1}$$

Moreover assume

$$|u| < \varepsilon_1 \quad \textit{and} \quad |(g \circ F)(u)| < \varepsilon_1. \tag{2.2}$$

Then either

$$\operatorname{cl} \Phi\left(g^{-1}\left(\frac{1}{2}\mathbb{D}\right)\right) \subset \frac{1}{2}\mathbb{D}, \tag{2.3}$$

or there exists $\varepsilon_2 : 0 < \varepsilon_2 < 1$ such that

$$\Phi(g^{-1}(\varepsilon_2 \mathbb{D})) \supset \varepsilon_2 \mathbb{D}. \tag{2.4}$$

Proof. — Suppose there exist $u_n \searrow 0$ and univalent Φ_n on $g_{u_n}^{-1}(\mathbb{D})$ satisfying (2.1) such that $g_{u_n}(F_n(u_n)) \to 0$ and both (2.3) and (2.4) fail. Then starting from some n the distortion of Φ_n on $V_n := g_{u_n}^{-1}(\frac{1}{2}\mathbb{D})$ is bounded by a constant Q. A reason for this, is for example the existence of a definite geometric annulus in $g_{u_n}^{-1}(\mathbb{D}) \setminus V_n$.

The sequence of the domains V_n converges in Carathéodory's sense, [**McM**, 5.1], to $V := (1/2)^{1/\nu}\mathbb{D}$ and, as all diameters of $\Phi_n(V_n)$ are uniformly bounded by $\frac{1}{2}\operatorname{diam}\overline{\mathbb{C}}$, and one can choose from (Φ_n, V_n) a subsequence convergent to certain (Φ, V).

Now notice that by (2.2) $\Phi(0) = 0$, in particular $0 \in \Phi(V)$. On the other hand by the failure of (2.3) we obtain $\operatorname{cl}\Phi(V) \not\subset \frac{1}{2}\mathbb{D}$. Hence $\operatorname{diam}\Phi(V) \geq 1/2$. Hence $\sup_V |\Phi'| \geq \frac{1}{2}2^{1/\nu}$. So $\inf_V |\Phi'| \geq Q^{-1}2^{1/(\nu-1)}$. A result is that for every $r : 0 < r \leq (1/2)^{1/\nu}$ the set $\Phi(r\mathbb{D})$ contains the disc of radius $Q^{-1}2^{1/(\nu-1)}r$ centered at 0.

Thus if $Q^{-1}2^{1/(\nu-1)}\tau^{1/\nu} > \tau$, or after rewriting:

$$\tau < \frac{1}{2}Q^{-\nu/(\nu-1)}, \tag{2.5}$$

we obtain for $g(z) = z^\nu$ the inclusion $\operatorname{cl}\Phi(g^{-1}(\tau\mathbb{D})) \supset \tau\mathbb{D}$. This implies the analogous inclusion for n large, what contradicts the assumption that (2.4) fails. □

Remark 2.2. — One could compute $\varepsilon_1, \varepsilon_2$ explicitly, however we have chosen above a more lazy way. In particular ε_2 can be chosen independent of Φ, *i.e.* the statement of the Lemma could start with: $(\forall\,\nu)(\exists\,\varepsilon_1, \varepsilon_2)\ldots$.

3. Hölder implies CE

An important role will be played by the following variant of a lemma proved in [**DPU**, Lemma 2.3 and (3.2)]

Lemma 3.1. — *Let X be a compact set in $\overline{\mathbb{C}}$ and $f : U \to \overline{\mathbb{C}}$ be a holomorphic map on U a neighbourhood of X such that $f(X) = X$. Fix $c \in \operatorname{Crit}(f) \cap X$. Assume that there is no periodic orbit in X attracting the point c.*

For every $y \in X$ write $k(y) = \max(0, -\log\operatorname{dist}(y, c))$. For $y = c$ write $k(y) = \infty$.

Then there exists a constant C_f such that for each $x \in X$ and $n \geq 1$

$$\sum_{j=0}^{n}{}' k(x_j) \leq nC_f, \tag{3.1}$$

where $\sum'$ denotes summation over all but at most one index j at which $k(x_j)$ is maximal, (∞ is also possible).

Proof. — To proceed as in [**DPU**] extend f to $\overline{\mathbb{C}}$ in a differentiable way. The observation used in [**DPU**] is that $f^n(U) \subset U$ for U small intersecting X is not possible. In case f is a rational map and X contained in Julia set, the family f^{jn} on U for $j = 1, 2, \ldots$ would be normal, what contradicts a property of Julia set. In general case we use also the additional property of f^n on U if (3.1) fails: $|(f^n)'| < 1/2$. This yields an attracting periodic orbit in X, what, together with the property $U \ni c$ also following from the construction in [**DPU**], contradicts the assumptions. $\square$

Definition 3.2. — We call A *regular* for f if for every small $r > 0$, for every $x \in \partial A$, every positive integer n and every component W of $f^{-n}(B(x,r))$ if $W \cap A \neq \varnothing$ then $f^n(W \cap A) = B(x,r) \cap A$.

The notion of *regular* is introduced *ad hoc* because we do not know how to prove our main theorem below without assuming this. Of course if A is completely invariant, *i.e.* $f^{-1}(A) = A$, then A is regular.

The reader will see that in the proof instead of $B = B(x,r)$ for all r it is sufficient to consider B boundedly distorted in many scales. To have such B satisfying Definition 3.2 it is sufficient to assume that A is Hölder and Jordan. The idea is that if B is large and $B \cap A$ is connected, then for pullbacks W_j (components of $f^{-j}(B)$) $W_j \cap A$ are also connected. If a critical value is met in ∂A then only one component of $f^{-1}(W_j \cap A) \cap W_{j+1}$ can intersect A. Otherwise their boundaries would be glued at a critical point, contradicting Jordan property. Bounded distortion and many scales are due to TCE property (see Sec.4).

Theorem 3.3. — *Let A be a Hölder immediate basin of attraction to a periodic sink for a rational map f. Assume A to be regular. Let $c \in \partial A$ be a critical point whose closure of the forward orbit is disjoint from* Crit $\setminus\{c\}$. *Then c satisfies (CE), with λ arbitrarily close to λ_{Ho}.*

Corollary 3.4. — *Let f be a polynomial and A_∞ be Hölder. Suppose there is only one critical point in $J(f)$. Then f is CE, with λ arbitrarily close to λ_{Ho}.*

Proof of Theorem 3.3. — The proof uses the procedure of the "reversed telescope" invented by Graczyk and Smirnov [**GS**, Appendix] to prove that CE2 (plus the so-called R-expansion property) implies CE. CE2 means $|(f^n)'(y)| \geq \mathrm{Const}\,\lambda^n$ for every $y \in J$ and n such that n is the smallest positive integer for which $f^n(y) \in \mathrm{Crit}$, $\lambda > 1$. Here instead of CE2 we shall use the definition of Hölder domain, the property (1.1).

Step 1. The block preceding the telescope. — Fix an arbitrary, large, n. Let $0 \leq m \leq n$ be the last time $\mathrm{dist}(x_m, \mathrm{Crit}) \leq \exp -n\delta\varepsilon$ for an arbitrary constant $\varepsilon : 0 < \varepsilon < 1$ and for δ from Lemma 1.5. Here $x = x_0 := c$. (We use the symbol x for c to distinguish the trajectory c_n of c from c it passes by.) A critical point c' such that $\mathrm{dist}(x_m, c') \leq \exp -n\delta\varepsilon$ must be c, supposed that n is large enough that $\exp -n\delta\varepsilon < \mathrm{dist}(O^+(c), \mathrm{Crit}\setminus(\{c\}))$, where $O^+(c)$ stands for the forward orbit of c (c included).

(1) If $n - m - 1 \geq \varepsilon n$ then $|(f^{n-m-1})'(x_{m+1})| \geq \lambda^{n-m-1}$ by Lemma 1.5.

(2) If $n - m - 1 < \varepsilon n$ then by Lemma 3.1 we have

$$\sum_{m<j<n} k(x_j) \leq (n-m-1)C_f,$$

the function k considered with respect to c. Hence

$$|(f^{n-m-1})'(x_{m+1})| \geq \exp\bigl(-(n-m-1)C'_f\bigr)$$

for a constant $C'_f > 0$.

Notice that by $k(x_m) > k(x_j)$ for every $j : m < j \leq n$, there is no need to exclude an exceptional j from the above sum, because the exceptional index in $\sum_{m \leq j < n}$ might be only $j = m$.

Suppose we know that there exists a constant $\lambda > 1$ such that

$$(3.2) \qquad |(f^m)'(x_1)| \geq \lambda^m.$$

Then in the case (1), (CE) for c is proved. In the case (2) we obtain

$$|(f^n)'(x_1)| \geq \lambda^m \exp\bigl(-(n-m-1)C'_f\bigr) \geq \tilde{\lambda}^n$$

for $1 < \tilde{\lambda} < \lambda$ with $\tilde{\lambda}$ arbitrarily close to λ for ε appropriately small, in particular we also obtain (CE).

Thus, we need to prove (3.2), provided

$$(3.3) \qquad \operatorname{dist}(x_m, c) \leq \exp -n\delta\varepsilon \leq \exp -m\delta\varepsilon.$$

We shall prove this with an arbitrary $\lambda : 1 < \lambda < \lambda_{\mathrm{Ho}}$ and for m large enough. More precisely, we shall prove (3.2) with the lower bound $\mathrm{Const}\, \lambda_{\mathrm{Ho}}^m$ where Const depends only on δ, ε.

Note that n large implies m large by the first inequality of (3.3) (c cannot too soon approach itself).

Step 2. Telescope: the first tube. — Define first some constants.

Let $T = [2(\delta\varepsilon\vartheta)^{-1}C_f] + 2$ for ϑ from Lemma 1.9. Let $C_4 = (\frac{1}{2}\varepsilon_1)^{-T}$ for ε_1 from Lemma 2.1.

Consider now $B := B(x_{m+1}, C_4 \operatorname{dist}(x_{m+1}, c_1))$. For every $j = 0, 1, \ldots$ define

$$W_j := \operatorname{Comp}_{x_{m+1-j}} f^{-j}(B).$$

Fix $j = j_0$ the first time W_{j+1} intersects Crit, at c' say. This can happen only with $c' = c$. Indeed, otherwise, using Lemma 1.9, we obtain

$$\operatorname{dist}(O^+(c), \operatorname{Crit} \setminus \{c\}) \leq \operatorname{dist}(c_m, c') < \operatorname{diam} W_{j+1} \leq C_4 \exp -m\delta\varepsilon\vartheta,$$

what for m large enough is not possible.

We conclude with $W_{j+1} \ni c$. We have two cases:

Case 1°. $f^j(c_1) \notin \frac{1}{2}\varepsilon_1 B$;

Case 2°. $f^j(c_1) \in \frac{1}{2}\varepsilon_1 B$.

Consider the case 2°. (Then we call $f^j : W_j \to B$ the first tube of our telescope.) In appropriate charts, in particular for B identified to $\mathbb{D}$, we can decompose f^j into $g \circ F$, the decomposition in the language of Lemma 2.1, where g corresponds to the $z \mapsto z^\nu + u$ part of f and F takes care of the rest, in particular it includes f^{j-1}.

We have also $c_1 \in \frac{1}{2}\varepsilon_1 B$ because $\frac{1}{2}\varepsilon_1 > C_4^{-1}$. Finally $\operatorname{diam} W_j \leq \frac{1}{2} \operatorname{diam} \widehat{\mathbb{C}}$ by Lemma 1.9. Thus we can apply Lemma 2.1. We multiplied c_1 by 1/2 because here we consider the spherical metric whereas in Sec.2 we considered discs $\tau\mathbb{D}$ with respect to the euclidean metric on $\mathbb{C}$.

We obtain two possibilities:

(1) The closure of $W_j' := \operatorname{Comp}_{x_{m+1-j}} f^{-j}(\tau_0 B)$ is contained in $\tau_0 B$.

τ_0 replaces here 1/2 resulting from the difference between the spherical metric on B and the euclidean on $\mathbb{D}$.

(2) There exists $\varepsilon_2 : 0 < \varepsilon_2 < 1$ such that

$$W_j'' := \operatorname{Comp}_{x_{m+1-j}} f^{-j}(\varepsilon_2 B) \supset \varepsilon_2 B \tag{3.4}$$

(Here is an explanation of the existence of $\varepsilon_1, \varepsilon_2$ that yield this alternative, for a reader who does not wish to decipher Section 2: If (1) does not hold we shrink B to $\varepsilon_2 B$ so that $\operatorname{diam} W_j'' \gg \operatorname{diam} \varepsilon_2 B$ and consider ε_1 small enough that $f^j(c_1)$ is close to x_{m+1}, hence c_1 is a "center" of the boundedly distorted W_j''. ε_1 small means also that c_1 is close to the center of $\varepsilon_2 B$. This gives (3.4).)

Notice now that (3.4) contradicts $x_{m+1-j} \in J(f)$. Thus, we can suppose that

$$\operatorname{cl} W_j' \subset \tau_0 B.$$

Step 3. The capture of expansion. — $f^j : W_j' \to \tau_0 B$ is polynomial-like, hence W_j' contains an f^j-fixed point p. In the case f is a polynomial (Corollary 3.4) $A = A_\infty$ is completely invariant, hence $p \in \partial A$ of course. Hence if we had UHPer on ∂A we would obtain for a constant $\tilde{\lambda} > 1$

$$|(f^j)'(p)| \geq \tilde{\lambda}^j. \tag{3.5}$$

(We shall come back to this discussion in Section 4. In particular UHPer on ∂A will be deduced from A Hölder, with $\tilde{\lambda} = \lambda_{\mathrm{Ho}}$.)

In the general case we do not know whether $p \in \partial A$, unfortunately. So, instead, we use the assumption A is regular. Denote f^j by F. As W_j' itersects ∂A at x_{m+1-j}, it intersects also A at, say, y^1. Write $F(y^1) = y^0$. Since y^0 has an F-preimage in $W_j' \cap A$, by the regularity of A also y^1 has an F-preimage y^2 in $W_j' \cap A$, next y^2 has, etc. Hence $|(F^i)'(y^i)| \geq R\lambda_{\mathrm{Ho}}^{ji} = (R^{1/i}\lambda_{\mathrm{Ho}}^j)^i$, where R is a constant dependent on $\operatorname{diam} B$. (If $\operatorname{diam} B$ is small then R is large. It arises from the ratio of derivatives of f^{-k} at y^0 and z in Def.1.2., resulting from a distortion bound.) Hence for an arbitrary $\tilde{\lambda} < \lambda_{\mathrm{Ho}}$ one can take i large enough and find $s : 0 \leq s < i$ such that

$$|(f^j)'(p)| \geq \tilde{\lambda}^j \quad \text{for } p = F^s(y^i). \tag{3.5$'$}$$

By construction the distortion of f^{j-1} on W_j' is bounded by a constant. Hence we have

$$|(f^{j-1})'(p)|/|(f^{j-1})'(x_{m+1-j})| \leq \text{Const}.$$

We have also $|f'(f^{j-1}(p))|/|f'(x_m)| \leq \text{Const}\, C_4^{(\nu-1)/\nu}$ for ν the multiplicity of f at c.

(We use the assumption that c_1 is peripheric in B with the factor C_4, *i.e.*

$$\operatorname{dist}(c_1, x_{m+1}) \geq C_4^{-1} \operatorname{diam} B.$$

Notice that earlier, in Step 2., we used the opposite inequality, that c_1 is close to the center.)

We conclude using (3.5) or (3.5') that

$$|(f^j)'(x_{m+1-j})| \geq \lambda_{\mathrm{Ho}}^j \operatorname{Const} C_4^{(\nu-1)/\nu}. \tag{3.6}$$

Step 4. Longer first tube. — Consider now the case 1^o. In this case instead of taking W_j for $j = j_0 + 1, \ldots$ we replace B by $\frac{1}{2}\varepsilon_1 B$ in the definition. Denote the resulting sets by $\widehat{W}_j$. We stop at $j = j_1$ such that for the first time $\widehat{W}_{j+1} \ni c$. Either we have the case 2^o now or again the case 1^o in which we continue with preimages of $\varepsilon_1^2 B$, etc. Notice that we finally arrive at the case 2^o because if we stop at $j = m$ we have $x_{m+1-j} = c_1$.

The conclusion (3.6), in the case 2^o, ending this procedure at some j_t, holds if $(\frac{1}{2}\varepsilon_1)^{t+1} < C_4^{-1}$.

We have fortunately, by Lemma 1.9, (compare Step 2.)

$$\operatorname{diam} \widehat{W}_j \leq C_4^{\vartheta} \exp -m\delta\varepsilon\vartheta.$$

Hence, by Lemma 3.1,

$$t \leq C_f(m+1)(-\vartheta \log C_4 + m\delta\varepsilon\vartheta)^{-1}$$

that is less than $T-1$ defined in Step 2. if m is large enough. The estimate $(\frac{1}{2}\varepsilon_1)^{t+1} > C_4^{-1}$ follows now from the definition of C_4.

Step 5. The number of tubes. Conclusion. — Thus, we have (3.6) for $j = j_t$. Denote this integer by k_1. We consider now $B := B(x_{m+1-k_1}, C_4 \operatorname{dist}(x_{m+1-k_1}, c_1))$ and repeat the above construction. We obtain the inequality (3.6) for x_{m+1-k_2} instead of x_{m+1-k_1} and for $j = k_2$. We continue until $\sum_{i=1}^{I} k_i = m$. We have constructed a reversed telescope **[GS**, Appendix**]**. Setting $\gamma := \operatorname{Const} C_4^{(\nu-1)/\nu}$ we conclude with

$$|(f^m)'(c_1)| \geq \gamma^I \lambda_{\mathrm{Ho}}^m. \tag{3.7}$$

Indeed, at each step $\gamma\lambda_{\mathrm{Ho}}^{k_i} \gg 1$ for m large enough, because k_i is large; c cannot too soon approach itself. So, by (3.6), using bounded distortion for the appropriate branch of $f^{-(k_i-1)}$ on the appropriate $B' = \operatorname{Comp} f^{-1}(\frac{1}{2}\varepsilon_1 B)$, we obtain

$$\cdots < \operatorname{dist}(x_{m-k_1-k_2}, c) < \operatorname{dist}(x_{m-k_1}, c) < \exp -m\delta\varepsilon, \tag{3.8}$$

resulting from the related inequalities concerning $\operatorname{dist}(x_{m-k_1-\cdots-k_i+1}, c_1)$.

Formally, (3.8), the construction of the i-th tube and k_i large, are proved alternately by induction over i. In particular $m - k_1 - \cdots - k_i \geq k_I$ for $i < I$ is also large. Here $\operatorname{dist}(x_{m-k_1-\cdots-k_i}, c) \leq \exp -m\delta\varepsilon \leq \exp -(m - k_1 - \cdots - k_i)\delta\varepsilon$ replaces (3.3)

By Lemma 3.1 we obtain a bound for I:

$$I \leq (\delta\varepsilon)^{-1} C_f + 1.$$

Thus (3.7) yields (3.2), with λ arbitrarily close to λ_{Ho} for m large enough. Therefore as noticed at the beginning we obtain (CE). $\square$

It is possible to prove Theorem 3.3 without refering to Lemmas 1.7, 1.9[(2)]. The method relies on the following fact (weaker than Lemma 1.9)

Lemma 3.5. — *For A Hölder, there exists $C > 0$ such that for every $Q \geq 1$, $r > 0$, positive integer n and $x \in \partial A$, for B a topological disc of diameter r boundedly distorted around $f^n(x)$, say containing $B(x_n, r/4)$, for W the component of $f^{-n}(B(x_n, r))$ containing x, if $f^n|_W$ is univalent and the distortion of f^{-n} on $B = B(x_n, r)$ is bounded by Q (namely $\sup|(f^{-n})'|/\inf|(f^{-n})'| \leq Q$), then*

$$\operatorname{diam} W \leq CQr^{\vartheta}, \quad \textit{for } \vartheta = \log \lambda_{\mathrm{Ho}} / \log L.$$

Proof. — Consider $\alpha \in \Sigma^d$ such that $b(\alpha)$ converges to x. W contains a round disc centered at x of diameter equal to $Q^{-1} \cdot \frac{1}{4} \operatorname{diam} W$. By Lemma 1.3 $\lambda_{\mathrm{Ho}}^{-t} <$ Const $\cdot Q^{-1} \operatorname{diam} W$, in particular $t = [(\mathrm{Const} + \log Q + \log(-\operatorname{diam} W))/\log \lambda_{\mathrm{Ho}}] + 1$, implies $z_t(\alpha) \in W$. Hence $z_{t-n}(\sigma^n(\alpha)) \in B$. So $t - n \geq (\log 1/r)/\log L -$ Const.

Now $t \geq t - n$ implies

$$[(\mathrm{Const} + \log Q + \log(-\operatorname{diam} W))/\log \lambda_{\mathrm{Ho}}] + 1 \geq (\log 1/r)/\log L - \mathrm{Const},$$

hence after exponentiating the both sides, $\operatorname{diam} W \leq \mathrm{Const}\, Qr^{\vartheta}$. $\square$

Now, in Proof of Theorem 3.3, in Step 2, one can define W_j with the use of the "shrinking neighbourhoods" procedure, see [**P2**, Sec.2]:

For $B := B(x_{m+1}, C_4 \operatorname{dist}(x_{m+1}, c_1))$, for every $j = 0, 1, \ldots$ write $B_{[j} := \beta_j B$ for $\beta_j = \prod_{i=1}^{j}(1 - b_i)$ for $b_i := \exp -\kappa i$ for an arbitrary $\kappa > 0$, close to 0.

Write $B' = \mathrm{Comp}_{x_m} f^{-1}(B)$ and $B'_{[j} = \mathrm{Comp}_{x_m} f^{-1}(B_{[j})$. Finally define $W_j :=$ $\mathrm{Comp}_{x_{m+1-j}} f^{-(j-1)}(B'_{[j})$

For $j \leq j_0 - 1$ we have by construction f^{j-1} univalent on $f(W_{j+1})$ with distortion bounded by $\exp(-\mathrm{Const}\, \kappa j)$ (using Koebe distortion theorem). Hence $\operatorname{diam} W_j$ can be estimated due to Lemma 3.5 by $\mathrm{Const}\, r^{\vartheta}$ for ϑ arbitrarily close to $\log \lambda_{\mathrm{Ho}} / \log L$.

We do the same trick in Step 4. in the definition of $\widehat{W}_j$.

4. TCE rational maps and the topological invariance of CE

In this section we shall provide a new proof of the theorem proved first in [**PR2**], that CE is a topological condition provided the following holds:

Condition $(*)$. — For every exposed critical point $c \in J(f)$ it holds

$$\mathrm{cl} \bigcup_{n>0} f^n(c) \cap (\mathrm{Crit} \setminus \{c\}) = \varnothing.$$

(2) We followed that way in the first distributed version of the paper.

In other words for no critical point in $J(f)$ its ω-limit set contains another critical point. This condition was already present in Theorem 3.3. (Recall that *exposed* means the forward trajectory of c does not meet other critical points.)

Let us introduce some properties of a rational map f related to CE and to *exponential shrinking of components*, compare Def. 1.6. (We follow the numeration and terminology from **[NP]**.)

2^o (*exponential shrinking of components*) There exist $0 < \xi_2 < 1$ and $r_2 > 0$ such that for every $x \in J$ every $n > 0$ and $W = \mathrm{Comp}_x f^{-n}(B(f^n(x), r_2))$ one has $\operatorname{diam} W \leq \xi_2^n$.

3^o (*exponential shrinking of components at critical points*) The same as above, but only for W containing a critical point.

4^o (*finite criticality* or *topological Collet-Eckmann*, abbr. TCE) There exist $M > 0, P > 1$ and $r > 0$ such that for every $x \in J$ there exists an increasing sequence of positive integers $n_j, j = 1, 2, \ldots$ such that $n_j \leq Pj$ and for each j

$$\#\{i : 0 \leq i < n_j,\ \mathrm{Comp}_{f^i(x)} f^{-(n_j - i)} B(f^{n_j}(x), r) \cap \mathrm{Crit} \neq \varnothing\} \leq M.$$

5^o (*mean exponential shrinking of components*) There exist $P > 1, 0 < \xi_5 < 1$ and $r_5 > 0$ such that for every $x \in J$ there exists an increasing sequence of positive integers $n_j = n_j(x)$, $j = 1, 2, \ldots$ such that $n_j \leq Pj$ and for each j one has

$$\operatorname{diam} \mathrm{Comp}_x f^{-n_j}(B(f^{n_j}(x), r_5)) \leq \xi_5^{n_j}.$$

Another interesting condition is *uniform hyperbolicity on periodic orbits* (abbr: UHPer): There exists $\lambda_{\mathrm{Per}} > 1$ such that every periodic $p \in J(f)$ satisfies

$$|(f^k)'(p)| \geq \lambda_{\mathrm{Per}}^k.$$

where k is a period of p.

Formally we do not restrict 3^o to critical points in J, but this condition implies there are no critical points outside J attracted to J (which is equivalent to the absence of parabolic periodic orbits).

Notice that 4^o is a topological condition, *i.e.* if it is satisfied by f and there exists a homeomorphism h from a neighbourhood of $J(f)$ to a neighbourhood of $J(g)$ such that $h(J(f)) = J(g)$ and $hf = gh$ then 4^o holds also for g.

The implications CE$\Rightarrow 2^o \Rightarrow 3^o \Rightarrow 4^o$ have been proven in **[PR1]** (see also **[NP]**). $4^o \Rightarrow 5^o$ has also been proven in **[PR1]**. Here we shall prove $5^o \Rightarrow 2^o$ and next $2^o \Rightarrow$CE provided $(*)$. Thus we shall prove:

Theorem 4.1. — *If f is topological Collet-Eckmann and satisfies $(*)$ (a particular case is that there is only one critical point in J), then f is CE.*

In view of the above discussion we shall obtain

Corollary 4.2. — *(CE & $(*)$) is a topological property.*

Remark that it is straightforward to prove $5^o \Rightarrow$ UHPer, see Lemma 4.7. below. Unfortunately we cannot prove UHPer$\Rightarrow$ CE, to mimic the interval case [**NP**], [**NS**].

Let us start the proofs with $5^o \Rightarrow 2^o$ which is surprisingly easy.

Lemma 4.3. — *Mean exponential shrinking of components implies exponential shrinking of components.*

Proof. — Fix an arbitrary $x \in J(f)$ and $n \geq 0$. Write $B_j := B(x_j, r_5)$ for $j = 0, 1, \ldots, n$. Set $t_0 := 0$ and define the increasing sequence of integers $0 < t_1 < t_2, \ldots$ by induction as follows: Given t_i take t_{i+1} such that $t_i + (n - t_i)/2P \leq t_{i+1} \leq n$ and for

$$K_{i+1} := \mathrm{Comp}_{x_{t_i}} f^{-(t_{i+1}-t_i)}(B_{t_{i+1}}),$$

$$\mathrm{diam}\, K_{i+1} \leq \xi_5^{t_{i+1}-t_i}. \tag{4.1}$$

This is possible by the definitions of the constants in 5^o. 5^o implies that the number of $n_j = n_j(x_{t_i})$'s not exceeding $m = Pk$ is at least k for every $k = 1, 2, \ldots$. So for every $m \geq 0$ we obtain $\#\{n_j : n_j \leq m\} \geq [m/P]$ (the integer part of m/P). In particular for $m \geq 2P$ we obtain $\#\{n_j : n_j \leq m\} \geq m/2P$, hence $\{n_j : m/2P \leq n_j \leq m\} \neq \varnothing$. Finally apply this to $m = n - t_i$ and choose as t_{i+1} any n_j from the latter nonempty set.

If $\xi_5^{t_{i+1}-t_i} \leq r_5$, *i.e.*

$$t_{i+1} - t_i \geq \frac{\log r_5}{\log \xi_5} \tag{4.2}$$

then by (4.1)

$$K_{i+1} \subset B_{x_{t_i}}. \tag{4.3}$$

Suppose $i = I$ is the smallest integer such that either $n - t_I < 2P$ so we may not find t_{I+1} satisfying (4.1), or (4.2) does not hold. The latter: $t_{I+1} - t_I < \log r_5 / \log \xi_5$, together with $t_{I+1} - t_I \geq (n - t_I)/2P$ imply $n - t_I \leq 2P \log r_5 / \log \xi_5$. Denote the maximum of this constant and $2P$ by C.

Due to (4.3) for every $i = 1, \ldots, I-1$ we have a "telescope" so we obtain

$$\mathrm{Comp}_{x_{t_1}} f^{-(t_I - t_1)}(B_{t_I}) \subset B_{t_1}.$$

Hence, applying also (4.1) for $i = 0$,

$$\mathrm{diam}\, \mathrm{Comp}_x (f^{-t_I}(B_{t_I})) \leq \mathrm{diam}\, \mathrm{Comp}_x f^{-t_1}(B_{t_1}) \leq \xi_5^{t_1} \leq \xi_5^{n/2P} \tag{4.4}$$

provided $n \geq 2P$ (otherwise $I = 0$, *i.e.* there is no t_1).

Finally, due to $n - t_I$ bounded by C, we can replace in (4.4) B_{t_I} by

$$\mathrm{Comp}_{x_{t_I}} f^{-(n-t_I)}(B(x_n, r_2))$$

for a constant r_2 small enough. We conclude with

$$\mathrm{diam}\, \mathrm{Comp}_x f^{-n}(B(x_n, r_2)) \leq \xi_5^{n/2P}$$

which proves 2^o with $\xi_2 = \xi_5^{1/2P}$. The case $n < 2P$ is trivial, r_2 small enough does the job. □

Lemma 4.4. — *Assume 2^o. Then there exist $\vartheta : 0 < \vartheta < 1$ such that for every $a > 0$ small enough, for every $x \in J(f)$ and every $n \geq 0$*

$$\operatorname{diam} \operatorname{Comp}_x f^{-n}(B(f^n(x), a)) \leq a^{\vartheta}. \tag{4.5}$$

Proof. — See Proof of Lemma 1.9. □

Analogously to Lemma 1.5 we have the following

Lemma 4.5. — *Assume 2^o. Then there exist $\lambda > 1$, $\delta > 0$ and an integer $n_0 > 0$ such that for every $n \geq n_0$ and $x \in J(f)$, if for every $j = 0, \ldots, n-1$*

$$\operatorname{dist}(x_j, \operatorname{Crit}) > \exp -\delta n,$$

then $|(f^n)'(x)| > \lambda^n$.

Remark 4.6. — This Lemma will be proved with λ arbitrarily close to ξ_2^{-1}. Notice that this, for ∂A rather than $J(f)$, together with Lemma 1.7, give a new proof of Lemma 1.5.

Proof of Lemma 4.5. — Consider arbitrary $\delta, \varepsilon > 0$. Let

$$s := \Big[\frac{\log \varepsilon}{\log \xi_2} + \frac{\delta n}{-\log \xi_2}\Big] + 1.$$

Then, by 2^o, for all $0 \leq j \leq n$

$$\operatorname{diam} \operatorname{Comp}_{x_{n-j}} f^{-s-j}(B(x_{n+s}, r_2)) \leq \xi_2^{s+j} \leq \xi_2^s \leq \varepsilon \exp -\delta n. \tag{4.6}$$

Now let $B = B(x_n, r_2 \exp -\delta M n)$ for $M = [\log L/(-\log \xi_2)] + 1$. Then for n large enough we obtain, using (4.6) for $j = 0$ and the definition of s,

$$\operatorname{Comp}_{x_n} f^{-s}(B(x_{n+s}, r_2)) \supset B.$$

Let $W_n = \operatorname{Comp}_x f^{-n}(B)$. Then there exists $y \in W_n$ such that

$$|(f^n)'(y)| \geq \frac{\operatorname{diam} B}{\operatorname{diam} W_n} \geq (2r_2 \exp -\delta M n)\xi_2^n,$$

where the second inequality follows from 2^o. Now, as in Proof of Lemma 1.5, for ε small, with the use of (4.6), we can switch from y to x, hence $|(f^n)'(x)| \geq \lambda^n$, for λ arbitrarily close to ξ_2^{-1} if δ, ε are appropriately small. □

Lemma 4.7. — *2^o implies UHPer on $J(f)$. Moreover $\lambda_{\mathrm{Per}} = \xi_2^{-1}$.*

Proof. — For each periodic point $x \in J(f)$, with $f^k(x) = x$, we consider the backward trajectory $x^j = f^{Nk-j}(x)$, N such that $Nk - j > 0$. By 2^o, for j large enough, W_j are so small that shrinking of W_j is comparable to decreasing of derivatives of the respective branches of f^{-j} (critical points are far away, so there is almost no distortion). Moreover we obtain $|(f^{Nk})'(x)| \geq \text{Const}\, \xi_2^{-Nk}$ for every positive N with the same Const, that implies $|(f^k)'(x)| \geq \xi_2^{-k}$. Compare [**NP**, section 2]. $\square$

Proof of Theorem 4.1. — It is sufficient to prove that 2^o and $(*)$ imply (CE) for every exposed critical point in J. The proof is the same as the proof of Theorem 3.3. Having obtained a mapping corresponding to the polynomial-like $f^j|_{W_j'} \to \tau_0 B$ we find a periodic point p and we refer to UHPer, compare (3.5). (We do not have the harder case (3.5')). $\square$

5. TCE repellers

Call a pair (X, f) a *holomorphic invariant set* if $X \subset \overline{\mathbb{C}}$ is compact and $f : X \to \overline{\mathbb{C}}$ is defined on a neighbourhood of X and $f(X) = X$. (Recall that in Lemmas 1.8 and 3.1 we considered already such pairs.) We say that holomorphic invariant sets (X, f) and (Y, g) are *topologically conjugate* if there exist neighbourhoods U_X, U_Y of X, Y respectively, and a homeomorphism $h : U_X \to U_V$ such that $hf = gh$. Recall that a property of holomorphic invariant sets is called *topological* if for every (X, f) and (Y, g) topologically conjugate, if (X, f) satisfies this property then (Y, g) satisfies this too. Sometimes we restrict the space of holomorphic invariant sets under consideration to those that satisfy certain property (not necessarily topological, for example to those f's that extend to rational functions).

For example it is easy to see (and is well-known) that the *expanding property* (namely $|(f^k)'| > 1$ for a positive integer k), is topological. An argument is that expanding is equivalent to 4^o with n_j being the sequence of all positive integers and $M = 0$, that is of course a topological condition.

Here we consider (X, f) with properties weaker than expanding, namely with: $2^o - 5^o$ with J replaced by X, W intersecting X and f not necessarily extendable to a rational function on the Riemann sphere.

One can define also CE as (CE) for every exposed $c \in W$ for W intersecting X.

We call a holomorphic invariant set (X, f) a *holomorphic repeller* if there exists a neighbourhood V of X in the domain of f such that $(\forall\, x \in V \setminus X)(\exists\, n > 0)$ such that $f^n(x) \notin V$.

We have the following

Proposition 5.1. — *For (X, f) holomorphic invariant sets, $5^o \Rightarrow 2^o \Rightarrow 3^o \Rightarrow 4^o$. Moreover, for (X, f) holomorphic repellers, or holomorphic invariant sets such that f extends to a rational function and $X \subset J(f)$, the properties 2^o, 3^o, 4^o and 5^o are equivalent. Then all of them are topological properties, CE implies each of $2^o - 5^o$ and conversely, provided $(*)$ from Section 4.*

Proof of Proposition 5.1. — As mentioned in Sec.4 the proof of $3^o \Rightarrow 4^o$ has been done in [**PR1**, Lemma 2.2] for f rational ($X = J$ has been considered there, but for

X strictly in J the proof is the same). For (X, f) an arbitrary holomorphic invariant set we need to refer to **[DPU]** as it is stated here in Lemma 3.1. The assumptions are satisfied: if there existed a periodic point $p \in X$ whose periodic orbit attracts a critical point $c \in X$, then for $x = c$ the condition 2^o would not hold.

As mentioned in Sec.4 the implication $4^o \Rightarrow 5^o$ for f rational has been proven in **[PR1]**. For (X, f) holomorphic repeller we refer to the following fact proved in [**PR2**, Appendix]:

If $X \neq \widehat{\mathbb{C}}$ then 4^o implies that X is nowhere dense.

Now repeating [**PR1**, (2.6)] one uses the repelling property to know that all the maps $f^n : W \to B$, for every $B(x,r), x \in X$, r small and every W a component of $f^{-n}(B)$ intersecting X, are proper. This is needed in the proof that if f^n have uniformly bounded criticalities on W, then the respective preimages of $\frac{1}{2}B$ have diameters shrinking to 0 as $n \to \infty$.

To prove the latter fact we find a little disc D in $\frac{1}{2}B \setminus X$, so that the components W' of f^n-preimages of D in W have diameters shrinking to 0. Such D exists due to X nowhere dense and $W' \to X$. Finally we use a bounded distortion lemma in a bounded criticality setting (for example [**PR1**, Lemma 2.1].

$5^o \Rightarrow 2^o$ is automatic, see Lemma 4.3. The proof that CE implies 4^o is the same as in **[PR1]** and the proof of the opposite implication is the same as in Section 4. □

We call a holomorphic repeller satisfying any of the properties $2^o - 5^o$ a *topological Collet-Eckmann repeller*, abbr. TCE repeller.

Proposition 5.2. — *Let $X = \partial A$ for a connected open domain $A \subset \overline{\mathbb{C}}$. Let f be a holomorphic map defined on a neighbourhood U of X such that $f(U \cap A) \subset A$, $f(X) = X$ Then A Hölder implies (X, f) is TCE (*i.e. *it satisfies 4^o). Coversely, if (X, f) is TCE and additionally it is a repeller or f extends to a rational map on $\overline{\mathbb{C}}$, then A is Hölder.*

Proof. — Assume that A is Hölder. Definition 1.2 is still valid except that one considers only $z \in A$ close to ∂A. Observe now that Hölder implies 2^o. The proof is similar to Proof of Lemma 1.7., except that in this situation one needs *Markov geometric coding tree.* Instead of one point z as in Sec.1, choose a finite family $Z \subset A$ in a small Hausdorff distance from the whole X and join each $z^j \in f^{-1}(Z)$ by a curve γ^j to a point in Z, so that γ^j is close to ∂A, in particular in the domain of f^{-1}.

Next 2^o implies TCE by Proposition 5.1.

In the opposite direction TCE implies 2^o by Proposition 5.1. Next we prove that A is Hölder: Consider a disc $B := B(x,r)$ for $x \in X$ such that $\operatorname{diam} W \leq \xi^n$ for W components of $f^{-n}(B)$ intersecting X (compare property 2^o) and consider $D = B(z, \delta) \subset A \cap B$. Let W' be a component of $f^{-n}(D)$ in W. Hence $\operatorname{diam} W' \leq \xi^n$ so by bounded distortion $|(f^n)'(y)| \geq \mathrm{Const}\, \xi^{-n}$ for $y \in W', f^n(y) = z$. Compare [**GS**, Sec.5] and [**PR1**, Sec.3]. □

We obtain an immediate

***Corollary 5.3* (Rigidity of Hölder domains).** — *Let A be a connected open domain $A \subset \overline{\mathbb{C}}$. Let f be a holomorphic map defined on a neighbourhood U of ∂A such that $f(U \cap A) \subset A$, $f(\partial A) = \partial A$. Suppose A is Hölder. Let g be a holomorphic map on a neighbourhood of a compact set $Y \subset \overline{\mathbb{C}}$ such that f on a neighbourhood of $X = \partial A$ is conjugated by a homeorphism h to g and $h(X) = Y$. Assume that g extends to a rational function on $\overline{\mathbb{C}}$ or assume that (X, f) (hence (Y, g)) are repellers. Then the component of $\overline{\mathbb{C}} \setminus Y$ intersecting $h(A)$ is Hölder.*

Let us underline that we allow above critical points in X to be in the ω-limit set of other critical points. Proposition 5.2 and Corollary 5.3 are much easier than the corresponding Theorem 3.3 and Corollary 4.2.

Remark finally that in between *holomorphic expanding repellers* (*i.e.* holomorphic repellers with expanding property) and *holomorphic TCE repellers* there lies the class of *holomorphic semihyperbolic repellers*, that is satisfying the property 4^o with n_j being the sequence of all positive integers. Semihyperbolicity is of course a topological condition.

Notice that this semihyperbolicity is equivalent to 2^o with all $f^n|_W$ of uniformly bounded criticality. Notice also that for compact nowhere dense repellers semihyperbolicity is equivalent to the assumption that critical points in X are nonrecurrent, see **[CJY]**. Parabolic points cannot happen for repellers.

If $X = \partial A$ for A a basin of a sink, I believe that f semihyperbolic is equivalent to A John. This has been proven in the case $A = A_\infty$ for polynomial f in **[CJY]**.

References

[CJY] L. Carleson, P. Jones, J.-C. Yoccoz, *Julia and John,* Bol. Soc. Bras. Mat. 25 (1994), 1–30.

[DPU] M. Denker, F. Przytycki, M. Urbański: On the transfer operator for rational functions on the Riemann sphere. Ergod. Th. and Dynam. Sys. 16.2 (1996), 255–266.

[GS] J. Graczyk, S. Smirnov, *Collet, Eckmann and Hölder,* Invent. Math. 133.1 (1998), 69–96.

[NP] T. Nowicki, F. Przytycki, *Topological invariance of the Collet-Eckmann property for S-unimodal maps,* Fund. Math. 155 (1998), 33–43.

[McM] C. McMullen, *Complex Dynamics and Renormalization,* Princeton University Press, Princeton NJ, 1994.

[NS] T. Nowicki, D. Sands, *Non-uniform hyperbolicity and universal bounds for S-unimodal maps,* Preprint, April 1996.

[Po] Ch. Pommerenke, *Boundary Behaviour of Conformal Maps,* Springer-Verlag, Berlin, 1992.

[P1] F. Przytycki, *On the Perron-Frobenius-Ruelle operator for rational maps on the Riemann sphere and for Hölder continuous functions,* Bol. Soc. Bras. Mat. 20.2 (1990), 95–125.

[P2] F. Przytycki, *Iterations of holomorphic Collet-Eckmann maps: Conformal and invariant measures,* Trans. AMS 350.2 (1998), 717–742.

[PR1] F. Przytycki, S. Rohde, *Porosity of Collet-Eckmann Julia sets,* Fund. Math. 155 (1998), 189–199.

[PR2] F. Przytycki, S. Rohde, *Rigidity of holomorphic Collet-Eckmann repellers,* Preprint, July 1997. To appear in Arkiv för Math.

[PUZ] F. Przytycki, M. Urbański, A. Zdunik, *Harmonic, Gibbs and Hausdorff measures for holomorphic maps* Part 1 in Annals of Math. 130 (1989), 1–40. Part 2 in Studia Math. 97.3 (1991), 189–225.

[PZ] F. Przytycki, A. Zdunik, *Density of periodic sources in the boundary of a basin of attraction for iteration of holomorphic maps: geometric coding trees technique,* Fund. Math. 145 (1994), 65–77.

F. Przytycki, Institute of Mathematics, Polish Academy of Sciences, ul. Śniadeckich 8, 00950 Warszawa, Poland • *E-mail :* `feliksp@impan.gov.pl`

DIERK SCHLEICHER

Rational parameter rays of the Mandelbrot set

Astérisque, tome 261 (2000), p. 405-443
<http://www.numdam.org/item?id=AST_2000__261__405_0>

Astérisque
261, 2000, p. 405–443

RATIONAL PARAMETER RAYS OF THE MANDELBROT SET

by

Dierk Schleicher

Abstract. — We give a new proof that all external rays of the Mandelbrot set at rational angles land, and of the relation between the external angle of such a ray and the dynamics at the landing point. Our proof is different from the original one, given by Douady and Hubbard and refined by P. Lavaurs, in several ways: it replaces analytic arguments by combinatorial ones; it does not use complex analytic dependence of the polynomials with respect to parameters and can thus be made to apply for non-complex analytic parameter spaces; this proof is also technically simpler. Finally, we derive several corollaries about hyperbolic components of the Mandelbrot set.

Along the way, we introduce partitions of dynamical and parameter planes which are of independent interest, and we interpret the Mandelbrot set as a symbolic parameter space of kneading sequences and internal addresses.

1. Introduction

Quadratic polynomials, when iterated, exhibit amazingly rich dynamics. Up to affine conjugation, these polynomials can be parametrized uniquely by a single complex variable. The Mandelbrot set serves to organize the space of (conjugacy classes of) quadratic polynomials. It can be understood as a "table of contents" to the dynamical possibilities and has a most beautiful structure. Much of this structure has been discovered and explained in the groundbreaking work of Douady and Hubbard [**DH1**], and a deeper understanding of the fine structure of the Mandelbrot set is a very active area of research. The importance of the Mandelbrot set is due to the fact that it is the simplest non-trivial parameter space of analytic families of iterated holomorphic maps, and because of its universality as explained by Douady and Hubbard [**DH2**]: the typical local configuration in one-dimensional complex parameter spaces is the Mandelbrot set (see also [**McM**]).

1991 ***Mathematics Subject Classification.*** — 30C20, 30D05, 30D40.
Key words and phrases. — Mandelbrot set, kneading sequence, internal address, external ray, parameter ray, hyperbolic component.

Unfortunately, most of the beautiful results of Douady and Hubbard on the structure of the Mandelbrot set are written only in preliminary form in the preprints **[DH1]**. The purpose of this article is to provide concise proofs of several of their theorems. Our proofs are quite different from the original ones in several respects: while Douady and Hubbard used elaborate perturbation arguments for many basic results, we introduce partitions of dynamical and parameter planes, describe them by symbolic dynamics, and reduce many of the questions to a combinatorial level. We

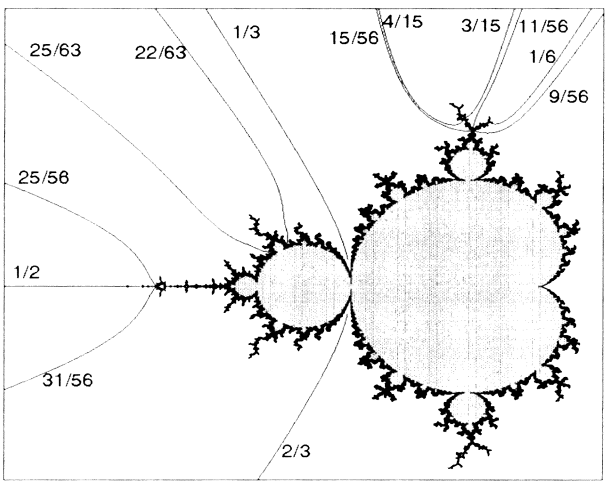


FIGURE 1. *The Mandelbrot set and several of its parameter rays which are mentioned in the text. Picture courtesy of Jack Milnor.*

feel that our proofs are technically significantly simpler than those of Douady and Hubbard. An important difference for certain applications is that our proof does not use complex analytic dependence of the maps with respect to the parameter and is therefore applicable in certain wider circumstances: the initial motivation for this research was a project with Nakane (see **[NS]** and the references therein) to understand the parameter space of antiholomorphic quadratic polynomials, which depends only real-analytically on the parameter. Of course, the "standard proof" using Fatou coordinates and Ecalle cylinders, as developed by Douady and Hubbard and elaborated by Lavaurs [**La2**], is a most powerful tool giving interesting insights; it has had many important applications. Our goal is to present an alternative approach in order to enlarge the toolbox for applications in different situations.

The fundamental result we want to describe in this article is the following theorem about landing properties of external rays of the Mandelbrot set, a theorem due to Douady and Hubbard; for background and terminology, see the next section.

Theorem 1.1 **(The Structure Theorem of $\mathcal{M}$).** — *Parameter rays of the Mandelbrot set at rational angles have the following landing properties:*

(1) *Every parameter ray at a periodic angle ϑ lands at a parabolic parameter such that, in its dynamic plane, the dynamic ray at angle ϑ lands at the parabolic orbit and is one of its two characteristic rays.*
(2) *Every parabolic parameter c is the landing point of exactly two parameter rays at periodic angles. These angles are the characteristic angles of the parabolic orbit in the dynamic plane of c.*
(3) *Every parameter ray at a preperiodic angle ϑ lands at a Misiurewicz point such that, in its dynamic plane, the dynamic ray at angle ϑ lands at the critical value.*
(4) *Every Misiurewicz point c is the landing point of a finite non-zero number of parameter rays at preperiodic angles. These angles are exactly the external angles of the dynamic rays which land at the critical value in the dynamic plane of c.*

(The parameter $c = 1/4$ is the landing point of a single parameter ray, but this ray corresponds to external angles 0 and 1; we count this ray twice in order to avoid having to state exceptions.)

The organization of this article is as follows: in Section 2, we describe necessary terminology from complex dynamics and give a few fundamental lemmas. Section 3 contains a proof of the periodic part of the theorem, and along the way it shows how to interpret the Mandelbrot set as a parameter space of kneading sequences. The preperiodic part of the theorem is then proved in Section 4, using properties of kneading sequences. In the final Section 5, we derive fundamental properties of hyperbolic components of the Mandelbrot set. Most of the results and proofs in this paper work also for "Multibrot sets": these are the connectedness loci of the polynomials $z^d + c$ for $d \geq 2$.

This article is an elaborated version of Chapter 2 of my Ph.D. thesis [**S1**] at Cornell University, written under the supervision of John Hubbard and submitted in the summer of 1994. It is part of a mathematical ping-pong with John Milnor: it builds at important places on the paper [**GM**]; recently Milnor has written a most beautiful new paper **[M2]** investigating external rays of the Mandelbrot set from the point of view of "orbit portraits", i.e., landing patterns of periodic dynamic rays. I have not tried to hide how much both I and this paper have profited from many discussions with him, as will become apparent at many places. This paper, as well as Milnor's new one, uses certain global counting arguments to provide estimates, but in different directions. It is a current project **[ES]** to combine both approaches to provide a new, more conceptual proof without global counting. Proofs in a similar spirit of further fundamental properties of the Mandelbrot set can be found in **[S2]**.

Acknowledgements. — It is a pleasure to thank many people who have contributed to this paper in a variety of ways. I am most grateful to John Hubbard for so much help and friendship over the years. To John Milnor, I am deeply indebted for the ping-pong mentioned above, as well as for many discussions and a lot of encouragement along the way. Tan Lei contributed very helpful corrections and suggestions most constructively. Many more people have shared their ideas and understanding with me and will recognize their contributions; they include Bodil Branner, Adrien Douady, Karsten Keller, Eike Lau, Jiaqi Luo, Misha Lyubich, Shizuo Nakane, Chris Penrose, Carsten Petersen, Johannes Riedl, Mitsu Shishikura and others. I am also grateful to the Institute for Mathematical Sciences in Stony Brook for its inspiring environment and support. Finally, special thanks go to Katrin Wehrheim for a most helpful suggestion.

2. Complex Dynamics

In this section, we briefly recall some results and notation from complex dynamics which will be needed in the sequel. For details, the notes **[M1]** by Milnor are recommended and, of course, the work **[DH1]** by Douady and Hubbard which is the source of most of the results mentioned below.

By affine conjugation, quadratic polynomials can be written uniquely in the normal form $p_c : z \mapsto z^2 + c$ for some complex parameter c. For any such polynomial, the filled-in Julia set is defined as the set of points z with bounded orbits under iteration. The Julia set is the boundary of the filled-in Julia set. It is also the set of points which do not have a neighborhood in which the sequence of iterates is normal (in the sense of Montel). Julia set and filled-in Julia set are connected if and only if the only critical point 0 has bounded orbit; otherwise, these sets coincide and are a Cantor set. The Mandelbrot set $\mathcal{M}$ is the *quadratic connected locus*: the set of parameters c for which the Julia set is connected. Julia sets and filled-in Julia sets, as well as the Mandelbrot set, are compact subsets of the complex plane. The Mandelbrot set is known to be connected and full (i.e. its complement is connected).

Douady and Hubbard have shown that Julia sets and the Mandelbrot set can profitably be studied using external rays: for a compact connected and full set $K \subset \mathbb{C}$, the Riemann mapping theorem supplies a unique conformal isomorphism Φ_K from the exterior of K to the exterior of a unique disk $\overline{D}_R = \{z \in \mathbb{C} : |z| \leq R\}$ subject to the normalization condition $\lim_{z\to\infty} \Phi(z)/z = 1$. The inverse of the Riemann map allows to transport polar coordinates to the exterior of K; images of radial lines and centered circles are called *external rays* and *equipotentials*, respectively. For a point $z \in \mathbb{C} - K$ with $\Phi(z) = re^{2\pi i\vartheta}$, the number ϑ is called the *external angle* and $\log r$ is called the *potential* of z. External angles live in $\mathbb{S}^1$; we will always measure them in full turns, i.e., interpreting $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$. Sometimes, it will be convenient to count the two angles 0 and 1 differently and have external angles live in $[0, 1]$. Potentials are parametrized by the open interval $(\log R, \infty)$. An external ray at angle ϑ is said to *land* at a point z if $\lim_{r\searrow\log R} \Phi_K^{-1}(re^{2\pi i\vartheta})$ exists and equals z. For general compact connected full sets K, not all external rays need to land. By Carathéodory's theorem, local connectivity of K is equivalent to landing of all the rays, with the landing points depending continuously on the external angle.

For all the sets we consider here, it turns out that the conformal radius R is necessarily equal to 1. In order to avoid confusion, we will replace the term "external ray" by *dynamic ray* or *parameter ray* according to whether it is an external ray of a filled-in Julia set or of the Mandelbrot set.

For $c \in \mathcal{M}$, the filled-in Julia set K_c is connected. For brevity, we will denote the preferred Riemann map by φ_c, rather than Φ_{K_c}. A classical theorem of Böttcher asserts that this map conjugates the dynamics outside of K_c to the squaring map outside the closed unit disk: $\varphi_c \circ p_c = (\varphi_c)^2$. A dynamic ray is periodic or preperiodic whenever its external angle is periodic or preperiodic under the doubling map on $\mathbb{S}^1$. The periodic and preperiodic angles are exactly the rational numbers. More precisely, a rational angle is periodic iff, when written in lowest terms, the denominator is odd; if the denominator is even, then the angle is preperiodic. It is well known [**M1**, Section 18] that dynamic rays of connected filled-in Julia sets always land whenever their external angles are rational. The landing points of periodic (resp. preperiodic) rays are periodic (resp. preperiodic) points on repelling or parabolic orbits. Conversely, every repelling or parabolic periodic or preperiodic point of a connected Julia set is the landing point of one or more rational dynamic rays; preperiods and periods of all the rays landing at the same point are equal.

If a quadratic Julia set is a Cantor set, then there still is a Böttcher map φ_c near infinity conjugating the dynamics to the squaring map. One can try to extend the domain of definition of the Böttcher map by pulling it back using the conjugation relation. However, there are problems about choosing the right branch of a square root needed in the conjugation relation. The absolute value of the Böttcher map is independent of the choices and allows to define potentials outside of the filled-in Julia set. The set of points at potentials exceeding the potential of the critical point is simply connected and the map φ_c can be defined there uniquely. This domain includes the critical value. In particular, the external angle of the critical value is defined uniquely. Douady and Hubbard have shown that the preferred Riemann map $\Phi_{\mathcal{M}}$ of the exterior of the Mandelbrot set is given by $\Phi_{\mathcal{M}}(c) = \varphi_c(c)$.

For disconnected Julia sets, the map φ_c defines dynamic rays at sufficiently large potentials. If a dynamic ray at angle ϑ is defined for potentials greater than $t > 0$, then one can pull back by the dynamics and obtain the dynamic rays at angles $\vartheta/2$ and $(\vartheta + 1)/2$ down to potential $t/2$, except if the ray at angle ϑ contains the critical value. In the latter case, the two pull-back rays will bounce into the critical point and the pull-back is no longer possible uniquely. This phenomenon has been studied by Goldberg and Milnor in the appendix of **[GM]**. Conversely, a dynamic ray at angle ϑ can be extended down to the potential $t > 0$ provided its image ray at angle 2ϑ can be extended down to the potential $2t$ and does not contain the critical value, or if the ray at angle 4ϑ can be extended down to the potential $4t$ without containing the critical value or its image, etc.. The ray can be defined for all potentials in $(0, \infty)$ if the external angle of the critical value is different from $2^k\vartheta$ for all $k = 1, 2, 3, \ldots$ This is the general situation, and in this case, the dynamic ray is known to land at a unique point of the Julia set, whether or not the angle ϑ is rational.

We rephrase these facts in a form which we will have many opportunities to use: if a parameter $c \notin \mathcal{M}$ has external angle ϑ, then the dynamic ray at angle ϑ for the parameter c will contain the critical value. If the angle ϑ is periodic, then this ray cannot possibly land: the ray must bounce into an inverse image of the critical point at a finite positive potential. The main focus of Sections 3 and 4 will be to transfer the landing properties of dynamic rays at rational angles into landing properties of parameter rays at rational angles: as so often in complex dynamics, the general strategy is "to plow in the dynamical plane and then to harvest in parameter space", as Douady phrased it.

When a periodic ray lands at a periodic point, the periods need not be equal: it is possible that the period of the ray is a proper multiple of the period of the point it is landing at. We will therefore distinguish *ray periods* and *orbit periods*. If only one ray lands at every periodic point on the orbit, then both periods are equal; in general, there is a relation between these periods and the number of rays landing at each point on the orbit; see Lemma 2.4. For our purposes, periodic orbits will be most interesting if at least two rays land at each of its points. Such periodic orbits have a distinguished point and two distinguished dynamic rays landing at this point; these play a prominent role in all the symbolic descriptions of the Mandelbrot set. Following the terminology of Milnor **[M2]**, we will call the distinguished point and rays the *characteristic periodic point of the orbit* and the *characteristic rays* (see below), and the corresponding external angles will be the two *characteristic angles* of the orbit. In Thurston's fundamental preprint **[T]**, the two characteristic rays and their common landing point are the "minor leaf" of a "lamination". We will not use or describe his notation here, but we note that it is very close in spirit to this article.

For our purposes, it will be sufficient to define characteristic points and rays only for parabolic periodic orbits.

***Definition 2.1* (Characteristic Components, Points and Rays).** — For a quadratic polynomial with a parabolic orbit, the unique Fatou component containing the critical value will be called the *characteristic Fatou component*; the only parabolic periodic point on its boundary will be the *characteristic periodic point* of the parabolic orbit. It is the landing point of at least two dynamic rays, and the two of them closest to the critical value on either side will be the *characteristic rays*.

The fact that every parabolic periodic point is the landing point of at least two dynamic rays will be shown after Lemma 3.6. Lemma 2.4 will describe the characteristic rays dynamically.

With hesitation, we use the term "Misiurewicz point" for a parameter c for which the critical point or, equivalently, the critical value, is (strictly) preperiodic. This terminology has been introduced long ago, but it is only a very special case of what Misiurewicz was investigating. In real dynamics, the term is used in a wider meaning. We have not been successful in finding an adequate substitution term and invite the reader for suggestions.

In this section, we provide two lemmas which are the engine of our proof: the first one is of analytical nature; it is a slight generalization of Lemma B.1 in Goldberg and Milnor **[GM]**, guaranteeing stability in the Julia set at repelling (pre)periodic points. The second lemma will make counting possible by estimating the number of parabolic parameters with given ray periods.

***Lemma 2.2* (Stability of Repelling Orbits).** — *Suppose that, for some parameter $c_0 \in \mathbb{C}$ (not necessarily in the Mandelbrot set), there is a repelling periodic point z_0 at which some periodic dynamic ray at angle ϑ lands. Then, for parameters c sufficiently close to c_0, the periodic point z_0 can be continued analytically as a function $z(c)$ and the dynamic ray at angle ϑ in the dynamic plane of c lands at $z(c)$. Moreover, the dynamic ray and its landing point form a closed set which is canonically homeomorphic to $[0, \infty]$ via potentials, and this parametrized ray depends continuously on the parameter.*

When $z(c)$ is continued analytically along some curve in parameter space along which the orbit remains repelling, then $z(c)$ can lose its dynamic ray at angle ϑ only on a parameter ray at some angle $2^k\vartheta$ for some integer $k \geq 0$, or else at any parabolic parameter; in the latter case, the dynamic ray at angle ϑ must then land at the parabolic orbit.

If z_0 is repelling and preperiodic, the analogous statement holds provided that neither the point z_0 nor any point on its forward orbit is the critical point.

Proof. — We first assume that z_0 is a periodic point. By the implicit function theorem, z_0 can be continued analytically as a function $z(c)$ in a neighborhood of c_0; the multiplier $\lambda(c)$ will also depend analytically on c so that the cycle is repelling sufficiently close to c_0. Let V be such a neighborhood of c_0 and denote the period of z_0 by n. Then for every $c \in V$ there exists a local branch g_c of the inverse map of $p_c^{\circ n}$ fixing $z(c)$. There is a neighborhood U of z_0 such that g_{c_0} maps $\overline{U}$ into U, and possibly by shrinking V, we may assume that all g_c have the same property for $c \in V$. Under iteration of g_c, any point in U then converges to $z(c)$. Let $t > 0$ be a potential such that, for the parameter c_0, the set U contains all the points of the dynamic ϑ-ray at potentials t and below, including the landing point.

Now we distinguish two cases, according to whether or not $c_0 \in \mathcal{M}$. If $c_0 \notin \mathcal{M}$, then the external angle of the parameter c_0 is well-defined and different from the finitely many angles $2^k\vartheta$ for $k = 1, 2, 3, \ldots$ because the dynamic ray at angle ϑ lands. If V is small enough so that all points in V are outside $\mathcal{M}$ and have their external angles different from all the $2^k\vartheta$, then for every $c \in V$, the dynamic ray at angle ϑ lands, and the point at potential t depends analytically on the parameter. It will therefore be contained in U for sufficiently small perturbations and thus converge to $z(c)$ under iteration of g_c, so the landing point of the ray is $z(c)$.

However, if $c_0 \in \mathcal{M}$, then we may assume V small enough so that all its points have potentials less than $t/2$ (with respect to the potential function of the Mandelbrot set). In the corresponding dynamic planes, the critical values then have potentials less than $t/2$, so every dynamic ray exists and depends analytically on the parameter for potentials greater than $t/2$ (the construction of the Böttcher coordinates around ∞ is analytic in the parameter where they exist). By shrinking V, we may then assume that for all $c \in V$, the segment between potentials $t/2$ and t in the dynamic ray at angle ϑ is contained in U. Iterating the map g_c, it follows that the dynamic ray at angle ϑ lands at $z(c)$. In both cases, rays and landing points depend continuously on the parameter, including the parametrization by potentials.

Suppose that $z(c)$ is continued analytically along some curve in parameter space along which the orbit remains repelling, and let $\widetilde{c}$ be the first point on which the dynamic ray at angle ϑ no longer lands at $z(c)$. For the parameter $\widetilde{c}$, the dynamic ray at angle ϑ then cannot land at any repelling periodic point because this point would keep the ray under perturbations. Therefore, the dynamic ray at angle ϑ must either land at a parabolic orbit, or it must fail to land entirely. The latter case happens if and only if the forward orbit of the dynamic ϑ-ray contains the critical value, so the parameter $\widetilde{c}$ is on a parameter ray as specified.

The statement about preperiodic points follows by taking inverse images and is straightforward, except if z_0 or any point on its forward orbit are the critical point. However, if some preperiodic dynamic ray lands at the critical value, then a small perturbation may bring the critical value onto this dynamic ray, and the inverse rays will bounce into the critical point (after that, both branches will land, and the landing points are two branches of an analytic function). □

***Lemma 2.3* (Counting Parabolic Orbits).** — *For every positive integer n, the number of parabolic parameters in $\mathbb{C}$ having a parabolic orbit of exact ray period n is at most half the number of periodic angles in $[0,1]$ having exact period n under doubling modulo 1.*

Proof. — We can calculate the exact number of periodic angles. If an angle $\vartheta \in [0,1]$ satisfies $2^n\vartheta \equiv \vartheta$ modulo 1, then we can write $\vartheta = a/(2^n-1)$ for some integer a, and there are 2^n such angles in $[0,1]$. Only a subset of these angles has exact period n: denoting the number of such angles by s'_n, we have $\sum_{k|n} s'_k = 2^n$, which allows to determine the s'_n recursively or via the Möbius inversion formula. We have $s'_1 = 2$, and all the s'_n are easily seen to be even. In the sequel, we will work with the integers $s_n := s'_n/2$. The first few terms of the sequence (s_n), starting with s_1, are $1, 1, 3, 6, 15, 27, 63, \ldots$ The specified number of periodic angles in $[0,1]$ is then exactly $2s_n$.

We consider the curve

$$\{(c,z) \in \mathbb{C}^2 : p_c^{\circ n}(z) = z\}$$

consisting of points z which are periodic under p_c with period dividing n. It factors as a product $\prod_{k|n} Q_k(c,z)$ according to exact periods. (The curves Q_k have been shown to be irreducible by Bousch **[Bo]** and by Lau and Schleicher **[LS]**, a fact we will not use.) For $|c| > 2$, the filled-in Julia set of p_c is a Cantor set containing all the periodic points. For $|c| > 4$, it is easy to verify that points z with $|z| > |c|^{1/2}+1$ escape to ∞, and so do points with $|z| < |c|^{1/2}-1$. Periodic points therefore satisfy $|z| = |c|^{1/2}(1+o(1))$ as $c \to \infty$. The multiplier of a periodic orbit of exact period n is the product of the periodic points on the orbit multiplied by 2^n, so it grows like $|4c|^{n/2}(1+o(1))$.

For any parameter c, the number of points which are fixed under the n-th iterate is obviously equal to 2^n, counting multiplicities. These points have exact periods dividing n, so the number of periodic points of exact period n equals $2s_n$ by the same recursion formula as above. These periodic points are grouped in orbits, so the number of orbits is $2s_n/n$ (which implies that $2s_n$ is divisible by n). For bounded parameters c, the periodic points and thus the multipliers are bounded; since there are $2s_n/n$ orbits, the multipliers of which are analytic and behave like $|c|^{n/2}$ near infinity, and since every orbit contains n points, it follows that sufficiently large multipliers are assumed exactly $(2s_n/n)(n/2)n = ns_n$ times on Q_n (we do not have to count multiplicities here because multiple orbits always have multiplier $+1$). Consider the multiplier map on Q_n which assigns to every point (c,z) the multiplier $(\partial/\partial z)p_c^{\circ n}(z)$. It is a proper map and thus has a mapping degree, so (counting multiplicities) every multiplier in $\mathbb{C}$ is assumed equally often, including the value $+1$. The number of points (c,z) having multiplier $+1$ therefore equals ns_n, counting multiplicities. Projecting onto the c-coordinate and ignoring multiplicities, a factor n is lost because points on the same orbit project onto the same parameter, and we obtain an upper bound of s_n for the number of parameters. (In fact, it is not too hard to show at this point that s_n provides an exact count [**M2**]. We will show this in Corollary 3.4 by a global counting argument.)

Consider a parabolic orbit of exact period k and multiplier $\mu = e^{2\pi i p/q}$ with $(p,q) = 1$. Then the exact ray period is $qk =: n$, and qk is also the smallest period such that, when interpreting the orbit as an orbit of this period, the multiplier becomes $+1$. Therefore, the periodic points on this orbit are on Q_n, and the number of parabolic parameters having exact ray period n therefore is at most s_n. □

A more detailed account of such counting arguments can be found in Section 5 of Milnor [**M2**].

The following standard lemma is folklore and at the base of every description of quadratic iteration theory. Our proof follows Milnor [**M2**]; compare also Thurston [**T**, Theorem II.5.3 case b) i) a)]. We do not assume that the Julia set has any particular property; it need not even be connected.

Lemma 2.4 **(Permutation of Rays).** — *If more than two periodic rays land at a periodic point, or if the orbit period is different from the ray period, then the first return map of the point permutes the rays transitively.*

Proof. — Denote the orbit period by k and the ray period by n. Since a periodic orbit has periodic rays landing only if the orbit is repelling or parabolic, the first return map of any of its periodic points is a local homeomorphism and permutes the rays landing there in such a way that their periods are all equal, and the number of rays landing at each point of the orbit is a constant s, say. If $s = 1$, then orbit period and ray period are equal. If $s = 2$, then either ray and orbit periods are equal, or the first return map of any point has no choice but to transitively permute the two rays landing at this point. We may hence assume $s \geq 3$. Then the s rays landing at any one of these periodic points separate the dynamic plane into s sectors. Every sector is bounded by two dynamic rays, so it has associated a width: the external angles of the two rays cut $\mathbb{S}^1$ into two open intervals, exactly one of which does not contain external angles of rays landing at the same point. The *width* of the sector will be the length of this interval (normalized so that the total length of $\mathbb{S}^1$ is 1).

Since the dynamics of the first return map is a local homeomorphism near the periodic point, every sector is periodic, and so is the sequence of the corresponding widths. More precisely, we will justify the following observations below: if a sector does not contain the critical point, then it maps homeomorphically onto its image sector (based at the image of its landing point), and the width of the sector doubles. However, if the sector does contain the critical point, then the sector maps in a two-to-one fashion onto the image sector, and it covers the remaining dynamic plane once. In this case, the width of the sector will decrease under this mapping, and the image sector contains the critical value. To justify these statements, first note that the rays bounding any sector are mapped to the rays bounding the image sector. Looking at external angles within the sector, it follows that either the sector maps forward homeomorphically, or it covers the entire complex plane once and the image sector twice. The latter must happen for the sector containing the critical point. Since all the sectors at any periodic point combined exactly cover the complex plane twice when mapped forward, all the other sectors must map homeomorphically onto the image sectors. We also see that among all the sectors based at any point, the sector containing the critical point must have width greater than $1/2$, and all the other sectors then have widths less than $1/2$ (the critical point cannot be on a sector boundary: if it is on a periodic dynamic ray, then this ray cannot land, and if it is on a periodic point, then this point is superattracting). The width of any sector doubles under the map if it does not contain the critical point; since the sum of the widths of all the sectors based at any point is 1, the width of the critical sector must decrease.

For each orbit of sectors, there must be at least one sector with minimal width. It must contain the critical value (or it would be the image of a sector with half the width), and it cannot contain the critical point (or its image sector would have smaller width). Therefore, all the shortest sectors of the various cycles of sectors must be bounded by pairs of rays separating the critical point from the critical value, and these sectors are all nested. Among them, there is one innermost sector S_1 based at some point z_1 of the periodic orbit. This sector S_1 cannot contain another point from the orbit of z_1: if there was such a point z', there would have to be a sector based at z' which was shorter than all the shortest sectors at points on the orbit of z_1, and this is obviously absurd.

If there is an orbit of sectors not involving S_1, then any shortest sector on this orbit must contain the critical value and thus S_1, but it cannot contain the critical point. This sector must then contain all sectors at z_1 except the one containing the critical point. The representative of this orbit of sectors at z_1 must then be the unique sector containing the critical point. Any cycle of sectors has then only two choices for its representative at z_1: the sector containing the critical point or the critical value. If there is more than one cycle, then it follows that there are just two cycles, and each of them has exactly one representative at each point of the periodic orbit, so $s = 2$ in contradiction to our assumption. □

Remark. — This lemma is at the heart of the general definition of characteristic rays: the main part of the proof works when at least two rays land at each of the periodic points, and it shows that there is a unique sector of minimal width containing the critical value. The rays bounding this sector are called the characteristic rays. For the special case of a parabolic orbit, this definition agrees with the one we have given above.

3. Periodic Rays

In this section, we will be concerned with parameter rays at periodic angles. The proof of the following weak form of the theorem is due to Goldberg, Milnor, Douady, and Hubbard; see [**GM**, Theorem C.7].

***Proposition 3.1* (Periodic Parameter Rays Land).** — *Every parameter ray at a periodic angle ϑ lands at a parabolic parameter c_0. In the dynamic plane of c_0, the dynamic ray at angle ϑ lands at the parabolic orbit.*

Proof. — Let c_0 be a point in the limit set of the parameter ray at angle ϑ and let n be the exact period of ϑ. In the dynamic plane of c_0, the dynamic ray at angle ϑ must land at a repelling or parabolic periodic point z of ray period n; see [**M1**, Theorem 18.1]. If z was repelling, Lemma 2.2 would imply that for parameters c sufficiently close to c_0, the dynamic ray at angle ϑ in the dynamic plane of c would land at a repelling periodic point $z(c)$, so it could not bounce off any precritical point. However, when c is on the parameter ray at angle ϑ, then the dynamic ray at angle ϑ must bounce off some precritical point, even infinitely often.

Therefore, c_0 is parabolic, and within its dynamics, the dynamic ray at angle ϑ lands at the parabolic orbit. Since limit sets are connected but parabolic parameters of given ray period form a finite set by Lemma 2.3, the parameter ray at angle ϑ lands and the statements follow. □

This proves half of the first assertion in Theorem 1.1. The remainder of the first and the second assertion will be shown in several steps. We want to show that at a parabolic parameter c_0, those two parameter rays land which have the same external angles as the two characteristic rays of the critical value Fatou component, and no other rational ray lands there. The first statement is usually shown using Ecalle cylinders. It turns out that it is much easier to show that some ray does *not* land at a given point, rather than to show where it does land. The idea in this paper will be to exclude all the wrong rays from landing at given parabolic parameters, using partitions in the dynamic and parameter planes. Using that the rays must land somewhere, a global counting argument will then prove the theorem.

Let c be a parabolic parameter and let Θ_c be the set of periodic angles ϑ such that the parameter ray at angle ϑ lands at c. A priori, it might be empty; if it is not, then all the angles in Θ_c have the same period by Proposition 3.1. We will prove the following two results later in this section.

***Proposition 3.2* (Necessary Condition).** — *If an angle ϑ is in Θ_c, then the dynamic ray at angle ϑ lands at the characteristic point of the parabolic orbit in the dynamic plane of c.*

***Proposition 3.3* (At Most Two Rays).** — *Suppose the set Θ_c contains more than one angle. Let n be the common period of these angles and suppose that all parameter ray pairs of periods $2, 3, \ldots, n-1$ land in pairs. Then Θ_c consists of exactly those two angles which are the characteristic angles of the parabolic orbit in the dynamic plane of c.*

These two propositions allow to prove the half of the theorem dealing with periodic rays; we will deal with the preperiodic half in the next section.

Proof of Theorem 1.1 (periodic case). — We will use induction on the period. For period $n = 1$, there are only two angles 0 and 1 which both describe the same parameter ray. This ray runs along the positive real axis and lands at the unique parabolic parameter $c = 1/4$ with ray period 1.

For any period $n \geq 2$, the number of parabolic parameters of any given ray period is at most half the number of parameter rays at periodic angles of the same period by Lemma 2.3. Since every ray lands at such a parabolic parameter by Proposition 3.1, and at most two rays land at any such point by Proposition 3.3 (using the inductive hypothesis), it follows that exactly two rays land at every parabolic point, and Proposition 3.3 says which ones these are. It also follows that the number of parabolic parameters of any given period is largest possible as allowed by Lemma 2.3. □

***Corollary 3.4* (Counting Parabolic Orbits Exactly).** — *Let s_k be the number of parameters having a parabolic orbit of exact ray period k. These numbers satisfy the recursive relation $\sum_{k|n} s_k = 2^{n-1}$, which determines them uniquely.* □

It remains to prove the two propositions. In both of them, we have to exclude that certain rays land at given parabolic parameters. We do that using appropriate partitions: first in the dynamic plane, then in parameter space. We start by discussing the topology of parabolic quadratic Julia sets and define a variant of the Hubbard tree on them. Hubbard trees have been introduced by Douady and Hubbard in **[DH1]** for postcritically finite polynomials. We will be interested in combinatorial statements about combinatorially described Julia sets, so these results could be derived in purely combinatorial terms. However, it will be more convenient to use topological properties of the Julia sets in the parabolic case, in particular that they are pathwise connected (which follows from local connectivity). This was originally proved by Douady and Hubbard **[DH1]**; proofs can also be found in Carleson and Gamelin **[CG]** and in Tan and Yin **[TY]**.

In a quadratic polynomial with a parabolic orbit, let z be any point within the filled-in Julia set and let U be a bounded Fatou component. We then define a *projection* of

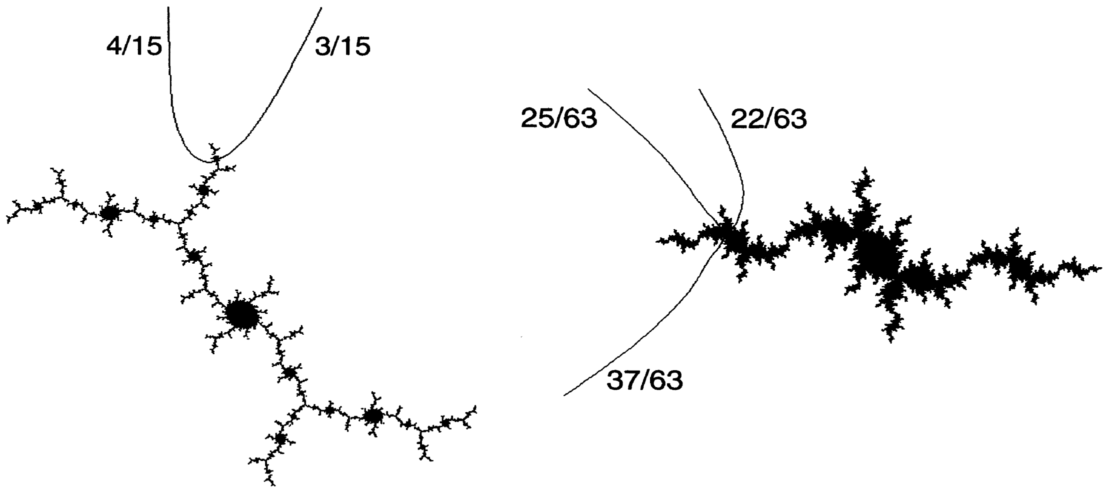


FIGURE 2. *Illustration of the theorem in the periodic case. The polynomials at the landing points of the parameter rays at angles* $3/15$ *and* $4/15$ *(left) and at angles* $22/63$ *and* $25/63$ *(right) are shown. In both pictures, the rays landing at the characteristic points are drawn. For the corresponding parameter rays, see Figure 1.*

z onto $\overline{U}$ as follows: If $z \in \overline{U}$, then the projection of z onto $\overline{U}$ is z itself. Otherwise, consider any path within the filled-in Julia set connecting z to an interior point of U (such a path exists because the filled-in Julia set is pathwise connected); then the projection of z onto $\overline{U}$ is the first point where this path intersects ∂U. There may be many such paths, but the projection is still well-defined: take any two paths from z to the interior of U and connect their endpoints within U. If the paths are different, they will bound some subset of $\mathbb{C}$, which must be in K because K is full. If the paths reach ∂U in different points, then these paths enclose part of the boundary of U, but the boundary of any Fatou component is always in the boundary of the filled-in Julia set. This contradiction shows that the projection is well-defined. Every parabolic periodic point is on the boundary of at least one periodic Fatou component, so the projection in this case is just the identity.

***Lemma 3.5* (Projection Onto Periodic Fatou Components).** — *In a quadratic polynomial with a parabolic orbit, the projections of all the parabolic periodic points onto the Fatou component containing the critical value take images in the same point, which is the characteristic point of the parabolic periodic orbit. Projections of the parabolic periodic points onto any other bounded Fatou component take images in at most two boundary points, which are periodic or preperiodic points on the parabolic orbit.*

Proof. — Let n be the period of the periodic Fatou components and number them $U_0, U_1, \dots, U_{n-1}, U_n = U_0$ in the order of the dynamics, so that $U_0 = U_n$ contains the critical point. Let a_k be the number of different images that the projections of all the parabolic periodic points onto $\overline{U}_k$ have, for $k = 0, 1, \dots, n$ (with $a_0 = a_n$). We first show that $a_{k+1} \geq a_k$ for $k = 1, 2, \dots, n-1$.

Let z be a parabolic periodic point and let $\pi(z)$ be its projection onto $\overline{U}_k$. We claim that $p(\pi(z))$ is the projection onto $\overline{U}_{k+1}$ of either $p(z)$ or the parabolic point on the boundary of the Fatou component containing the critical value (i.e. the characteristic point on the parabolic orbit). Indeed, if the path between z and $\pi(z)$ maps forward homeomorphically under p, then $\pi(p(z)) = p(\pi(z))$. If it does not, then the path must intersect the component containing the critical point, and $\pi(z) = \pi(0)$. But then $p(\pi(z))$ is the projection of the characteristic point on the parabolic orbit. Therefore, for $k \in \{0, 1, 2, \dots, n-1\}$, all the a_k image points of the projections of parabolic periodic points onto $\overline{U}_k$ will be mapped under p to image points of the projection onto $\overline{U}_{k+1}$. Since for $k \neq 0$, the polynomial p maps $\overline{U}_k$ homeomorphically onto $\overline{U}_{k+1}$, we get $a_n \geq a_{n-1} \geq \cdots \geq a_2 \geq a_1$. Similarly, since p maps $\overline{U}_0$ in a two-to-one fashion onto $\overline{U}_1$, we have $a_1 \geq a_0/2$.

Now we connect the parabolic periodic points by a tree: first, there is a path between the critical point and the critical value, and all the other parabolic periodic points which are not on this path can be connected, one by one, to the subtree which has been constructed thus far. We can require that every path which we are adding intersects the boundary of any bounded Fatou component in the least number of points (at most two). After finitely many steps, all the parabolic periodic points are connected by a finite tree, and all the endpoints of this tree are parabolic periodic points. It is not hard to check that this tree intersects the boundary of any bounded Fatou component exactly in the image points of the projections of the parabolic periodic points.

We now claim that there is a periodic Fatou component whose closure does not disconnect the tree. Indeed, any component which does disconnect the tree has at least one parabolic periodic point and thus at least one periodic Fatou component in each connected component of the complement. Pick one connected component, and within it pick a periodic Fatou component that is "closest" to the removed one (in the sense that the path between these two components does not contain further periodic Fatou components). Remove this component and continue; this process can be continued until we arrive at a component which does not disconnect the tree. Let $U_{\widetilde{k}}$ be such a component.

It follows that $a_{\widetilde{k}} \leq 2$: all the parabolic periodic points which are not on the boundary of $U_{\widetilde{k}}$ must project to the same boundary point, say b, which may or may not be the parabolic periodic point on the boundary of $U_{\widetilde{k}}$. We will now show that it will be, so $a_{\widetilde{k}} = 1$.

Since b is the image of a projection onto $\overline{U}_{\tilde{k}}$ of some parabolic periodic point which is not in $\overline{U}_{\tilde{k}}$, it follows from the argument above that $p(b)$ is the image of a projection onto $\overline{U}_{\tilde{k}+1}$ of some parabolic periodic point which is not in $\overline{U}_{\tilde{k}+1}$. But since b is the only boundary point of $\overline{U}_{\tilde{k}}$ with this property, it follows that b is fixed under the first return map of this Fatou component. The point b must then be the unique parabolic periodic point on the boundary of $\overline{U}_{\tilde{k}}$, and we have $a_{\tilde{k}} = 1$.

Since $a_1 \leq a_2 \leq a_n \leq 2a_1$, it follows in particular that $a_1 = 1$ and all $a_k \leq 2$, and all projections onto the Fatou component containing the critical value take values in the same point, which is the characteristic point of the parabolic orbit. The remaining claims follow. □

Remark. — The tree just constructed is similar to the *Hubbard tree* introduced in **[DH1]** for postcritically finite polynomials. An important difference is that our tree does not connect the critical orbit. Moreover, Hubbard trees in **[DH1]** are specified uniquely, while our trees still involve the choice of how to traverse bounded Fatou components. We will suggest a preferred tree below.

However, some properties are independent of the choice of the tree. Assume that two simple curves γ_1 and γ_2 within the filled-in Julia set connect the same two points z_1 and z_2, such that a point w is on one of the curves but not on the other. Then w is on the closure of a bounded Fatou component because the region which is enclosed by the two curves must be in the filled-in Julia set. The tree intersects the boundary of any bounded Fatou component in at most two points which are projection images and thus well-defined. Therefore, the choice for the curves and thus for the tree is only in the interior of bounded Fatou components.

For any point w in the Julia set (not in a bounded Fatou component), it follows that the number of branches of the tree (i.e. the number of components the point disconnects the tree into) is independent of the choice of the tree. Similarly, the number of branches is "almost" non-decreasing under the dynamics, so that $p(w)$ has at least as many branches as w: all the different branches at w will yield different branches at $p(w)$, except if w is on the boundary of the Fatou component containing the critical point. At such boundary points, only the branch leading into the critical Fatou component can get lost. (However, it does happen that $p(w)$ has extra branches in the tree.) It follows that the characteristic point on the parabolic orbit has at most one branch on any tree, and all the other parabolic periodic points can have up to two branches.

A branch point of a tree is a point w which disconnects the tree into at least three complementary components.

***Lemma 3.6* (Branch Points of Tree)**. — *Branch points of the tree between parabolic periodic points are periodic or preperiodic points on repelling orbits.*

Proof. — Branch points are never on the parabolic orbit, as we have just seen. Therefore, the image of a branch point is always a branch point with at least as many branches. Since there are only finitely many branch points, every branch point is periodic or preperiodic and hence on a repelling orbit. □

We can now proceed to select a preferred tree, which we will call a *(parabolic) Hubbard tree.* We only have to specify how it will traverse bounded Fatou components. In fact, since every bounded Fatou component will eventually map homeomorphically onto the critical Fatou component, we only have to specify how the tree has to traverse this component; for the remaining components, we can pull back.

Let U be the critical Fatou component and let w be the parabolic periodic point on its boundary. First we want to connect the critical point in U to w by a simple curve which is forward invariant under the dynamics. We will use *Fatou coordinates* for the attracting petal of the dynamics [**M1**, Section 7]. In these coordinates, the dynamics is simply addition of $+1$, and our curve will just be a horizontal straight line connecting the critical orbit. This curve can be extended up to the critical point. The other point on the boundary of the critical Fatou component which we have to connect is $-w$, and we use the symmetric curve. With this choice, we have specified a preferred tree which is invariant under the dynamics, except that the image of the tree connects the characteristic periodic point on the parabolic orbit to the critical value. Removing this curve segment from the image tree, we obtain the same tree as before.

It is well known that, if a repelling or parabolic periodic point disconnects the Julia set into several parts, then this point is the landing point of as many dynamic rays as it disconnects the Julia set into. It follows that any branch point of the Hubbard tree has dynamic rays landing between any two branches; any periodic point on the interior of the tree is the landing point of at least two dynamic rays separating the tree. It now follows that the characteristic point on the parabolic orbit, and thus every parabolic periodic point, is the landing point of at least two dynamic rays. The two characteristic rays of the parabolic orbit are the two rays landing at the characteristic point of the orbit and closest possible to the critical value on either side. A different description of the characteristic rays has been given in Lemma 2.4.

***Lemma 3.7* (Orbit Separation Lemma).** — *Any two parabolic periodic points of a quadratic polynomial can be separated by two (pre)periodic dynamic rays landing at a common repelling (pre)periodic point.*

Proof. — It suffices to prove the lemma when one of the two parabolic periodic points is the characteristic point of the orbit; this is also the only case we will need here. Let z be this characteristic point and let z' be a different parabolic periodic point. Consider the tree of the polynomial as constructed above. It contains a unique path connecting z and z'. We may assume that this path does not traverse a periodic

Fatou component except at its ends; if it does, we replace z' by the parabolic periodic point on that Fatou component. Similarly, we may assume that the path does not traverse another parabolic periodic point. If the path from z to z' contains a branch point of the tree, then by Lemma 3.6, this branch point is periodic or preperiodic and repelling, and it is therefore the landing point of rational dynamic rays separating the parabolic orbit as claimed.

If the Hubbard tree does not have a branch point between z and z', then it takes a finite number k of iterations to map z' for the first time onto z. Denoting the path from z to z' by γ, then the k-th iterate of γ must traverse itself and possibly more in an orientation reversing way: denoting the k-th image of z by z'', then the image curve connects z and z''; near z, it must start along the end of γ because z is an endpoint of the Hubbard tree, and it cannot branch off because we had assumed no branch point of the tree to be on γ. There must be a unique point z_f in the interior of γ which is fixed under the k-th image of γ. This point is a repelling periodic point, and it is the landing point of two dynamic rays with the desired separation properties. □

Now we can prove Proposition 3.2.

Proof of Proposition 3.2. — In Proposition 3.1, we have shown that Θ_c can contain only angles ϑ of dynamic rays landing at the parabolic cycle in the dynamic plane of c. By the Orbit Separation Lemma 3.7, all the rays not landing at the characteristic point of the parabolic orbit are separated from the critical value by a partition formed by two dynamic rays landing at a common repelling (pre)periodic point. This partition is stable in a neighborhood in parameter space by Lemma 2.2. But the parameter c being a limit point of the parameter ray at angle ϑ means that, for parameters arbitrarily close to c, the critical value is on the dynamic ray at angle ϑ. □

The set Θ_c of external angles of the parabolic parameter c can thus contain only such periodic angles which are external angles of the characteristic periodic point of the parabolic orbit in the dynamical plane of c. If there are more than two such angles, we want to exclude all those which are not characteristic. This is evidently impossible by a partition argument in the dynamic plane. In order to prove Proposition 3.3, we will use a partition of parameter space; for that, we have to look more closely at parameter space and incorporate some symbolic dynamics using kneading sequences. The partition of parameter space according to kneading sequences and, more geometrically, into internal addresses is of interest in its own right (see below) and has been investigated by Lau and Schleicher **[LS]**; related ideas can be found in Thurston **[T]**, Penrose **[Pr1]** and **[Pr2]** and in a series of papers by Bandt and Keller (see **[Ke1]**, **[Ke2]** and the references therein). This partition will also be helpful in the next section, establishing landing properties of preperiodic parameter rays.

***Definition 3.8* (Kneading Sequence).** — To an angle $\vartheta \in \mathbb{S}^1$, we associate its *kneading sequence* as follows: divide $\mathbb{S}^1$ into two parts at $\vartheta/2$ and $(\vartheta+1)/2$ (the two inverse images of ϑ under angle doubling); the open part containing the angle 0 is labeled 0, the other open part is labeled 1 and the boundary gets the label $\star$. The kneading sequence of the angle ϑ is the sequence of labels corresponding to the angles ϑ, 2ϑ, 4ϑ, 8ϑ, ...

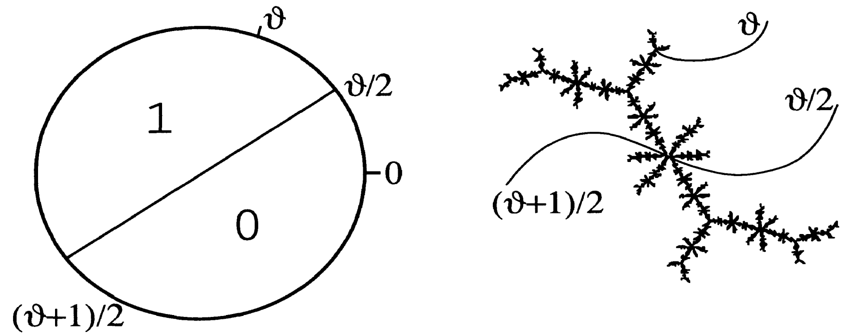


FIGURE 3. *Left: the partition used in the definition of the kneading sequence. Right: a corresponding partition of the dynamic plane by dynamic rays, shown here for the example of a Misiurewicz polynomial.*

It is easy to check that, for $\vartheta \neq 0$, the first position always equals 1. If ϑ is periodic of period n, then its kneading sequence obviously has the same property and the symbol $\star$ appears exactly once within this period (at the last position). The symbol $\star$ occurs only for periodic angles. However, it may happen that an irrational angle has a periodic kneading sequence (see e.g. [**LS**]). As the angle ϑ varies, the entry of the kneading sequence at any position n changes exactly at those values of ϑ for which $2^{n-1}\vartheta$ is on the boundary of the partition, i.e. where the kneading sequence has the entry $\star$. This happens if and only if the angle ϑ is periodic, and its exact period is n or divides n.

Another useful property which will be needed in Section 4 is that the pointwise limits $\mathbb{K}_-(\vartheta) := \lim_{\vartheta' \nearrow \vartheta} \mathbb{K}(\vartheta')$ and $\mathbb{K}_+(\vartheta) := \lim_{\vartheta' \searrow \vartheta} \mathbb{K}(\vartheta')$ exist for every ϑ. If ϑ is periodic, then $\mathbb{K}_\pm(\vartheta)$ is also periodic with the same period (but its exact period may be smaller). Both limiting kneading sequences coincide with $\mathbb{K}(\vartheta)$ everywhere, except that all the $\star$-symbols are replaced by 0 in one of the two sequences and by 1 in the other. The reason is simple: if ϑ' is very close to ϑ, then the orbits under doubling, as well as the partitions in the kneading sequences are close to each other, and any symbol 0 or 1 at any finite position will be unchanged provided ϑ' is close enough to ϑ. However, if the period of ϑ is n so that $2^{n-1}\vartheta$ is on the boundary of the partition in the kneading sequence, then $2^{n-1}\vartheta'$ will barely miss the boundary in its own partition, and the $\star$ will turn into a 0 or 1. As long as the orbit of ϑ' is close to the orbit of ϑ, all the symbols $\star$ will be replaced by the same symbol.

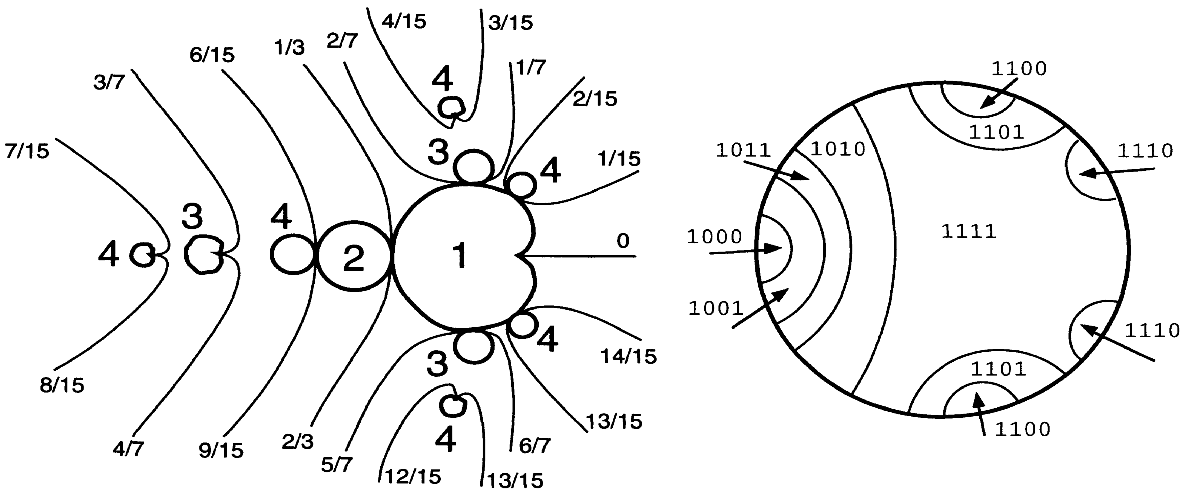


FIGURE 4. *Left: the partition $\mathcal{P}_n$ used in the proof of Proposition 3.3, for $n = 4$. The corresponding hyperbolic components are drawn in for clarity and do not form part of the partition. Right: a corresponding symbolic picture, showing how the partition yields a parameter space of initial segments of kneading sequences. The same pairs of rays are drawn in as on the left hand side, but the angles are unlabeled for lack of space.*

Proposition 3.1 asserts in particular that all the periodic parameter rays landing at the same parameter have equal period. All the rays of period at most $n-1$ divide the plane into finitely many pieces. We denote this partition by $\mathcal{P}_{n-1}$; it is illustrated in Figure 4. Parabolic parameters of ray period n and parameter rays of period n have no point in common with the boundary of this partition.

***Lemma 3.9* (Kneading Sequences in the Partition).** — *Fix any period $n \geq 1$ and suppose that all the parameter rays of periods at most $n-1$ land in pairs. Then all parameter rays in any connected component of $\mathcal{P}_{n-1}$ have the property that the first $n-1$ entries in their kneading sequences coincide and do not contain the symbol $\star$. In particular, rays of period n with different kneading sequences do not land at the same parameter.*

Proof. — The first statement is trivial for two periodic rays which do not have rays of lower periods between them, i.e., for rays from the same "access to infinity" of the connected component in $\mathcal{P}_{n-1}$: the first $n-1$ entries in the kneading sequences are stable for angles within every such access. The claim is interesting only for a connected component with several "accesses to infinity".

The hypothesis of the theorem asserts that parameter rays of periods up to $n-1$ land in pairs. Therefore, whenever two rays at angles ϑ_1, ϑ_2 are in the same connected component of $\mathcal{P}_{n-1}$, the parameter rays of any period $k \leq n-1$ on either side (in

$\mathbb{S}^1$) between these two angles must land in pairs. The number of such rays is thus even, and the k-th entry in the kneading sequence changes an even number of times between 0 and 1. □

Remark. — This lemma allows to interpret $\mathcal{P}_n$ as a parameter space of initial segments of kneading sequences. In Figure 4, the partition is indicated for $n = 4$, together with the initial four symbols of the kneading sequence. The entire parameter space may thus be described by kneading sequences, as noted above. To any parameter $c \in \mathbb{C}$, we may associate a kneading sequence as follows: it is a one-sided infinite sequence of symbols, and the k-th entry is 1 if and only if the parameter is separated from the origin by an even number of parameter ray pairs of periods k or dividing k; if the number of such ray pairs is odd, then the entry is 0, and if the parameter is exactly on such a ray pair, then the entry is $\star$. Calculating the kneading sequence of any point is substantially simplified by the observation that, in order to know the entire kneading sequence at a parameter ray pair of some period n, it suffices to know the first $n-1$ entries in the kneading sequence, so we only have to look at ray pairs of periods up to $n-1$. This leads to the following algorithm: for any point $c \in \mathbb{C}$, find consecutively the parameter ray pairs of lowest periods between the previously used ray pair and the point c. The periods of these ray pairs will form a strictly increasing sequence of integers and allow to reconstruct the kneading sequence, encoding it very efficiently. If we extend this sequence by a single entry 1 in the beginning, we obtain the *internal address* of c. For details, see **[LS]**. In the context of real quadratic polynomials, this internal address is known as the sequence of *cutting times* in the Hofbauer tower.

The figure shows that certain initial segments of kneading sequences appear several times. This can be described and explained precisely and gives rise to certain symmetries of the Mandelbrot set; see **[LS]**.

***Lemma 3.10* (Different Kneading Sequences).** — *Let c be a parabolic parameter and let z_1 be the characteristic periodic point on the parabolic orbit. Among the dynamic rays landing at z_1, only the two characteristic rays can have angles with identical kneading sequences.*

Proof. — Let $\vartheta_1, \vartheta_2, \ldots, \vartheta_s$ be the angles of the dynamic rays landing at z_1. If their number s is 2, then both angles are characteristic, and there is nothing to show. We may hence assume $s \geq 3$. All the rays ϑ_i are periodic of period n, say. By Lemma 2.4, the orbit period of the parabolic orbit is exactly $n/s =: k$. Let z_0 and z_0' be the two (different) immediate inverse images of z_1 such that z_0 is periodic. If any one of the rays $R(\vartheta_i)$ is chosen, its two inverse images, together with any simple path in the critical Fatou component connecting z_0 and z_0', form a partition of the complex plane into two parts. We label these parts again by 0 and 1 so that the dynamic ray at angle 0 is in part 0, and we label the boundary by $\star$. Now the labels of the parts containing the rays $R(\vartheta_i)$, $R(2\vartheta_i)$, $R(4\vartheta_i)$, ... again reflect the kneading sequence of ϑ_i because the partitions are bounded at the same angles.

Above, we have constructed a tree connecting the parabolic orbit. By Lemma 3.6, branch points of this tree are on repelling orbits, so z_0 and z_0' have at most two branches of the tree. One branch always goes into the critical Fatou component to the critical point. For z_0 or z_0', the other branch goes to the critical value which is always in the region labeled 1. The second branch at the other point (z_0' or z_0) must leave in the symmetric direction, so it will always lead into the region labeled 0 (if the second branch at z_0' were to lead into a direction other than the symmetric one, then the common image point z_1 of z_0 and z_0' could not be an endpoint of the tree, which it is by Lemma 3.6). It follows that the entire tree, except the part between z_0 and z_0', is in a subset of the filled-in Julia set whose label does not depend on which of the angles ϑ_i have been used to define the partition and the kneading sequence. For positive integers l, let z_l be the l-th forward image of z_0. If l is not divisible by k, it follows that the label of z_l is independent of ϑ_i; since z_l is the landing point of all the dynamic rays at angles $2^{l-1}\vartheta_i$, it follows that the $l-1$-st entries of the kneading sequences of all the ϑ_i are the same. Therefore, we can restrict our attention to the rays at angles $\vartheta_1', \vartheta_2', \ldots, \vartheta_s'$ landing at z_0, where ϑ_i' is an immediate inverse image of ϑ_i. The first return dynamics among the angles is multiplication by 2^k; for the rays, this must be a cyclic permutation with combinatorial rotation number r/s for some integer r (Lemma 2.4); compare Figure 5.

Depending on which of the rays ϑ_i is used for the kneading sequence, i.e. which of the rays ϑ_i' defines the partition, a given ray ϑ_j' may have label 0 or label 1. In particular, the total number among these rays which are in region 0 may be different. But then the number of symbols 0 within any period of the kneading sequence will be different and the kneading sequences cannot coincide. Two angles among the ϑ_i can thus have the same kneading sequence only if the corresponding partition has equally many rays in the region labeled 0. This leaves only various pairs of angles at symmetric positions around the critical value as candidates to have identical kneading sequences. But it is not hard to verify, looking at the cyclic permutation of the rays ϑ_i', that if two such angles define a partition in which at least one of the ϑ_i' is in region 0, then the two corresponding kneading sequences are different at some position which is a multiple of k. The only two angles with identical kneading sequences are therefore those for which all the ϑ_i' are in region 1 (or on its boundary), so the partition boundary is adjacent to the Fatou component containing the critical point. The angles ϑ_i are hence those for which the dynamic rays land at z_1 adjacent to the critical value, so the corresponding rays are the characteristic rays of the parabolic orbit. They do in fact have identical kneading sequences. □

Proof of Proposition 3.3. — For period $n = 1$, there is nothing to show. For periods $n \geq 2$, we may suppose that all parameter rays of periods up to $n-1$ land in pairs.

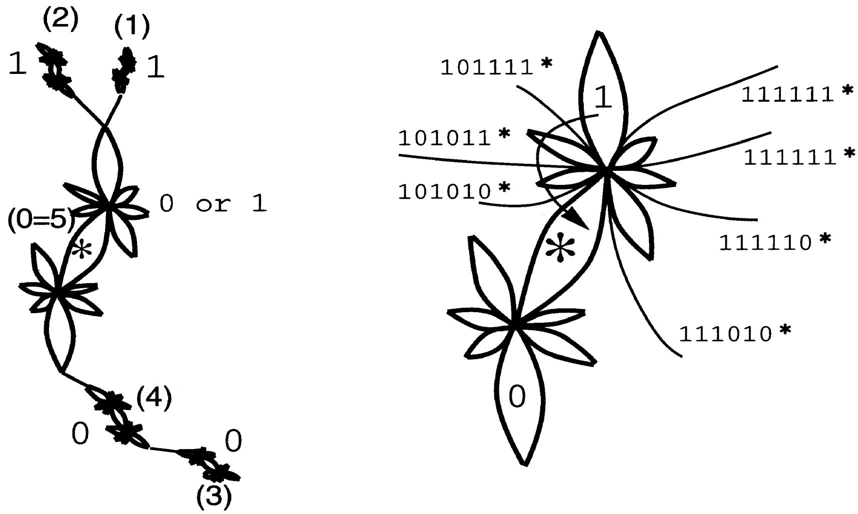


FIGURE 5. *Illustration of the proof of Lemma 3.10. Left: coarse sketch of the entire Julia set; numbers in parentheses describe the parabolic orbit (in this case, of period 5), and symbols 0 or 1 specify the corresponding entries in the kneading sequences of external angles of rays landing at the characteristic periodic point. Right: blow-up near the critical point (center, marked by* $*$*) with the periodic point* z_0*, considered as a fixed point of the first return map. In this case, seven rays land at* z_0 *with combinatorial rotation number* 3/7*. The rays are labeled by the corresponding kneading sequences. The symbols 0 and 1 indicate regions of the Julia set which are always on the same side of the partition, independently of which ray is chosen.*

We have a parabolic parameter c of ray period n and the set Θ_c contains the angles of periodic parameter rays landing at c. All these angles have the same period n and, by Lemma 3.9, identical kneading sequences. Since the corresponding dynamic rays all land at the characteristic point of the parabolic periodic orbit by Proposition 3.2, Lemma 3.10 says that if Θ_c contains more than a single element, it contains the two characteristic angles. □

We have now finished the proof of the periodic part of the theorem, describing which periodic parameter rays land at common points. This prepares the ground for combinatorial descriptions such as Lavaurs' algorithm [**La1**] or internal addresses.

4. Preperiodic Rays

In this section, we will turn to parameter rays at preperiodic angles and show at which Misiurewicz points they land. We will use again kneading sequences. Recall that if the angle ϑ is periodic of period n, then its kneading sequence $\mathbb{K}(\vartheta)$ will be periodic of the same period; it will have the symbol $\star$ exactly at positions $n, 2n, 3n, \ldots$ Moreover, the pointwise limits $\mathbb{K}_-(\vartheta) := \lim_{\vartheta' \nearrow \vartheta} \mathbb{K}(\vartheta')$ and $\mathbb{K}_+(\vartheta) := \lim_{\vartheta' \searrow \vartheta} \mathbb{K}(\vartheta')$ both exist; in one of them, all the symbols $\star$ are replaced by 1 and in the other by 0 throughout. Both are still periodic; in fact, their period is n (or a divisor thereof). More precisely, if the parameter rays at periodic angles ϑ_1 and ϑ_2 both land at the same parameter value, then $\mathbb{K}_\pm(\vartheta_1) = \mathbb{K}_\mp(\vartheta_2)$: it suffices to verify this statement for a single period within the kneading sequences, and this follows from Lemma 3.9. We can thus imagine every pair of periodic parameter rays being replaced by two pairs, infinitesimally close on either side to the given pair and having periodic kneading sequences without the symbol $\star$.

Kneading sequences of preperiodic rays are themselves preperiodic; the lengths of the preperiods of external angle and kneading sequence are equal (this is easy to verify; or see the proof of Lemma 4.1). However, the lengths of the periods do not have to be equal: the ray $9/56 = 0.001\overline{010}$ has kneading sequence $110\overline{111} = 110\overline{1}$. This fact is directly related to the number of parameter rays landing at the same Misiurewicz point; see below.

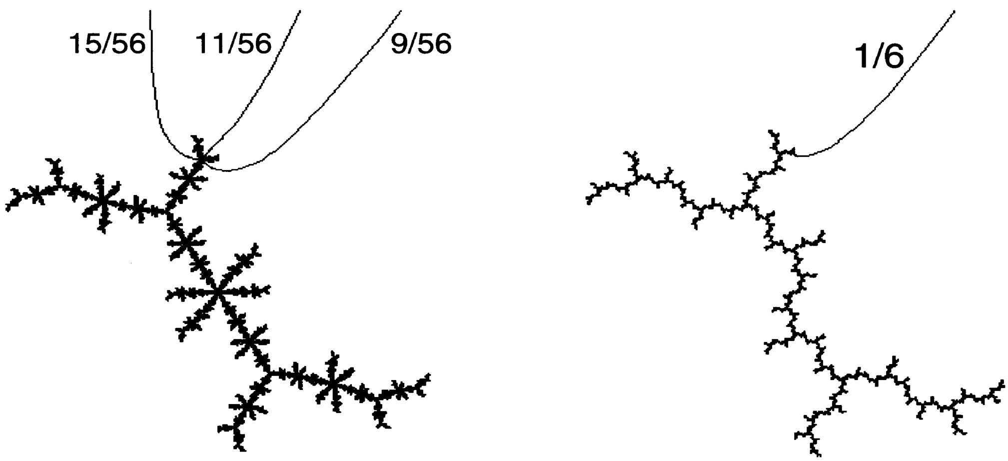


FIGURE 6. *Illustration of the theorem in the preperiodic case. Shown are the Julia sets of the polynomials at the landing points of the parameter rays at angles* 9/56, 11/56 *and* 15/56 *(left) and at angle* 1/6 *(right). In both pictures, the dynamic rays landing at the critical values are drawn.*

Proof of Theorem 1.1 (preperiodic case). — Consider any preperiodic parameter ray at angle ϑ and let c be one of its limit points. First suppose that c is a parabolic parameter. We know that there are two parameter rays at periodic angles ϑ_1, ϑ_2 which land at c. We can imagine two parameter ray pairs infinitesimally close to the ray pair $(\vartheta_1, \vartheta_2)$ on both sides, and these two parameter ray pairs have periodic kneading sequences without symbols $\star$. Each of these two periodic kneading sequences must differ at some finite position from the preperiodic kneading sequence of ϑ. But we had seen in the previous section that the regions of constant initial segments of kneading sequences are bounded by pairs of parameter rays at periodic angles (see Lemma 3.9), so there is a pair of periodic parameter rays landing at the same point separating the parameter ray at angle ϑ from the two rays at angles $\vartheta_{1,2}$ and from the parabolic point c. Therefore, c cannot be a limit point of the parameter ray at angle ϑ. This contradiction shows that no limit point of a preperiodic parameter ray is parabolic.

Now we argue similarly as in the proof of Proposition 3.1. For the parameter c, there is no parabolic orbit, so the dynamic ray at angle ϑ lands at a repelling preperiodic point. We want to show that the landing point is the critical value. Since c is a limit point of the parameter ray at angle ϑ, there are arbitrarily close parameters for which the critical value is on the dynamic ray at angle ϑ.

If, for the parameter c, the dynamic ray at angle ϑ does not land at the critical point or at a point on the backwards orbit of the critical point, then the dynamic ray at angle ϑ and its landing point depend continuously on the parameter by Lemma 2.2, so the critical value must be the landing point of the dynamic ray at angle ϑ for the parameter c. If, however, the landing point of the dynamic ϑ-ray is on the backwards orbit of the critical value, then some finite forward image of this ray will depend continuously on the parameter, and pulling back may yield a dynamic ray bouncing once into the critical value or a point on its backward orbit, but after that the two continuations will land at well-defined points. The ray with both continuations and both landing points will still depend continuously on the parameter, so again the dynamic ϑ-ray must land at the critical value for the parameter c. (However, this contradicts the assumption that the landing point is on the backwards orbit of the critical value because that would force the critical value to be periodic.)

We see that, for any limit point c of the parameter ray at angle ϑ, the number c is preperiodic under $z \mapsto z^2 + c$ with fixed period and preperiod, and c is a Misiurewicz point. Since any such point c satisfies a certain polynomial equation, there are only finitely many such points. The limit set of any ray is connected, so the parameter ray at angle ϑ lands, and the landing point is a Misiurewicz point with the required properties. This shows the third part of Theorem 1.1.

For the last part, we have already shown that a Misiurewicz point cannot be the landing point of a periodic parameter ray, or of a preperiodic ray with external angle different from the angles of the dynamic rays landing at the critical value. It remains to show that, given a Misiurewicz point c_0 such that the critical value is the landing point of the dynamic ray at angle ϑ, then the parameter ray at angle ϑ lands at c_0. We will use ideas from Douady and Hubbard **[DH1]**. By Lemma 2.2, there is a simply connected neighborhood V of c_0 in parameter space such that c_0 can be continued analytically as a repelling preperiodic point, yielding an analytic function $z(c)$ with $z(c_0) = c_0$ such that the dynamic ray at angle ϑ for the parameter $c \in V$ lands at $z(c)$. The relation $z(c) = c$ is certainly not satisfied identically on all of V, so the solutions are discrete and we may assume that c_0 is the only one within V.

Now we consider the winding number of the dynamic ray at angle ϑ around the critical value, which is defined as follows: denoting the point on the dynamic ϑ-ray at potential $t \geq 0$ by z_t and decreasing t from $+\infty$ to 0, the winding number is the total change of $\arg(z_t - c)$ (divided by 2π so as to count in full turns). Provided that the critical value is not on the dynamic ray or at its landing point, the winding number is well-defined and finite and depends continuously on the parameter. When the parameter c moves in a small circle around c_0 and if the winding number is defined all the time, then it must change by an integer corresponding to the multiplicity of c as a root of $z(c) - c$. However, when the parameter returns back to where it started, the winding number must be restored to what it was before. This requires a discontinuity of the winding number, so there are parameters arbitrarily close to c_0 for which the critical value is on the dynamic ray at angle ϑ, and c_0 is a limit point of the parameter ray at angle ϑ. Since this parameter ray lands, it lands at c_0. This finishes the proof of Theorem 1.1. □

Remark. — There is no partition in the dynamic plane showing that preperiodic parameter rays can not land at parabolic parameters: there are countably many preperiodic dynamic rays landing at the boundary of the characteristic Fatou component, for example at preperiodic points on the parabolic orbit, and they cannot be separated by a stable partition.

For the final part of the theorem, we used that a repelling preperiodic point $z(c)$ depends analytically on the parameter. As mentioned before, this proof started with a need to describe parameter spaces of antiholomorphic polynomials like the Tricorn and Multicorns, and there we do not have analytic dependence on parameters. Here is another way to prove that every Misiurewicz point is the landing point of all the parameter rays whose angles are the external angles of the critical value in the dynamic plane. We start with any Misiurewicz point c_0 and external angle ϑ of its critical value. Let c_1 be the landing point of the parameter ray at angle ϑ. Then both parameters c_0 and c_1 have the property that in the dynamic plane, the ray at angle ϑ lands at the critical value. It suffices to prove that this property determines the parameter uniquely. This is exactly the content of the Spider Theorem, which is an iterative procedure to find postcritically finite polynomials with assigned external angles of the critical value. In Hubbard and Schleicher **[HS]**, there is an easy proof for polynomials with a single critical point. While the existence part of that proof works only if the critical point is periodic, all we need here is the uniqueness part, and that works in the preperiodic case just as well, both for holomorphic and for antiholomorphic polynomials.

The last part could probably also be done in a more combinatorial but rather tedious way, using counting arguments like in the periodic case. This would, however, be quite delicate, as the number of Misiurewicz points and the number of parameter rays landing at them require more bookkeeping: the number of parameter rays landing at Misiurewicz points varies and can be any positive integer. The following lemma makes this more precise.

***Lemma 4.1* (Number of Rays at Misiurewicz Points).** — *Suppose that a preperiodic angle ϑ has preperiod l and period n. Then the kneading sequence $\mathbb{K}(\vartheta)$ has the same preperiod l, and its period k divides n. If $n/k > 1$, then the total number of parameter rays at preperiodic angles landing at the same point as the ray at angle ϑ is n/k; if $n/k = 1$, then the number of parameter rays is 1 or 2.*

In the example above, we had seen that the angle $9/56$ has period 3, while its kneading sequence has period 1. Therefore, the total number of rays landing at the corresponding Misiurewicz point is three: their external angles are $9/56, 11/56$ and $15/56$. If more than one ray lands at a given Misiurewicz point, it is not hard to determine all the angles knowing one of them, using ideas from the proof below.

Proof. — In the dynamic plane of ϑ, the dynamic ray at angle ϑ lands at the critical value, so the two inverse image rays at angles $\vartheta/2$ and $(\vartheta + 1)/2$ land at the critical point and separate the dynamic plane into two parts; this partition cuts the external angles of dynamic rays in the same way as in the partition defining the kneading sequence, see Definition 3.8 and Figure 3. We label the two parts by `0` and `1` in the analogous way, assigning the symbol $\star$ to the boundary. The partition boundary intersects the Julia set only at the critical point.

The critical value jumps after exactly l steps onto a periodic orbit of ray period n. Denote the critical orbit by $c_0, c_1, c_2, \ldots$ with $c_0 = 0$ and $c_1 = c$, so that $c_{l+1} = c_{l+n+1}$, while $c_l = -c_{l+n}$. The points c_l and c_{l+n} are on different sides of the partition. The periodic part of the kneading sequence starts exactly where the periodic part of the external angles start, so the preperiods are equal.

We know that n is the ray period of the orbit the critical value falls onto. The orbit period is exactly k: periodic rays which have their entire forward orbits on equal sides of the partition land at the same point, for the following reason: we can connect the landing points of two such rays by a curve which avoids all the preperiodic rays landing at the critical point, all the finitely many rays on their forward orbits, and their landing points (if we have to cross some of these ray pairs, they must also visit the same sides of the partition, and we can reduce the problem). Now inverse images of the rays are connected by inverse images of the curve, which avoids the same rays. Continuing to take inverse images in this way, the periodic landing points must converge to each other, so they cannot be different.

The number of dynamic rays on the orbit of ϑ landing at every point of the periodic orbit is therefore n/k, and the critical value jumps onto this orbit as a local homeomorphism, so it is the landing point of equally many preperiodic rays. But these rays reappear in parameter space as the rays landing at the Misiurewicz point.

It remains to show that there are no extra rays at the periodic orbit. By Lemma 2.4, more than two dynamic rays can land at the same periodic point only if these rays are on the same orbit, i.e., the dynamics permutes the rays transitively. The number of dynamic rays can therefore be greater than n/k only if $n/k = 1$, and in that case, there can be at most two rays. $\square$

Remark. — It does indeed happen that $n/k = 1$ while the number of rays is two. An example is given by the two parameter rays at angles $25/56 = 0/011\,\overline{100}$ and $31/56 = 0.100\,\overline{011}$; their common kneading sequence is $100\,\overline{101}$, so $n = k = 3$, but these two rays land together at a point on the real axis. On the other hand, for the angle $1/2 = 0.0\overline{1} = 0.1\overline{0}$, the kneading sequence is $1\overline{0}$, so $n = k = 1$; the parameter ray at angle $1/2$ is the only ray landing at the leftmost antenna tip $c = -2$ of the Mandelbrot set. These rays are indicated in Figure 1.

For a related discussion of rays landing at common points, from the point of view of "Thurston obstructions", see [**HS**].

5. Hyperbolic Components

Most, if not all, of the interior of the Mandelbrot set consists of what is known as hyperbolic components. A hyperbolic rational map is one where all the critical points are attracted by attracting or superattracting periodic orbits. The dynamic significance is that this is equivalent to the existence of an expanding metric in a neighborhood of the Julia set, which has many important consequences such as local connectivity of connected Julia sets (see Milnor [**M1**]). For a polynomial, the critical point at ∞ is always superattracting, and in the quadratic case, the polynomial is hyperbolic if the unique finite critical point either converges to ∞ or to a finite (super)attracting orbit. Hyperbolicity is obviously an open condition. A *hyperbolic component of the Mandelbrot set* is a connected component of the hyperbolic interior. The period of the attracting orbit is constant throughout the component and defines the *period of the hyperbolic component.* We will see in Corollary 5.4 that every boundary point of a hyperbolic component is a boundary point of the Mandelbrot set, so a hyperbolic component is also a connected component of the interior of $\mathcal{M}$. There is no example known of a non-hyperbolic component; it is conjectured that there are none. A *center* of a hyperbolic component is a polynomial for which there is a superattracting orbit; a *root* of such a component of period n is a parabolic boundary point where the parabolic orbit has ray period n. We will show in Corollary 5.4 that every hyperbolic component has a unique center and a unique root. It is easy to verify that the multiplier of the attracting orbit on a hyperbolic component is a proper map from the component to the open unit disk, so it has a finite mapping degree; we will see that this map is in fact a conformal isomorphism. The relation between centers and roots of hyperbolic components is important; the difficulty in establishing it lies in the discontinuity of Julia sets at parabolic parameters. Proposition 5.2 will help to overcome this difficulty. First we need to have a closer look at parabolic parameters.

***Lemma 5.1* (Roots of Hyperbolic Components).** — *Every parabolic parameter with ray period n is a root of at least one hyperbolic component of period n. If the orbit period k is smaller than n, then this parameter is also on the boundary of a hyperbolic component of period k, and the parabolic orbit breaks up under perturbation into exactly one orbit of period n and k each. If orbit and ray periods are equal, then the parabolic orbit is a merger of exactly two orbits of period n. In no case is such a parabolic parameter on the boundary of a hyperbolic component of different period.*

Proof. — First suppose that orbit period and ray period are equal. Then the first return map of any parabolic periodic point z leaves all the dynamic rays landing at z fixed, so its multiplier is $+1$. In local coordinates, the map has the form $\zeta \mapsto \zeta + \zeta^{q+1} + \cdots$ for some integer $q \geq 1$. The point z then has q attracting and repelling petals each, and every attracting petal must absorb a critical orbit. Since there is a unique critical point, we have $q = 1$. Under perturbation, the parabolic orbit then breaks up into exactly two orbits of exact period n, and no further orbit is involved. Denote the parabolic parameter by c_0 and let V be a simply connected neighborhood of c_0 not containing further parabolics of equal ray period. In $V - \{c_0\}$, all periodic points of exact period n can be continued analytically because their multipliers are different from $+1$. Among these periodic points, those which are repelling at c_0 can be continued analytically throughout all of V, while the two colliding orbits might be interchanged by a simple loop in $V - \{c_0\}$ (in fact, they will be: see Corollary 5.7). Their multipliers are therefore defined on a two-sheeted covering of $V - \{c_0\}$ and are analytic, even when the point c_0 is put back in. By the open mapping principle, the parameter c_0 is on the boundary of at least one hyperbolic component of period n. Since, for the parameter c_0, all the orbits of periods not divisible by n are repelling, the parameter can be only on the boundary of hyperbolic components with periods divisible by n. If it was on the boundary of a hyperbolic component with period rn for some integer $r > 1$, then the rn-periodic orbit would have to be indifferent at c_0; since there can be only one indifferent orbit, it would have to merge with the indifferent orbit of period n, and this orbit would get higher multiplicity than 2, a contradiction. This contradiction shows that c_0 is not on the boundary of any hyperbolic component of period other than n.

If the orbit period strictly divides the ray period, so that $s := n/k \geq 2$, then the first return map of the orbit must permute the rays transitively by Lemma 2.4. The least iterate which fixes the rays must also be the least iterate for which the multiplier is $+1$: the landing point of a periodic ray is either repelling or has multiplier $+1$ (this is the Snail Lemma, see **[M1]**); conversely, whenever the multiplier is $+1$, then all the finitely many rays must be fixed. It follows that the multiplier of the first return map of any of the parabolic periodic points is an exact s-th root of unity. The periodic orbit can then be continued analytically in a neighborhood of the root. Since the multiplier map is analytic, the parabolic parameter is on the boundary of a hyperbolic component of period k. The s-th iterate of the first return map has multiplier $+1$ and hence again the form $\zeta \mapsto \zeta + \zeta^{q+1} + \cdots$ in local coordinates, for an integer $q \geq 1$. The number of coalescing fixed points of this iterate is then exactly $q+1$.

Since there is only one critical orbit, the first return map of the parabolic orbit must permute the q attracting petals transitively and we have $q = s$. For the first, second, $\ldots$, $s-1$-st iterate of the first return map, the multiplier is different from $+1$, so the respective iterate has a single fixed point. The s-th iterate, however, corresponding to the $sk = n$-th iterate of the original polynomial, has a fixed point of multiplicity $q+1 = s+1$: exactly one of these points has exact period k; all the other points can have no lower periods than n, so they are on a single orbit of period n of which s points each are coalesced. There is no further orbit involved (or some iterate would have to have a parabolic fixed point of higher multiplicity with more attracting petals attached, as above). Since there is a single indifferent orbit of period n, its multiplier is well defined and analytic in a neighborhood of the parabolic parameter, which is hence on the boundary of a hyperbolic component of period n as well. □

A root of a hyperbolic component is called *primitive* if its parabolic orbit has equal orbit and ray periods, so it is the merger of exactly two orbits of equal period. If orbit and ray periods are different, then the root is called *non-primitive* or a *bifurcation point*: at this parameter, an attracting orbit bifurcates into another attracting orbit of higher period (the terminology *bi*-furcation comes from the dynamics on the real line, where the ratio of the periods is always two). We will see below that every hyperbolic component has a unique root. It will therefore make sense to call a hyperbolic component primitive or non-primitive according to whether or not its root is primitive.

Our next goal is to relate the dynamics at root points to the adjacent hyperbolic components. Perturbations of parabolics are a subtle issue because both the Julia set and the filled-in Julia set behave drastically discontinuously. We show that nonetheless the landing points of all the dynamic rays at rational angles behave continuously wherever the rays land. However, we will explain below that the rays themselves do not depend continuously on the parameter. In a way, the following proposition is the parabolic analogue to Lemma 2.2, which dealt with repelling periodic points. This proposition has first been shown by Lavaurs: it is the main result of Exposé XVII in **[DH1]**. Our proof is different; one of our main ingredients will be the Orbit Separation Lemma 3.7.

Proposition 5.2 **(Continuous Dependence of Landing Points).** — *For any rational angle ϑ, the landing point of the dynamic ray at angle ϑ depends continuously on the parameter on the entire subset of parameter space for which the ray lands.*

The dynamic ray at angle ϑ for the polynomial p_c fails to land if and only if it bounces into the critical point or into a point on the inverse orbit of the critical point, which happens if and only if the parameter c is outside the Mandelbrot set on a parameter ray at one of the finitely many angles $\{2\vartheta, 4\vartheta, 8\vartheta, \dots\}$. All these parameter rays land, and at the landing parameters, the dynamic ray at angle ϑ lands as well. These landing points are the interesting cases of the proposition.

Proof. — It suffices to discuss the case of a periodic angle ϑ: the statement for preperiodic angles follows by taking inverse images because the pull-back is continuous (if the landing point of a preperiodic ray visits a critical point along its preperiodic orbit, which happens at Misiurewicz points, then several preperiodic points may merge and split up with different rays, but this happens in a continuous way).

If the landing point of the dynamic ray at angle ϑ is repelling, then the proposition reduces to Lemma 2.2. We may thus assume that the landing point is parabolic. Under perturbation, any parabolic periodic point splits up into several periodic points which may be attracting, repelling, or indifferent, and these periodic points depend continuously on the parameter. We need to show that the landing point of the ray after perturbation is one of the continuations of the parabolic periodic point it was landing at.

Denote the parabolic parameter before perturbation by c_0, let n be its ray period and let $V \subset \mathbb{C}$ be a simply connected open neighborhood which does not contain further parabolics of equal or lower ray periods, and which does not meet parameter rays of period n other than those landing at c_0. Then analytic continuation of repelling periodic points of periods up to n is possible in $V - \{c_0\}$.

First we discuss the case that the parabolic parameter c_0 is non-primitive with orbit period k and $q := n/k \geq 2$. Then c_0 is on the boundary of at least one hyperbolic component of period n and k each. The multipliers μ_n and μ_k of the unique orbits of periods n and k which are indifferent at c_0 are well-defined and analytic in a neighborhood of c_0: for the period k orbit, this is clear because the periodic point $z = z(c)$ can be continued analytically near c_0 and the multiplier depends analytically on c and z; for the period n orbit, this is also true because the multiplier is well-defined and analytic near c_0, except at c_0 itself, and has a removable isolated singularity at c_0. Critical points of the multiplier maps are thus isolated and the boundaries of both components are piecewise smooth analytic curves.

Within a hyperbolic component of period n near c_0, let $z_k(c)$ the periodic point of period k which merges into the parabolic orbit of c_0 at the characteristic point; it is on a repelling orbit and can be continued analytically within the component. Because of the piecewise analytic boundary, there is a curve within the hyperbolic component converging to c_0. The point $z_k(c)$ is the landing point of at least one periodic dynamic ray at some angle ϑ', and it will keep this ray throughout the hyperbolic component by Lemma 2.2. In fact, it can lose the ray only at c_0 and only in favor of the parabolic orbit. Since $z_k(c)$ merges into the parabolic orbit and all the parabolic periodic points are separated by stable partitions by the Orbit Separation Lemma 3.7, the point $z_k(c)$ will keep the ray ϑ' even at the parabolic orbit. Therefore, ϑ' is one of the external angles of the characteristic point on the parabolic orbit. Since we are in the non-primitive case, all the rays are on the same orbit by Lemma 2.4, and $z_k(c)$ is the landing point of all the dynamic rays of the characteristic point of the parabolic orbit. Under perturbation into the period n component, the landing points of all dynamic rays from the parabolic orbit will thus depend continuously on the parameter. The same argument works for perturbations into a hyperbolic component of period k: we consider a repelling periodic point $z_n(c)$ which merges into the characteristic point of the parabolic orbit. It must be the landing point of one of the dynamic rays from the characteristic point of the parabolic orbit as above, and all the other dynamic rays from the parabolic orbit must land at points on the orbit of $z_n(c)$.

The two parameter rays landing at c_0 together with their landing point cut $\mathbb{C}$ into two connected components, and continuity of the landing points of the dynamic rays from the parabolic orbit is true for any connected component which contains a hyperbolic component of period n or k by the arguments above. But the components of different periods must be in different connected components because analytic continuation of $z_k(c)$ within $V - \{c_0\}$ from a period n component to a period k component must change the landing pattern of the dynamic rays, and the only place where this can happen is at a parameter ray landing at c_0. This proves the proposition in the non-primitive case (and as a bonus result, we see that c_0 can be on the boundary of only one component of period n and k each).

In the primitive case, the rays landing at the parabolic orbit are on two different ray orbits, and the characteristic point of the parabolic orbit is the landing point of exactly one ray from each orbit. The proof above shows continuity of the landing points of at least one of the two ray orbits. Instead, we will follow a suggestion of Tan Lei and use an Orbit Separation Property: *in the dynamics of c_0, any repelling periodic point of period at most n can be separated from the characteristic point of the parabolic orbit by a ray pair landing at a repelling periodic or preperiodic point.* This prevents the dynamic rays landing at the parabolic orbit from jumping onto other orbits, and the Orbit Separation Lemma 3.7 prevents them from jumping onto other points on the parabolic orbit. This proves the proposition also in the primitive case. (This Orbit Separation Property does not hold in the non-primitive case.)

In order to prove the Orbit Separation Property, we will use the tree constructed in Section 3. Let w be a repelling periodic point of period at most n, let z_1 be the characteristic point of the parabolic orbit and let U_1 be the characteristic Fatou component. The only periodic point of period at most n on $\overline{U}_1$ is the characteristic parabolic point, so $w \notin \overline{U}_1$. Within the filled-in Julia set of c_0, connect w to z_1 by a simple curve γ; the point where the curve first meets $\overline{U}_1$ is the image of the projection $\pi(w)$ onto U_1. Let γ be the curve between w and $\pi(w)$.

The curve γ is unique except on the closures of bounded Fatou components (compare the remark before Lemma 3.6). If it does not meet any bounded Fatou components (except for U_1 at its end), then it is unique and each of its points except w is the landing point of at least two dynamic rays. Then every iterate of the dynamics must map γ homeomorphically which contradicts the expansion from angle doubling. Therefore, an interior point of γ will be on the closure of some periodic or preperiodic bounded Fatou component. Let n' be the least number of iterations for an interior point of γ to hit a periodic Fatou component. Then γ maps forward homeomorphically for at least n' iterations, and the periodic Fatou component it meets on its interior is different from the one at its end. Therefore, the n'-th image of γ connects two different periodic Fatou components. These two Fatou components can be separated by a periodic or preperiodic dynamic ray pair landing on a repelling orbit (in the primitive case, no two periodic Fatou components have a boundary point in common, so the periodic Fatou components inherit the separation property from the parabolic periodic points as shown in Lemma 3.7). But since γ maps homeomorphically onto its n'-th image, the two endpoints of γ are also separated by a ray pair on a repelling orbit. This proves the Orbit Separation Property and thus the proposition. □

Remark. — Unlike their landing points, the dynamic rays themselves may depend discontinuously on the perturbation. The simplest possible example occurs near the parabolic parameter $c_0 = 1/4$: for this parameter, the dynamic ray at angle $0 = 1$ lands at the parabolic fixed point $z = 1/2$, and the ray is the real line to the right of $1/2$. The critical point 0 is in the interior of the filled-in Julia set. Perturbing the parameter to the right on the real axis, i.e., on the parameter ray at angle $0 = 1$, the dynamic ray will bounce into the critical point and thus fail to land. But for arbitrarily small perturbations near this parameter ray, the dynamic ray at angle $0 = 1$ will get very close to the critical point before it turns back and lands near $1/2$. The closer the parameter is to $c_0 = 1/4$, the lower will the potential of the critical point be, and while the dynamic ray keeps reaching out near the critical point, it does so at lower and lower potentials, and in the limit the part of the ray at real parts less than $1/2$ will be squeezed off. Points at any potential $t > 0$ will depend continuously on the parameter, and so does the landing point at potential $t = 0$; however, this continuity is not uniform in t, and the dynamic ray as a whole can and does change discontinuously with respect to the Hausdorff metric.

Continuous dependence of landing points of rays requires a single critical point (of possibly higher multiplicity). It is false already for cubic polynomials; an example can be found in the appendix of Goldberg and Milnor **[GM]**.

We can now draw a couple of useful conclusions.

***Corollary 5.3* (Stability at Roots of Hyperbolic Components).** — *For any hyperbolic component, the landing pattern of periodic and preperiodic dynamic rays is the same for all polynomials from the component and at any of its roots.*

Proof. — Again, it suffices to discuss periodic rays; the statement about preperiodic rays follows simply by taking inverse images because for the considered parameters, all the preperiodic dynamic rays land, and they never land at the critical value. Throughout the component, all the periodic rays land at repelling periodic points, so no orbit can lose a ray under perturbations, and consequently no orbit can gain a ray, either. Hence we only have to look at the roots of the component. Upon perturbation into the component, the proof of the previous proposition shows that all the rays from the parabolic orbit will land in the same way at a single repelling orbit, and all the repelling periodic points at the root parameter will keep their rays by Lemma 2.2. □

Remark. — Perturbing a parabolic orbit with orbit period k and ray period $n > k$ into the component of period k changes the landing pattern of rational rays: the parabolic orbit creates an attracting orbit of period k, so a repelling orbit of period n remains, and all the dynamic rays from the parabolic orbit land at different points. Phrased differently, when moving from the component of period k into the one of period n, then n/k periodic rays each start landing at common periodic points; of course, this forces the obvious relations for the preperiodic rays. The landing pattern of all other periodic rays remains stable.

This discussion not only describes the landing patterns of periodic rays within hyperbolic components, but on the entire parameter space except at parameter rays at periodic angles: since the landing points depend continuously on the parameter and periodic orbits are simple except at parabolics, the pattern can change only at parabolic parameters or at parameter rays where the dynamic rays fail to land. The relation between landing patterns of periodic rays and the structure of parameter space has been investigated and described by Milnor in **[M2]**. The landing pattern of preperiodic rays also changes at Misiurewicz points and at preperiodic parameter rays.

***Corollary 5.4* (The Multiplier Map).** — *The multiplier map on any hyperbolic component is a conformal isomorphism onto the open unit disk, and it extends as a homeomorphism to the closures. In particular, every hyperbolic component has a unique* *root and a unique center. The boundary of a hyperbolic component is contained in the boundary of the Mandelbrot set.*

Proof. — The multiplier map is a proper analytic map from the hyperbolic component to the open unit disk and extends continuously to the boundary. Any point on the boundary has a unique indifferent orbit. If the multiplier at such a boundary point is different from $+1$, then orbit and multiplier extend analytically in a neighborhood of the boundary point. The multiplier is obviously not constant. This shows in particular that parabolic parameters are dense on the boundary of the component; since parabolics are landing points of parameter rays and thus in the boundary of $\mathcal{M}$, the boundary of every hyperbolic component is contained in the boundary of the Mandelbrot set.

The number of parabolic parameters with fixed orbit period and multiplier $+1$ is finite, so the boundary of any hyperbolic component consists of a finite number of analytic arcs (which might contain critical points) limiting on finitely many parabolic parameters with multipliers $+1$. Since the multiplier map is proper onto $\mathbb{D}$, the component has at least one root.

By Corollary 5.3, the landing pattern of periodic dynamic rays has to be the same at all the roots of a given hyperbolic component. It follows that at all the roots, the angles of the characteristic rays of the parabolic orbits have to coincide: in every case, the characteristic ray pair lands at the Fatou component containing the critical value and separates the critical value from the rest of the parabolic orbit. Among all ray pairs separating the critical point from the critical value, the characteristic ray pair must be the one closest to the critical value. Hence the landing pattern of periodic dynamic rays determines the characteristic angles. Since any root of a hyperbolic component must be the landing point of the parameter rays at the characteristic angles of the parabolic orbit, every hyperbolic component has a unique root.

We can now determine the mapping degree d, say, of the multiplier map. The hyperbolic component is simply connected because the Mandelbrot set is full. Then the multiplier map μ has exactly $d-1$ critical points, counting multiplicities. If $d > 1$, let $v \in \mathbb{D}$ be a critical value of μ and connect v and $+1$ by a simple smooth curve $\gamma \subset \mathbb{D}$ avoiding further critical values. Then $\mu^{-1}(\gamma)$, together with the unique root of the component, contains a simple closed curve Γ enclosing an open subset of the hyperbolic component. This subset must map at least onto all of $\mathbb{D}-\gamma$, so Γ surrounds boundary points of the component and thus of $\mathcal{M}$. But this contradicts the fact that the Mandelbrot set is full, so the multiplier map is a conformal isomorphism onto $\mathbb{D}$ and the component has a unique center. It extends continuously to the closure and is surjective onto $\partial\mathbb{D}$ because it is surjective on the component. Near every non-root of the component, the boundary of the component is an analytic arc (possibly with critical points) and the multiplier is not locally constant on the arc, so it is locally injective. Global injectivity now follows from injectivity on the component. The multiplier map is thus invertible, and continuity of the inverse is a generality. □

***Proposition 5.5* (No Shared Roots).** — *Every parabolic parameter is the root of a single hyperbolic component.*

Proof. — Since the period of a component equals the ray period of its root, we can restrict attention to any fixed period n. We have well defined maps from centers to hyperbolic components (which we have just seen is a bijection) and from hyperbolic components to their roots. This gives a surjective map from centers of period n to parabolic parameters of ray period n. Denote the number of centers of period n by s_n.

A center of a hyperbolic component of period n is a point c such that the critical point 0 is periodic of exact period n under $z \mapsto z^2 + c$; therefore, c must satisfy a polynomial equation $(\ldots((c^2+c)^2+c)\ldots)^2+c=0$ of degree 2^{n-1}. Since this polynomial is also solved by centers of components of periods k dividing n, we get the recursive relation $\sum_{k|n} s_k = 2^{n-1}$. By Lemma 2.3, this is exactly the number of parabolic parameters of ray period k. Since a surjective map between finite sets of equal cardinality is a bijection, every parabolic parameter is the root of a single hyperbolic component. □

Remark. — This proposition shows even without resorting to Corollary 5.4 that every hyperbolic component has a unique center, so that the only critical point of the multiplier map (if it had mapping degree greater than one) could be the center. This is indeed what happens for the "Multibrot sets": the connectedness loci for the maps $z \mapsto z^d + c$ with $d \geq 2$.

Before continuing the study of hyperbolic components, we note an algebraic observation following from the proof we have just given.

***Corollary 5.6* (Centers of Components as Algebraic Numbers)**

Every center of a hyperbolic component of degree n is an algebraic integer of degree at most s_n. It is a simple root of its minimal polynomial. □

A neat algebraic proof for this fact has been given by Gleason; see **[DH1]**. As far as I know, the algebraic structure of the minimal polynomials of the centers of hyperbolic components is not known: when factored according to exact periods, are they irreducible? What are their Galois groups? Manning (unpublished) has verified irreducibility for $n \leq 10$, and he has determined that the Galois groups for the first few periods are the full symmetric groups. Giarrusso (unpublished) has observed that this induces a Galois action between the Riemann maps of hyperbolic components of equal periods, provided that their centers are algebraically conjugate.

Now we can describe the boundary hyperbolic components much more completely.

***Corollary 5.7* (Boundary of Hyperbolic Components).** — *No non-parabolic parameter can be on the boundary of more than one hyperbolic component. Every parabolic parameter is either a primitive root of a hyperbolic component and on the boundary of no further component, or it is a "point of bifurcation": a non-primitive root of a hyperbolic component and on the boundary of a unique further hyperbolic component. In particular, if two hyperbolic components have a boundary point in common, then this point is the root of exactly one of them. The boundary of a hyperbolic component is a smooth analytic curve, except at the root of a primitive component. At a primitive root, the component has a cusp, and analytic continuation of periodic points along a small loop around this cusp interchanges the two orbits which merge at this cusp.*

Proof. — If two hyperbolic components have a non-parabolic parameter in their common boundary, then the landing patterns of periodic rays must be the same within both components. This must then also be true at their respective roots, which yields a contradiction: on the one hand, the roots must be different by Proposition 5.5; on the other hand, the parabolic orbits at the roots must have the same characteristic angles (compare the proof of Corollary 5.4), so they must be the landing points of the same parameter rays. It follows that the multiplier map of the indifferent orbit cannot have a critical point at c_0: if it had a critical point there, then c_0 would connect locally two regions of hyperbolic parameters which cannot belong to different hyperbolic components; however, if they belonged to the same component, then the closure of the component would separate part of its boundary from the exterior of the Mandelbrot set, a contradiction. Therefore, the boundary of every hyperbolic component is a smooth analytic curve near every non-parabolic boundary point.

Now let c_0 be a parabolic parameter of ray period n and orbit period k. We know that it is the root of a unique hyperbolic component of period n.

In the non-primitive case (when k strictly divides n), the point c_0 cannot be on the boundary of a hyperbolic component of period different from n and k by Lemma 5.1. It is on the boundary of a single hyperbolic component of period n (Proposition 5.5). In a small punctured neighborhood of c_0 avoiding further parabolics of ray period n, the multiplier map of the n-periodic orbit is analytic and cannot have a critical point at c_0, for the same reason as above, so the component of period n occupies asymptotically (on small scales) a half plane near c_0. The parameter c_0 is also on the boundary of a component of period k the multiplier of which is analytic near c_0. Since hyperbolic components cannot overlap, this component must asymptotically be contained in a half plane, so the multiplier map cannot have a critical point and must then be locally injective near c_0. The boundaries of both components must then be smooth analytic curves near c_0.

In the primitive case $k = n$, the parameter c_0 cannot be on the boundary of a hyperbolic component of period different from n by Lemma 5.1, and it cannot be on the boundary of two hyperbolic components of period n because otherwise it would have to be their simultaneous root, contradicting Proposition 5.5. In a small simply connected neighborhood V of c_0, analytic continuation of the two orbits colliding at c_0 is possible in $V - \{c_0\}$ (compare the proof of Lemma 5.1), so their multipliers can be defined on a two-sheeted cover of V ramified at c_0. If analytic continuation of these two orbits along a simple loop in V around c_0 did not interchange the two orbits, then both multipliers could be defined in V, and both would define different hyperbolic components intersecting V in disjoint regions, yielding the same contradiction as in the non-primitive case above. Therefore, small simple loops around c_0 do interchange the two orbits. In order to avoid the same contradiction again, the multiplier must be locally injective on the two-sheeted covering on V. Projecting down onto V, the component must asymptotically occupy a full set of directions, so the component has a cusp. □

Remark. — The basic motor for many of these proofs about hyperbolic components was uniqueness of parabolic parameters with given combinatorics, via landing properties of parameter rays at periodic angles. A consequence was uniqueness of centers of hyperbolic components with given combinatorics. One can also turn this discussion around and start with centers of hyperbolic components: the fact that hyperbolic components must have different combinatorics, and that they have unique centers, is a consequence of Thurston's topological characterization of rational maps, in this case most easily used in the form of the Spider Theorem [**HS**].

References

[Bo] T. Bousch: *Sur quelques problèmes de la dynamique holomorphe.* Thèse, Université de Paris-Sud (1992).

[CG] L. Carleson, T. Gamelin: *Complex dynamics.* Universitext, Springer Verlag (1993).

[DH1] A. Douady, J. Hubbard: *Etude dynamique des polynômes complexes.* Publications mathématiques d'Orsay 84-02 (1984) (première partie) and 85-04 (1985) (deuxième partie).

[DH2] A. Douady, J. Hubbard: *On the dynamics of polynomial-like mappings.* Ann. Scient. Ec. Norm. Sup. 18 (1985), 287–343.

[ES] D. Eberlein, D. Schleicher: *Rational parameter rays of Multibrot sets.* Manuscript (1999).

[GM] L. Goldberg, J. Milnor: *Fixed points of polynomial maps I/II.* Ann. Scient. Ec. Norm. Sup. 25 (1992) and 26 (1993), 51–98.

[HS] J. Hubbard, D. Schleicher: *The spider algorithm.* In: *Complex dynamics: the mathematics behind the Mandelbrot and Julia sets.* AMS, Proceedings in Applied Mathematics **49** (1994), 155–180.

[Ke1] K. Keller: *Symbolic dynamics for angle-doubling on the circle III. Sturmian sequences and the quadratic map.* Ergod. Th. Dyn. Sys. 14 (1994), 787–805.

[Ke2] K. Keller: *Invariante Faktoren, Juliaäquivalenzen und die abstrakte Mandelbrotmenge.* Habilitationsschrift, Universität Greifswald (1996).

[La1] P. Lavaurs: *Une description combinatoire de l'involution definie par M sur les rationnels à dénominateur impair.* Comptes Rendus Aca. Sci. Paris (4) 303 (1986) 143–146.

[La2] P. Lavaurs: *Systèmes dynamiques holomorphes: explosion de points périodiques paraboliques.* Thèse, Université de Paris-Sud (1989).

[LS] E. Lau, D. Schleicher: *Internal addresses in the Mandelbrot set and irreducibility of polynomials.* Preprint, Institute for Mathematical Sciences, Stony Brook, #19 (1994).

[M1] J. Milnor: *Dynamics in one complex variable: introductory lectures.* Vieweg Verlag (1999).

[M2] J. Milnor: *Periodic orbits, external rays and the Mandelbrot set; an expository account.* In this volume, 277–333.

[McM] C. McMullen: *The Mandelbrot set is universal.* Preprint (1997).

[NS] S. Nakane, D. Schleicher: *On multicorns and unicorns I: antiholomorphic dynamics and hyperbolic components.* In preparation.

[Pr1] C. Penrose: *On quotients of the shift associated with dendrite Julia sets of quadratic polynomials.* Thesis, University of Coventry (1990).

[Pr2] C. Penrose: *Quotients of the shift associated with dendrite Julia sets.* Manuscript (1994).

[S1] D. Schleicher: *Internal addresses in the Mandelbrot set and irreducibility of polynomials.* Thesis, Cornell University (1994).

[S2] D. Schleicher: *On Fibers and Local Connectivity of the Mandelbrot and Multibrot Sets.* Preprint, Institute for Mathematical Sciences, Stony Brook, #13 (1998).

[T] W. Thurston: *On the geometry and dynamics of iterated rational maps.* Preprint, Princeton University (1985).

[TY] Tan Lei and Yin Yongcheng: *Local connectivity of the Julia set for geometrically finite rational maps.* Science in China (Series A) **39** 1 (1996), 39–47.

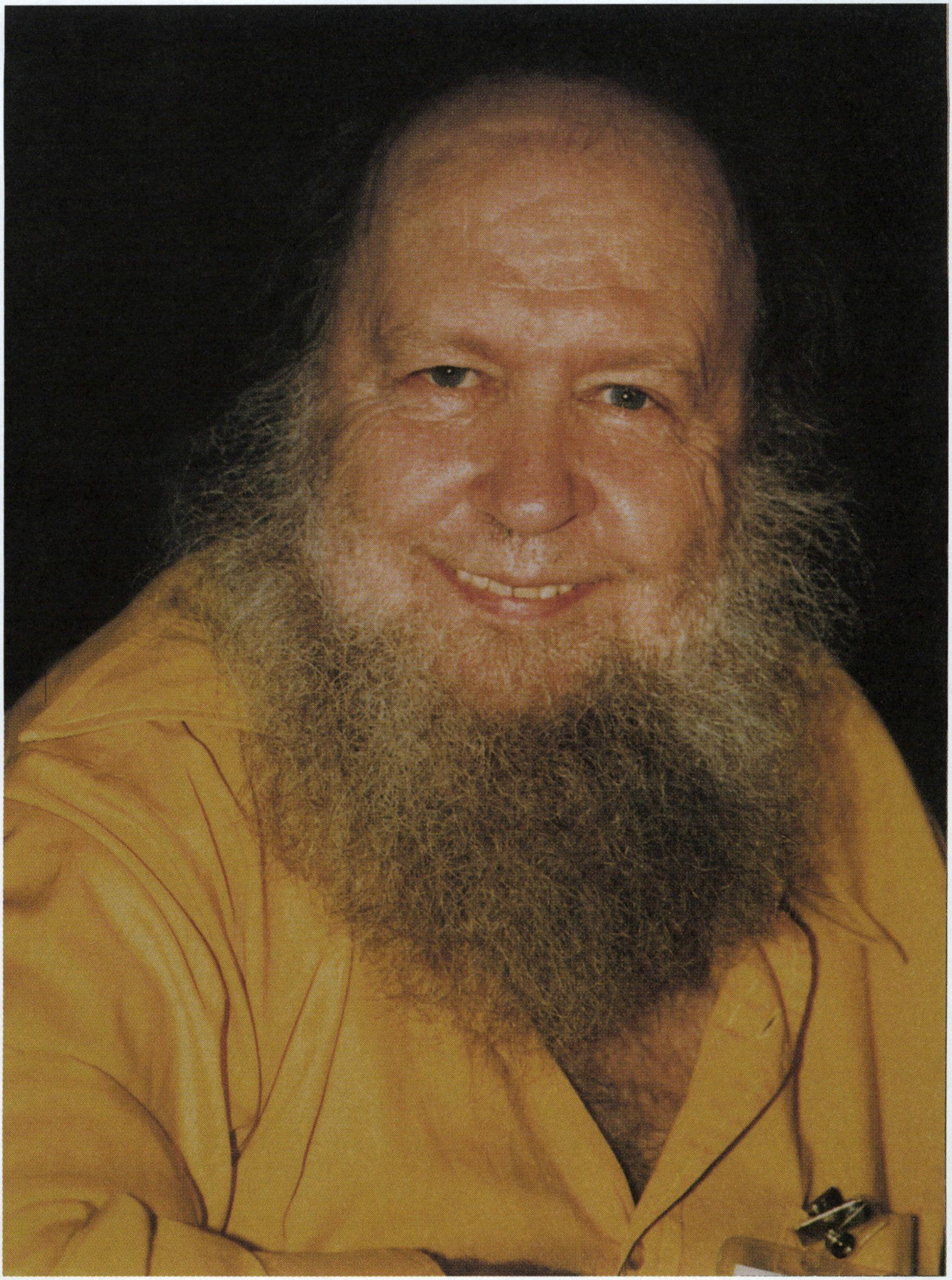
D. Schleicher, Fakultät für Mathematik, Technische Universität, Barer Straße 23, D-80290 München, Germany • *E-mail :* `dierk@mathematik.tu-muenchen.de`
Url : `http://pckoenig1.mathematik.tu-muenchen.de/~dierk`

AST

Géométrie complexe et systèmes dynamiques - Colloque en l’honneur d’Adrien Douady Orsay, 1995 - Pages préliminaires

Astérisque, tome 261 (2000), p. II-XXI

<http://www.numdam.org/item?id=AST_2000__261__R2_0>



Adrien Douady, 1995

ASTÉRISQUE 261

GÉOMÉTRIE COMPLEXE ET SYSTÈMES DYNAMIQUES COLLOQUE EN L'HONNEUR D'ADRIEN DOUADY,

ORSAY 1995

édité par

Marguerite Flexor

Pierrette Sentenac

Jean-Christophe Yoccoz

Société Mathématique de France 2000

Publié avec le concours du Centre National de la Recherche Scientifique

Marguerite Flexor
Université Paris-Sud, CNRS UMR 8628, Bâtiment 430, Mathématiques, 91405 Orsay Cedex.
E-mail : `Marguerite.Flexor@math.u-psud.fr`

Pierrette Sentenac
Université Paris-Sud, CNRS UMR 8628, Bâtiment 430, Mathématiques, 91405 Orsay Cedex.
E-mail : `Pierrette.Sentenac@math.u-psud.fr`

Jean-Christophe Yoccoz
Collège de France, 3 rue d'Ulm, 75005 Paris.
Université Paris-Sud, CNRS UMR 8628, Bâtiment 430, Mathématiques, 91405 Orsay Cedex.

***Classification mathématique par sujets* (1991).** — 30C, 30D, 30F, 32A, 32B, 32G, 32S, 32L, 37A, 37C, 37D, 37E, 37XX, 52A, 53C, 58F.

***Mots clefs*.** — Théorème central limite, Collet-Eckmann, équisingularité, ergodicité, application de Hénon, feuilletages holomorphes, mouvements holomorphes, orbites homoclines, hyperbolicité, mesures invariantes, itération, ensembles de Julia, applications de Lorenz, ensemble de Mandelbrot, exposant de Lyapounov positif, applications rationnelles, mesures SRB, volume.

GÉOMÉTRIE COMPLEXE ET SYSTÈMES DYNAMIQUES COLLOQUE EN L'HONNEUR D'ADRIEN DOUADY, ORSAY 1995

édité par Marguerite Flexor, Pierrette Sentenac, Jean-Christophe Yoccoz

Résumé. — Ce volume est issu du colloque « Géométrie complexe et Systèmes dynamiques » qui s'est tenu à l'université Paris-sud (Orsay) du 3 au 8 juillet 1995, à l'occasion du soixantième anniversaire d'Adrien Douady. Les articles qui le composent illustrent de nombreuses facettes de l'activité actuelle en systèmes dynamiques : itération des polynômes (en particulier quadratiques), des fractions rationnelles, feuilletages holomorphes, dynamique non uniformément hyperbolique.

***Abstract* (Conference on Complex Geometry and Dynamical Systems in honor of Adrien Douady, Orsay 1995)**

On account of Adrien Douady's sixtieth birthday, a meeting on "Complex Geometry and Dynamical Systems" was held in July 1995 at the University of Paris-Sud (Orsay). The present book is an outcome of that meeting. The many-faceted activity in the field of dynamical systems is reflected in the papers gathered in the volume. Topics covered include iteration of polynomials (with special regards to quadratic ones), rational fractions, holomorphic foliations, and non uniformly hyperbolic dynamics.

TABLE DES MATIÈRES

RÉSUMÉS DES ARTICLES

Le but de cet article est double : discuter le concept de mouvements holomorphes et des phénomènes de type Mañé-Sad-Sullivan en plusieurs variables complexes ; comparer les différentes notions de formes de Beltrami en CR géométrie qui sont apparues dans [4] et [7].

Dans cet article, on considère les applications de Hénon pour lesquelles l'analyse de [BC2] est valable. On construit des ensembles munis de bonnes propriétés hyperboliques et de bonnes structures de retour, et on montre que leurs fonctions de temps de retour ont des restes à décroissance exponentielle. Ceci permet d'appliquer les résultats de [Y]. Des propriétés statistiques telles que la décroissance exponentielle des corrélations et le théorème central limite sont établies.

Nous considérons le problème de la classification des feuilletages projectifs complexes ayant un ensemble limite algébrique. Nous démontrons le résultat suivant :

Soit $\mathcal{F}$ *un feuilletage holomorphe par des courbes dans le plan projectif complexe* $\mathbb{C}P(2)$ *dont l'ensemble limite se compose d'une courbe algébrique* Λ *et de singularités. Si les singularités* $\operatorname{sing}\mathcal{F} \cap \Lambda$ *sont génériques alors ou bien* $\mathcal{F}$ *est donné par une* 1*-forme rationnelle fermée ou bien* $\mathcal{F}$ *est l'image réciproque par une application rationnelle d'un feuilletage de Riccati* $\mathcal{R} : p(x)dy - (a(x)y^2 + b(x)y)dx = 0$ *(où* Λ *correspond à* $\overline{(y=0)} \cup \overline{(p(x)=0)}$*) dans* $\overline{\mathbb{C}} \times \overline{\mathbb{C}}$*.*

La preuve repose sur la résolubilité des groupes d'holonomie généralisée associés à un processus de réduction des singularités de $\operatorname{sing}\mathcal{F} \cap \Lambda$ et sur la construction d'une structure affine transverse à $\mathcal{F}$ en dehors de la courbe algébrique invariante contenant Λ.

Nous donnons une minoration du volume d'un domaine compact convexe d'un espace euclidien dont le bord est de classe $C^{1,1}$. Nous caractérisons le cas d'égalité.

Nous montrons l'existence de mesures de Sinaï-Ruelle-Bowen pour une classe d'applications C^2 par morceaux d'un rectangle dans lui-même dont les dérivées ne sont pas nécessairement bornées. Ces résultats peuvent être considérés comme une généralisation d'un théorème bien connu en dimension 1 sur l'existence de mesures absolument continues invariantes. Dans un article précédent [8], des résultats semblables étaient énoncés et les preuves esquissées pour des systèmes inversibles. Nous donnons ici des preuves complètes dans le cas général de systèmes non inversibles ; nous développons en particulier la théorie des variétés stables et instables lorsque les dérivées des applications considérées ne sont pas bornées.

Dans cet article, nous considérons des polynômes à coefficients réels avec un point critique éventuellement dégénéré d'ordre pair et dont l'orbite reste bornée. Nous donnons des conditions nécessaires et suffisantes pour que leur ensemble de Julia soit totalement discontinu. Nous montrons aussi que ces ensembles de Julia ne portent pas de champs de droites invariants. Dans le cas réel, ceci généralise les résultats de B. Branner et J.H. Hubbard sur les polynômes cubiques et les résultats de C. McMullen sur l'absence de champs de droites invariants.

Cet article fait partie d'une série sur la dynamique des polynômes quadratiques. Nous transportons notre résultat géométrique précédent [L3] au plan des paramètres. À toute valeur c (en dehors de la cardioïde principale et des copies qui s'y rattachent) est associée « une suite principale de pièces gigognes du parapuzzle ». Nous montrons alors que les modules des anneaux entre deux pièces consécutives croissent au moins linéairement. D'après ([L2])) et le critère géométrique de Martens & Nowicki (*cf.* ce volume) ceci implique que presque tout polynôme quadratique réel (au sens de la mesure de Lebesgue) est hyperbolique ou possède une mesure finie absolument continue invariante ou est infiniment renormalisable. Dans des articles ultérieurs [L5,L7] nous montrons que l'ensemble des paramètres réels infiniment renormalisables est de mesure

nulle, ce qui complète la description de la dynamique pour presque tout polynôme quadratique réel.

Nous introduisons une classe de familles d'applications réelles dépendant d'un paramètre qui étend les modèles géométriques classiques de Lorenz. Ces applications sont à la fois singulières (discontinuités avec derivées infinies) et possèdent des points critiques ; elles sont basées sur le comportement du flot de Lorenz pour un ensemble de paramètres important. Notre résultat principal dit qu'une expansion non-uniforme est le type de dynamique que l'on retrouve le plus souvent même s'il y a formation de points critiques.

Nous discutons une condition suffisante, de nature géometrique, pour l'existence de mesures de probabilité invariantes et absolument continues pour des applications $S-$unimodales. Il en résulte qu'une application quadratique typique, au sens de la mesure de Lebesgue, admet une mesure de SRB.

Après avoir étudié la dépendance analytique des séparatrices d'une famille « équisingulière » de germes de feuilletages holomorphes à l'origine de $\mathbb{C}^2$, nous définissons la quasihomogénéité comme une propriété de rigidité. Nous obtenons un théorème de type K. Saito pour les germes de feuilletages quasihomogènes et un théorème de type Briançon-Skoda dans le cas général.

Nous expliquons quelques résultats fondamentaux de Douady-Hubbard sur l'ensemble de Mandelbrot en utilisant l'idée de « portrait orbital » c'est-à-dire le modèle des rayons externes qui aboutissent sur une orbite périodique d'une application polynomiale quadratique.

On présente, à travers des résultats récents, des problèmes ouverts et des conjectures, une perspective globale pour l'étude des systèmes dynamiques dissipatifs (flots, difféomorphismes ou transformations d'une variété compacte sans bord ou de l'intervalle). Cette perspective est couronnée par une description conjecturale de la dynamique d'un ensemble dense de systèmes : pour ceux-ci,

il n'y a qu'un nombre fini d'attracteurs, périodiques ou sensibles aux conditions initiales ; ces attracteurs sont stochastiquement stables et l'union de leurs bassins est de mesure totale. Cette conjecture, formulée pour la première fois au début de 1995, fournit un schéma pour la compréhension des familles paramétrées de systèmes dynamiques. On peut la considérer comme une version probabiliste d'un vieux rêve des années soixante, l'existence d'un ouvert dense de systèmes dynamiquement stables, rêve qui s'était évanoui à la fin de cette décade. Seul le cas de la dimension 1 a survécu : Swiatek avec l'aide de Graczyk ([GS]) et Lyubich l'ont indépendamment établi pour la famille quadratique réelle ; plus récemment Kozlovski ([Ko]) a annoncé le même résultat pour la famille des applications unimodales de classe C^3.

Pour les applications unidimensionnelles réelles ou complexes, notre conjecture est plus précise, prédisant pour la plupart des valeurs des paramètres un nombre fini d'attracteurs, qui sont périodiques ou supportent une mesure de probabilité invariante absolument continue. Remarquablement, Lyubich ([Ly2]) vient avec l'aide de Martens et Nowicki (*cf.* ce volume) d'établir ce résultat pour la famille des polynômes quadratiques réels.

Soit f une fraction rationnelle hyperbolique. On suppose que son ensemble de Julia $\mathcal{J}$ n'est pas connexe. Nous allons montrer que, à l'exception d'un nombre fini de composantes périodiques de $\mathcal{J}$, et la collection dénombrable de leurs composantes préimages, toute composante de $\mathcal{J}$ est soit un point soit une courbe de Jordan. Par conséquent, toute composante de $\mathcal{J}$ est localement connexe. Nous discutons également quand une telle courbe de Jordan est aussi un quasi-cercle. Nous donnerons un exemple explicite d'une fraction rationnelle ayant une composante de Julia qui est une courbe de Jordan mais pas un quasi-cercle.

Soit f un polynôme dont l'ensemble de Julia contient un seul point critique c. Nous montrons que si le bassin de l'infini est Hölderien, la condition de Collet-Eckmann est vérifiée : il existe $\lambda > 1, C > 0$ tel qu'on ait $|(f^n)'(f(c))| \geq C\lambda^n$ pour tout $n \geq 0$. Nous introduisons également les notions d'application rationnelle de type topologique Collet-Eckmann et de répulseur.

Nous présentons une nouvelle démonstration du fait que tous les rayons externes à arguments rationnels de l'ensemble de Mandelbrot aboutissent et nous relions l'argument externe d'un tel rayon à la dynamique du paramètre où le rayon aboutit. Notre démonstration est différente de celle donnée à l'origine par Douady et Hubbard et élaborée par P. Lavaurs : elle remplace des arguments analytiques par des arguments combinatoires ; elle n'utilise pas la dépendance analytique des polynômes par rapport au paramètre et peut donc être appliquée aux espaces de paramètres qui ne sont pas analytiques complexes ; la démonstration est aussi techniquement plus facile. Finalement, nous déduisons quelques corollaires sur les composantes hyperboliques de l'ensemble Mandelbrot.

Chemin faisant, nous construisons des partitions du plan dynamique et de l'espace des paramètres, intéressantes en elles-mêmes, et nous interprétons l'ensemble de Mandelbrot comme un espace de paramètres symboliques contenant des *kneading sequences* et des adresses internes.

ABSTRACTS

The purpose of this note is twofold : to discuss the concept of holomorphic motions and phenomena of Mañé-Sad-Sullivan type in several complex variables and secondly, to compare the different notions of Beltrami differentials in CR-geometry which have appeared in [4] and [7].

Hénon maps for which the analysis in [BC2] applies are considered. Sets with good hyperbolic properties and nice return structures are constructed and their return time functions are shown to have exponentially decaying tails. This sets the stage for applying the results in [Y]. Statistical properties such as exponential decay of correlations and central limit theorem are proved.

We regard the problem of classification for complex projective foliations with algebraic limit sets and prove the following :

Let $\mathcal{F}$ be a holomorphic foliation by curves in the complex projective plane $\mathbb{C}P(2)$ having as limit set some singularities and an algebraic curve $\Lambda \subset \mathbb{C}P(2)$. If the singularities $\operatorname{sing}\mathcal{F} \cap \Lambda$ are generic then either $\mathcal{F}$ is given by a closed rational 1-form or it is a rational pull-back of a Riccati foliation $\mathcal{R} : p(x)dy - (a(x)y^2 + b(x)y)dx = 0$, where Λ corresponds to $\overline{(y=0)} \cup \overline{(p(x)=0)}$, on $\overline{\mathbb{C}} \times \overline{\mathbb{C}}$.

The proof is based on the solvability of the generalized holonomy groups associated to a reduction process of the singularities $\operatorname{sing}\mathcal{F} \cap \Lambda$ and the construction of an affine transverse structure for $\mathcal{F}$ outside an algebraic curve containing Λ.

We give a lower bound for the volume of a compact convex domain in a Euclidean space with boundary of class $C^{1,1}$. We characterize the equality case.

We prove the existence of Sinai-Ruelle-Bowen measures for a class of C^2 self-mappings of a rectangle with unbounded derivatives. The results can be regarded as a generalization of a well-known one dimensional Folklore Theorem on the existence of absolutely continuous invariant measures. In an earlier paper [8] analogous results were stated and the proofs were sketched for the case of invertible systems. Here we give complete proofs in the more general case of noninvertible systems, and, in particular, develop the theory of stable and unstable manifolds for maps with unbounded derivatives.

In this paper we shall consider real polynomials with one (possibly degenerate) non-escaping critical (folding) point. Necessary and sufficient conditions are given for the total disconnectedness of the Julia set of such polynomials. Also we prove that the Julia sets of such polynomials do not carry invariant linefields. In the real case, this generalises the results by Branner and Hubbard for cubic polynomials and by McMullen on absence of invariant linefields.

This is a continuation of notes on the dynamics of quadratic polynomials. In this part we transfer our previous geometric result [L3] to the parameter plane. To any parameter value c (outside the main cardioid and the little Mandelbrot sets attached to it) we associate a "principal nest of parapuzzle pieces". We then prove that the moduli of the annuli between two consecutive pieces grow at least linearly. This implies, using Martens & Nowicki (*cf.* this volume) geometric criterion for existence of an absolutely continuous invariant measure together with [L2], that Lebesgue almost every real quadratic polynomial is either hyperbolic, or has a finite absolutely continuous invariant measure, or is infinitely renormalizable. In the further papers [L5,L7] we show that the latter set has zero Lebesgue measure, which completes the measure-theoretic picture of the dynamics in the real quadratic family.

We introduce a class of one-parameter families of real maps extending the classical geometric Lorenz models. These families combine singular dynamics (discontinuities with infinite derivative) with critical dynamics (critical points) and are based on the behaviour displayed by Lorenz flows over a fairly wide range of parameters. Our main result states that – nonuniform – expansion is the prevalent form of dynamics even after the formation of the criticalities.

A sufficient geometrical condition for the existence of absolutely continuous invariant probability measures for $S-$unimodal maps will be discussed. The Lebesgue typical existence of Sinai-Bowen-Ruelle-measures in the quadratic family will be a consequence.

We study the analytic dependance of separatrices for an equisingular family of germs of holomorphic foliation at the origin of $\mathbb{C}^2$. We define the quasi-homogeneity by a rigidity property. We obtain a K. Saito type Theorem for quasi-homogeneous foliations and a of Briançon-Skoda type Theorem in the general case.

A presentation of some fundamental results from the Douady-Hubbard theory of the Mandelbrot set, based on the idea of "orbit portrait" : the pattern of external rays landing on a periodic orbit for a quadratic polynomial map.

A view on dissipative dynamics, i.e. flows, diffeomorphisms, and transformations in general of a compact boundaryless manifold or the interval is presented here, including several recent results, open problems and conjectures. It culminates with a conjecture on the denseness of systems having only finitely many attractors, the attractors being sensitive to initial conditions (chaotic) or just periodic sinks and the union of their basins of attraction having total probability. Moreover, the attractors should be stochastically stable in their basins of attraction. This formulation, dating from early 1995, sets the scenario for the understanding of most nearby systems in parametrized form. It can be considered as a probabilistic version of the once considered possible existence of an open and dense subset of systems with dynamically stable structures, a dream of the sixties that evaporated by the end of that decade. The collapse of such a previous conjecture excluded the case of one dimensional dynamics : it is true at least for real quadratic maps of the interval as shown independently by Swiatek, with the help of Graczyk [GS], and Lyubich [Ly1] a few years ago. Recently, Kozlovski [Ko] announced the same result for C^3 unimodal mappings, in a meeting at IMPA. Actually, for one-dimensional real or complex dynamics, our main conjecture goes even further : for most values of parameters, the corresponding dynamical system displays finitely many attractors which are periodic sinks or carry an absolutely continuous invariant probability measure.

Remarkably, Lyubich [Ly2] has just proved this for the family of real quadratic maps of the interval, with the help of Martens and Nowicki [MN].

We show that if f is a hyperbolic rational map with disconnected Julia set $\mathcal{J}$, then with the possible exception of finitely many periodic components of $\mathcal{J}$ and their countable collection of preimages, every connected component of $\mathcal{J}$ is a point or a Jordan curve. As a corollary, every component of $\mathcal{J}$ is locally connected. We also discuss when a Jordan curve Julia component is a quasicircle and give an explicit example of a hyperbolic rational map with a Jordan curve Julia component which is not a quasicircle.

We prove that for every polynomial f if its basin of attraction to ∞ is Hölder and Julia set contains only one critical point c then f is Collet-Eckmann, namely there exists $\lambda > 1$, $C > 0$ such that, for every $n \geq 0$, $|(f^n)'(f(c))| \geq C\lambda^n$. We introduce also *topological Collet-Eckmann rational maps* and *repellers.*

We give a new proof that all external rays of the Mandelbrot set at rational angles land, and of the relation between the external angle of such a ray and the dynamics at the landing point. Our proof is different from the original one, given by Douady and Hubbard and refined by P. Lavaurs, in several ways : it replaces analytic arguments by combinatorial ones ; it does not use complex analytic dependence of the polynomials with respect to parameters and can thus be made to apply for non-complex analytic parameter spaces ; this proof is also technically simpler. Finally, we derive several corollaries about hyperbolic components of the Mandelbrot set.

Along the way, we introduce partitions of dynamical and parameter planes which are of independent interest, and we interpret the Mandelbrot set as a symbolic parameter space of kneading sequences and internal addresses.

PRÉFACE

Le colloque « Géométrie complexe et Systèmes dynamiques » s'est tenu à Orsay du 3 au 8 juillet 1995. Organisé en l'honneur d'Adrien Douady, à l'occasion de son soixantième anniversaire, il a réuni environ 250 participants.

Elève d'Henri Cartan, Adrien Douady a d'abord consacré ses recherches à la géométrie analytique : la structure analytique universelle sur l'ensemble des sous-espaces analytiques fermés d'un espace analytique compact, la déformation universelle d'un espace analytique, le principe « platitude et privilège » sont les points saillants de ce premier versant de son œuvre mathématique. Vient ensuite une période intermédiaire où ses intérêts sont très divers : densité des formes de Strebel, dimension des attracteurs (avec J. Oesterlé), théorème de Manin-Drinfeld...

À partir de 1980 s'ouvre l'époque de la dynamique holomorphe. L'étude de l'itération des fractions rationnelles initiée par Pierre Fatou et Gaston Julia au début du siècle était tombée en désuétude jusqu'à la fin des années soixante. C'est alors que les expériences numériques, rendues possibles par la puissance des ordinateurs, vont provoquer un extraordinaire regain d'activité dans cette direction, qui jusqu'à aujourd'hui ne s'est pas démenti. Sur le plan théorique, ce sont les résultats fondamentaux de D. Sullivan d'une part, de A. Douady et J. Hubbard d'autre part, qui vont ouvrir la voie.

La théorie des applications à allure polynomiale fournit le cadre naturel qui assure la flexibilité nécessaire à l'étude de l'itération des polynômes. Les techniques de chirurgie holomorphe ont permis de belles avancées. La description combinatoire, par rayons externes et équipotentielles, des ensembles de Julia et de l'ensemble de Mandelbrot recèle une puissance impressionnante et une harmonie magique.

Le présent volume ne prétend pas être un reflet fidèle du colloque d'Orsay. Certains auteurs n'avaient pu être présents, d'autres ont écrit sur un sujet différent de leur exposé oral. Si la dynamique holomorphe est très présente, d'autres domaines extrêmement actifs à l'heure actuelle (dynamique non-uniformément hyperbolique, feuilletages...) y apparaissent, l'ensemble reflétant la richesse, l'unité et la diversité des recherches contemporaines sur les systèmes dynamiques.