STATISTICAL PROPERTIES OF UNIMODAL MAPS: THE QUADRATIC FAMILY

ARTUR AVILA AND CARLOS GUSTAVO MOREIRA

Abstract. We prove that almost every non-regular real quadratic map is Collet-Eckmann and has polynomial recurrence of the critical orbit (proving a conjecture by Sinai). It follows that typical quadratic maps have excellent ergodic properties, as exponential decay of correlations (Keller and Nowicki, Young) and stochastic stability in the strong sense (Baladi and Viana). This is an important step to get the same results for more general families of unimodal maps.

Contents

1. Introduction

Here we consider the quadratic family, $f_a = a - x^2$, where $-1/4 \leq a \leq 2$ is the parameter, and we analyze its dynamics in the invariant interval.

The quadratic family has been one of the most studied dynamical systems in the last decades. It is one of the most basic examples and exhibits a very rich behavior. It was also studied through many different techniques. Here we are interested in describing the dynamics of a typical quadratic map from the statistical point of view.

Date: February 24, 2002.

Partially supported by Faperj and CNPq, Brazil.

1.1. **The probabilistic point of view in dynamics.** In the last decade Palis [Pa] described a general program for (dissipative) dynamical systems in any dimension. In short, 'typical' dynamical systems can be modeled stochastically in a robust way. More precisely, one should show that such typical system can be described by finitely many attractors, each of them supporting a (ergodic) physical measure: time averages of Lebesgue almost every orbit should converge to spatial averages according to one of the physical measures. The description should be robust under (sufficiently) random perturbations of the system: one asks for stochastic stability.

Moreover, a typical dynamical system was to be understood in the Kolmogorov sense: as a set of full measure in generic parametrized families.

Besides the questions posed by this conjecture, much more can be asked about the statistical description of the long term behavior of a typical system. For instance, the definition of physical measure is related to the validity of the Law of Large Numbers. Are other theorems still valid, like the Central Limit or Large Deviation theorems? Those questions are usually related to the rates of mixing of the physical measure.

1.2. **The richness of the quadratic family.** While we seem still very far away of any description of dynamics of typical dynamical systems (even in one-dimension), the quadratic family has been a remarkable exception. We describe briefly results which show the richness of the quadratic family from the probabilistic point of view.

The initial step in this direction was the work of Jakobson [J], where it was shown that for a positive measure set of parameters the behavior is stochastic, more precisely, there is an absolutely continuous invariant measure (the physical measure) with positive Lyapunov exponent: for Lebesgue almost every x, $|Df^n(x)|$ grows exponentially fast. On the other hand, it was later shown by Lyubich [L2] and Graczyk-Swiatek [GS] that regular parameters (with a periodic hyperbolic attractor) are (open and) dense. While stochastic parameters are predominantly expanding (in particular have sensitive dependence to initial conditions), regular parameters are deterministic (given by the periodic attractor). So at least two kinds of very distinct observable behavior are present in the quadratic family, and they alternate in a complicate way.

It was later shown that stochastic behavior could be concluded from enough expansion along the orbit of the critical value: the Collet-Eckmann condition, exponential growth of $|Df^n(f(0))|$, was enough to conclude a positive Lyapunov exponent of the system. A different approach to Jakobson's Theorem in [BC1] and [BC2] focused specifically on this property: the set of Collet-Eckmann maps has positive measure. After these initial works, many others studied such parameters (sometimes with extra assumptions), obtaining refined information of the dynamics of CE maps, particularly information about exponential decay of correlations[1] (Keller and Nowicki in [KN] and Young in [Y]), and stochastic stability (Baladi and Viana in [BV]). The dynamical systems considered in those papers have generally been shown to have excellent statistical description[2].

[1]CE quadratic maps are not always mixing and finite periodicity can appear in a robust way. This phenomena is related to the map being renormalizable, and this is the only obstruction: the system is exponentially mixing after renormalization.

[2]It is now known that weaker expansion than Collet-Eckmann is enough to obtain stochastic behavior for quadratic maps, on the other hand, exponential decay of correlations is actually

Many of those results also generalized to more general families and sometimes to higher dimensions, as in the case of Hénon maps [BC2].

The motivation behind all the results on CE maps is certainly the fact that they were valid for a class of maps that was known to have positive measure. It was known however that very different (sometimes wild) behavior coexisted ([Jo] and [HK]). It was shown for instance the existence of parameters without a physical measure or such that the physical measure is concentrated on a repelling hyperbolic fixed point. It remained to see if wild behavior was observable.

In a big project in the last decade, Lyubich [L3] together with Martens and Nowicki [MN] showed that almost all parameters have physical measures: more precisely, besides regular and stochastic behavior, only one more possibility could happen with positive measure, namely infinitely renormalizable maps (which always have a uniquely ergodic physical measure). Later Lyubich in [L5] showed that infinitely renormalizable parameters have measure zero, thus concluding the regular or stochastic duality, giving a further advance in the comprehension of the nature of the statistical behavior of typical quadratic maps. This celebrated result is remarkably linked to the progress obtained by Lyubich on the answer of the Feigenbaum conjectures ([L4]).

1.3. **Statements of the results.** In this work we describe the asymptotic behavior of the critical orbit. Our first result is an estimate of hyperbolicity:

Theorem A. *Almost every non-regular real quadratic map satisfies the Collet-Eckmann condition:*

$$\liminf_{n\to\infty} \frac{\ln(|Df^n(f(0))|)}{n} > 0.$$

The second is an estimate on the recurrence of the critical point. For regular maps, the critical point is non-recurrent (it actually converges to the periodic attractor). Among non-regular maps, however, the recurrence occurs at a precise rate which we estimate:

Theorem B. *Almost every non-regular real quadratic map has polynomial recurrence of the critical orbit with exponent* 1*:*

$$\limsup_{n\to\infty} \frac{-\ln(|f^n(0)|)}{\ln(n)} = 1.$$

In other words, the set of of n such that $|f^n(0)| < n^{-\gamma}$ is finite if $\gamma > 1$ and infinite if $\gamma < 1$.

As far as we know, this is the first proof of polynomial estimates for the recurrence of the critical orbit valid for a positive measure set of non-hyperbolic parameters (although subexponential estimates were known before). This also answers a long standing conjecture of Sinai.

Theorems A and B show that a typical non regular quadratic maps have enough good properties to conclude the results on exponential decay of correlations (which can be used to prove Central Limit and Large Deviation theorems) and stochastic stability in the sense of L^1 convergence of the densities (of stationary measures of perturbed systems). Many other properties also follow, like existence of a spectral

equivalent to the CE condition [NS], and all current results on stochastic stability use the Collet-Eckmann condition.

gap in [KN] and the recent results on almost sure (stretched exponential) rates of convergence to equilibrium in [BBM]. In particular, this answers positively Palis conjecture for the quadratic family.

1.4. **Unimodal maps.** Another reason to deal with the quadratic family is that it seems to open the doors to the understanding of unimodal maps. Its universal behavior was first realized in the topological sense, with Milnor-Thurston theory. The Feigenbaum-Coullet-Tresser observations indicated a geometric universality ([L4]).

A first result in the understanding of measure-theoretical universality was the work of Avila, Lyubich and de Melo [ALM], where it was shown how to relate metrically the parameter spaces of non-trivial analytic families of unimodal maps to the parameter space of the quadratic family. This was proposed as a method to relate observable dynamics in the quadratic family to observable dynamics of general analytic families of unimodal maps. In that work the method is used successfully to extend the regular or stochastic dichotomy to this broader context.

We are also able to adapt those methods to our setting. The techniques developed here and the methods of [ALM] are the main tools used in [AM] to obtain the main results of this paper (except the exact value of the polynomial recurrence) for non-trivial real analytic families of unimodal maps (with negative Schwarzian derivative and quadratic critical point). This is a rather general set of families, as trivial families form a set of infinite codimension. For a different approach (still based in [ALM]) which does not use negative Schwarzian derivative and obtains the exponent 1 for the polynomial recurrence, see [A].

In [AM] we also prove a version of Palis conjecture in the smooth setting. There is a residual set of k-parameter C^3 (for the equivalent C^2 result, see [A]) families of unimodal maps with negative Schwarzian derivative such that almost every parameter is either regular or Collet-Eckmann with subexponential bounds for the recurrence of the critical point. The final steps use techniques that date back to Jakobson, in a form advanced by Tsujii. Those techniques are needed to deal in the non-analytic setting.

Acknowledgements: We thank Viviane Baladi, Mikhail Lyubich and Marcelo Viana for helpful discussions about this work. We are grateful to Juan Rivera-Letelier for a valuable suggestion which led to the concept of gape renormalization and for listening to a first version.

2. General definitions

2.1. **Maps of the interval.** Let $f : I \to I$ be a C^1 map defined on some interval $I \subset \mathbb{R}$. The orbit of a point $p \in I$ is the sequence $\{f^k(p)\}_{k=0}^{\infty}$. We say that p is recurrent if there exists a subsequence $n_k \to \infty$ such that $\lim f^{n_k}(p) = p$.

We say that p is a periodic point of period n of f if $f^n(p) = p$, and $n \geq 1$ is minimal with this property. In this case we say that p is hyperbolic if $|Df^n(p)|$ is not 0 or 1. Hyperbolic periodic orbits are attracting or repelling according to $|Df^n(p)| < 1$ or $|Df^n(p)| > 1$.

We will often consider the restriction of iterates f^n to intervals $T \subset I$, such that $f^n|_T$ is a diffeomorphism. In this case we will be interested on the distortion of $f^n|_T$,

$$\operatorname{dist}(f^n|_T) = \frac{\sup_T |Df^n|}{\inf_T |Df^n|}.$$

This is always a number bigger than or equal to 1, we will say that it is small if it is close to 1.

2.2. **Trees.** We let Ω denote the set of finite sequences of non-zero integers (including the empty sequence). Let Ω_0 denote Ω without the empty sequence.

We denote $\sigma^+ : \Omega_0 \to \Omega$ by $\sigma^+(j_1, ..., j_m) = (j_1, ..., j_{m-1})$ and $\sigma^- : \Omega_0 \to \Omega$ by $\sigma^-(j_1, ..., j_m) = (j_2, ..., j_m)$.

2.3. **Growth of functions.** Let X be a class of functions $g : \mathbb{N} \to \mathbb{R}$ such that $\lim_{n\to\infty} g(n) = \infty$. We say that a function $f : \mathbb{N} \to \mathbb{R}$ grows at least rate X if there exists a function $g \in X$ such that $f(n) \geq g(n)$ for n sufficiently big. We say that it grows at rate X if there are $g_1, g_2 \in X$ such that $g_1(n) \leq f(n) \leq g_2(n)$ for n sufficiently big. We say that f decreases with rate (at least) X if $1/f$ grows at rate (at least) X.

Standard classes are the following. Linear for linear functions with positive slope. Polynomial for functions $g(n) = n^k, k > 0$. Exponential for functions $g(n) = e^{kn}, k > 0$.

The standard torrential function T is defined recursively by $T(1) = 1, T(n+1) = 2^{T(n)}$. The torrential class is the set of functions $g(n) = T(\max\{n + k, 1\}), k \in \mathbb{Z}$.

Torrential growth can be detected from recurrent estimates easily. A sufficient condition for a function which is unbounded from above to grow at least torrentially is an estimate as

$$f(n+1) > e^{f(n)^a}$$

for some $a > 0$. Torrential growth is implied by an estimate as

$$e^{f(n)^a} < f(n+1) < e^{f(n)^b}$$

with $0 < a < b$.

2.4. **Quasisymmetric maps.** Let $k \geq 1$ be given. We say that a homeomorphism $f : \mathbb{R} \to \mathbb{R}$ is quasisymmetric with constant k if for all $h > 0$

$$\frac{1}{k} \leq \frac{f(x+h) - f(x)}{f(x) - f(x-h)} \leq k.$$

The space of quasisymmetric maps is a group under composition, and the set of quasisymmetric maps with constant k preserving a given interval is compact in the uniform topology of compact subsets of $\mathbb{R}$. It also follows that quasisymmetric maps are Hölder.

To describe further the properties of quasisymmetric maps, we need the concept of quasiconformal maps and dilatation (see §4.2 for definitions), so we just mention a result of Ahlfors-Beurling which connects both concepts: any quasisymmetric map extends to a quasiconformal real-symmetric map of $\mathbb{C}$ and, conversely, the restriction of a quasiconformal real-symmetric map of $\mathbb{C}$ to $\mathbb{R}$ is quasisymmetric. Furthermore, it is possible to work out upper bounds on the dilatation (of an optimal extension) depending only on k and conversely.

The constant k is awkward to work with: the inverse of a quasisymmetric map with constant k may have a larger constant. We will therefore work with a less standard constant: we will say that h is γ-quasisymmetric (γ-qs) if h admits a quasiconformal symmetric extension to $\mathbb{C}$ with dilatation bounded by γ. This definition behaves much better: if h_1 is γ_1-qs and h_2 is γ_2-qs then $h_2 \circ h_1$ is $\gamma_2\gamma_1$-qs.

If $X \subset \mathbb{R}$ and $h : X \to \mathbb{R}$ has a γ-quasisymmetric extension to $\mathbb{R}$ we will also say that h is γ-qs.

Let $QS(\gamma)$ be the set of γ-qs maps of $\mathbb{R}$.

3. Real quadratic maps

If $\lambda \in \mathbb{C}$ we let $f_\lambda : \mathbb{C} \to \mathbb{C}$ denote the (complex) quadratic map $\lambda - z^2$. If $\lambda \in \mathbb{R}$ is such that $-1/4 \leq \lambda \leq 2$ there exists an interval $I_\lambda = [\beta, -\beta]$ with

$$\beta = \frac{-1 - \sqrt{1 + 4\lambda}}{2}$$

such that $f_\lambda(I_\lambda) \subset I_\lambda$ and $f_\lambda(\partial I_\lambda) \subset \partial I_\lambda$. For such a λ, the map $f = f_\lambda|_{I_\lambda}$ is unimodal, that is, it is a self map of I_λ with a unique turning point.

We will keep the notation f_λ to refer to a quadratic map when we discuss its complex extension and f to denote a fixed quadratic map when we discuss its unimodal restriction, and we let $I = I_\lambda$.

3.1. **The combinatorics of unimodal maps.**

In this subsection we fix a real quadratic map f and define some objects related to it.

3.1.1. *Return maps.* Given an interval $T \subset I$ we define the first return map $R_T : X \to T$ where $X \subset T$ is the set of points x such that there exists $n > 0$ with $f^n(x) \in T$, and $R_T(x) = f^n(x)$ for the minimal n with this property.

3.1.2. *Nice intervals.* An interval T is nice if it is symmetric around 0 and the iterates of ∂T never intersect $\operatorname{int} T$. Given a nice interval T we notice that the domain of the first return map R_T decomposes in a union of intervals T^j, indexed by integer numbers (if there are only finitely many intervals, some indexes will be corresponded to the empty set). If 0 belongs to the domain of R_T, we say that T is proper. In this case we reserve the index 0 to denote the component of the critical point: $0 \in T^0$.

If T is nice, it follows that for all $j \in \mathbb{Z}$, $R_T(\partial T^j) \subset \partial T$. In particular, $R_T|_{T^j}$ is a diffeomorphism onto T unless $0 \in T^j$ (and in particular $j = 0$ and T is proper). If T is proper, $R_T|_{T^0}$ is symmetric (even) with a unique critical point 0. As a consequence, T^0 is also a nice interval.

If $R_T(0) \in T^0$, we say that R_T is central.

3.1.3. *Landing maps.* Given a proper interval T we define the landing map $L_T : X \to T^0$ where $X \subset T$ is the set of points x such that there exists $n \geq 0$ with $f^n(x) \in T^0$, and $L_T(x) = f^n(x)$ for the minimal n with this property. We notice that $L_T|_{T^0} = \mathrm{id}$.

3.1.4. *Trees.* If T is a proper interval, the first return map to T naturally relates to the first landing to T^0 in the following way.

If $\underline{d} \in \Omega$, we define $T^{\underline{d}}$ inductively in the following way. We let $T^{\underline{d}} = T$ if $\underline{d}$ is empty and if $\underline{d} = (j_1, ..., j_m)$ we let $T^{\underline{d}} = (R_T|_{T^{j_1}})^{-1}(T^{\sigma^-(\underline{d})})$.

We denote $R_T^{\underline{d}} = R_T^{|\underline{d}|}|_{T^{\underline{d}}}$ which is always a diffeomorphism over T.

We denote by $C^{\underline{d}} = (R_T^{\underline{d}})^{-1}(T^0)$. We notice that the domain of the first landing map L_T coincides with the union of the $C^{\underline{d}}$, and furthermore $L_T|_{C^{\underline{d}}} = R_T^{\underline{d}}$.

Notice that this allows us to relate R_T to R_{T^0} since $R_{T^0} = L_T \circ R_T$.

3.1.5. *Renormalization.* We say that f is renormalizable if there is an interval T and $m > 1$ such that $f^m(T) \subset T$ and $f^j(\operatorname{int} T) \cap \operatorname{int} T = \emptyset$ for $1 \leq j < m$. The maximal such interval is called the renormalization interval of period m, it has the property that $f^m(\partial T) \subset \partial T$.

The set of renormalization periods of f gives an increasing (possibly empty) sequence of numbers m_i, $i = 1, 2, \ldots$, each related to a unique renormalization interval $T^{(i)}$ which form a nested sequence of intervals. We include $m_0 = 1$, $T^{(0)} = I$ in the sequence to simplify the notation.

We say that f is finitely renormalizable if there is a smallest renormalization interval $T^{(k)}$. We say that $f \in \mathcal{F}$ if f is finitely renormalizable and 0 is recurrent but not periodic. We let $\mathcal{F}_k$ denote the set of maps f in $\mathcal{F}$ which are exactly k times renormalizable.

3.1.6. *Principal nest.* Let Δ_k denote the set of all maps f which have (at least) k renormalizations and which have an orientation reversing non-attracting periodic point of period m_k which we denote p_k (that is, p_k is the fixed point of $f^{m_k}|_{T^{(k)}}$ p_k with $Df^{m_k}(p_k) \leq -1$). For $f \in \Delta_k$, we denote $T_0^{(k)} = [-p_k, p_k]$. We define by induction a (possibly finite) sequence $T_i^{(k)}$, such that $T_{i+1}^{(k)}$ is the component of the domain of $R_{T_i^{(k)}}$ containing 0. If this sequence is infinite, then either it converges to a point or to an interval.

If $\cap_i T_i^{(k)}$ is a point, then f has a recurrent critical point which is not periodic, and it is possible to show that f is not $k+1$ times renormalizable. Obviously in this case we have $f \in \mathcal{F}_k$, and all maps in $\mathcal{F}_k$ are obtained in this way: if $\cap_i T_i^{(k)}$ is an interval, it is possible to show that f is $k+1$ times renormalizable.

We can of course write $\mathcal{F}$ as a disjoint union $\cup_{i=0}^{\infty} \mathcal{F}_i$. For a map $f \in \mathcal{F}_k$ we refer to the sequence $\{T_i^{(k)}\}_{i=1}^{\infty}$ as the principal nest.

It is important to notice that the domain of the first return map to $T_i^{(k)}$ is always dense in $T_i^{(k)}$. Moreover, the next result shows that, outside a very special case, the return map has a hyperbolic structure.

Lemma 3.1. *Assume $T_i^{(k)}$ does not have a non-hyperbolic periodic orbit in its boundary. For all $T_i^{(k)}$ there exists $C > 0$, $\lambda > 1$ such that if $x, f(x), \ldots, f^{n-1}(x)$ do not belong to $T_i^{(k)}$ then $|Df^n(x)| > C\lambda^n$.*

This lemma is a simple consequence of a general theorem of Guckenheimer on hyperbolicity of maps of the interval without critical points and non-hyperbolic periodic orbits (Guckenheimer considers unimodal maps with negative Schwarzian derivative, so this applies directly to the case of quadratic maps, the general case is also true by Mañé's Theorem, see [MvS]). Notice that the existence of a non-hyperbolic periodic orbit in the boundary of $T_i^{(k)}$ depends on a very special combinatorial setting, in particular, all $T_j^{(k)}$ must coincide (with $[-p_k, p_k]$), and the k-th renormalization of f is in fact renormalizable of period 2.

It is easy to see that the first return map to $T_i^{(k)}$ has infinitely many components except in the special case above or if $i = 0$. This lemma shows that the complement of the domains of the return map to $T_i^{(k)}$ (outside the special case or if $i = 0$) form a regular Cantor set. This has many useful consequences (for instance, the image of a regular Cantor set by a quasisymmetric map has always zero Lebesgue measure).

3.1.7. *Lyubich's Regular or Stochastic dichotomy.* A map $f \in \mathcal{F}_k$ is called *simple* if the principal nest has only finitely many central returns, that is, there are only finitely many i such that $R|_{T_i^{(k)}}$ is central. Such maps have many good features, in particular, they are stochastic (this is a consequence of [MN] and [L1]).

In [L3], it was proved that almost every quadratic map is either regular or simple or infinitely renormalizable. It was then shown in [L5] that infinitely renormalizable maps have 0 Lebesgue measure, which establishes the Regular or Stochastic dichotomy.

Due to Lyubich's results, we can completely forget about infinitely renormalizable maps, we just have to prove the claimed estimates for almost every simple map.

During our discussion, for notational reasons, we will fix a renormalization level κ, that is, we will only analyze maps in Δ_κ. This allows us to fix some convenient notation: given $g \in \Delta_\kappa$ we define $I_i[g] = T_i^{(\kappa)}[g]$, so that $\{I_i[g]\}$ is a sequence of intervals (possibly finite). We use the notation $R_i[g] = R_{I_i[g]}$, $L_i[g] = L_{I_i[g]}$ and so on (so that the domain of $R_i[g]$ is $\cup I_i^j[g]$ and the domain of $L_i[g]$ is $\cup C_i^{\underline{d}}[g]$). When doing phase analysis (working with fixed f) we usually drop the dependence on the map and write R_i for $R_i[f]$.

(Notice that, once we fix the renormalization level κ, for $g \in \Delta_\kappa$, the notation $I_i[g]$ stands for $T_i^{(\kappa)}[g]$, even if g is more than κ times renormalizable.)

3.1.8. *Strategy.* To motivate our next steps, let us describe the general strategy behind the proofs of Theorems A and B.

(1) We consider a certain set of non-regular parameters of full measure and describe (in a probabilistic way) the dynamics of the principal nest. This is our phase analysis.

(2) From time to time, we transfer the information from the phase space to the parameter, following the description of the parapuzzle nest which we will make in the next subsection. The rules for this correspondence are referred as phase-parameter relation (which will be proved using techniques of complex dynamics).

(3) This correspondence will allow us to exclude parameters whose critical orbit behaves badly (from the probabilistic point of view) at infinitely many levels of the principal nest. The phase analysis coupled with the phase-parameter relation will assure us that the remaining parameters have still full measure.

(4) We restart the phase analysis for the remaining parameters with extra information.

After many iterations of this procedure we will have enough information to tackle the problems of hyperbolicity and recurrence.

We will first tackle the problem of explaining and proving the phase-parameter relation, and we will delay all statistical arguments until §6.

3.2. **Parameter partition.** Part of our work is to transfer information from the phase space of some map $f \in \mathcal{F}$ to a neighborhood of f in the parameter space. This is done in the following way. We consider the first landing map L_i: the domain of this map induces a partition of the interval I_i. The complement of the domain of L_i is a hyperbolic Cantor set $K_i = I_i \setminus \cup C_i^{\underline{d}}$. This Cantor set persists in a small parameter neighborhood J_i of f, changing in a continuous way.

Along J_i, the first landing map is topologically the same (in a way that will be clear soon). However the critical value $R_i[g](0)$ moves relative to the partition (when g moves in J_i). This allows us to partition the parameter piece J_i in smaller pieces, each corresponding to a region where $R_i(0)$ belongs to some fixed domain of the first landing map.

Theorem 3.2 (Topological Phase-Parameter relation). *Let $f \in \mathcal{F}_\kappa$. There is a sequence $\{J_i\}_{i\in\mathbb{N}}$ of nested parameter intervals (the principal parapuzzle nest of f) with the following properties.*

1. *J_i is the maximal interval containing f such that for all $g \in J_i$ the interval $T^{(\kappa)}_{i+1}[g]$ is defined and changes in a continuous way. (Since the first return map to $R_i[g]$ has a central domain, the landing map $L_i[g] : \cup C^{\underline{d}}_i[g] \to I_i[g]$ is defined.)*
2. *$L_i[g]$ is topologically the same along J_i: there exists homeomorphisms $H_i[g] : I_i \to I_i[g]$, such that $H_i[g](C^{\underline{d}}_i) = C^{\underline{d}}_i[g]$. The maps $H_i[g]$ change continuously.*
3. *There exists a homeomorphism $\Xi_i : I_i \to J_i$ such that $\Xi_i(C^{\underline{d}}_i)$ is the set of g such that $R_i[g](0)$ belongs to $C^{\underline{d}}_i[g]$.*

The homeomorphisms H_i and Ξ_i are not uniquely defined, it is easy to see that we can modify them inside each $C^{\underline{d}}_i$ window keeping the above properties. However, H_i and Ξ_i are well defined maps if restricted to K_i.

This fairly standard phase-parameter result can be proved in many different ways. The most elementary proof is probably to use the monotonicity of the quadratic family to deduce the Topological Phase-Parameter relation from Milnor-Thurston's kneading theory by purely combinatorial arguments. Another approach is to use Douady-Hubbard's description of the combinatorics of the Mandelbrot set (restricted to the real line) as does Lyubich in [L3] (whose results are partially reproduced in this paper, see Remark 4.6 in §4.7 for the connection with the Topological Phase-Parameter relation).

With this result we can define for any $f \in \mathcal{F}_\kappa$ intervals $J^j_i = \Xi_i(I^j_i)$ and $J^{\underline{d}}_i = \Xi_i(I^{\underline{d}}_i)$. From the description we gave it immediately follows that two intervals $J_{i_1}[f]$ and $J_{i_2}[g]$ associated to maps f and g are either disjoint or nested, and the same happens for intervals J^j_i or $J^{\underline{d}}_i$. Notice that if $g \in \Xi_i(C^{\underline{d}}_i) \cap \mathcal{F}_\kappa$ then $\Xi_i(C^{\underline{d}}_i) = J_{i+1}[g]$.

We will concentrate on the analysis of the regularity of Ξ_i for the special class of simple maps f: one of the good properties of the class of simple maps is better control of the phase-parameter relation. Even for simple maps, however, the regularity of Ξ_i is not great: there is too much dynamical information contained in it. A solution to this problem is to forget some dynamical information. With this intent we will introduce in §3.2.2 an interval which will be used to erase information.

3.2.1. *Geometric interpretation.* Before getting into those technical details, it will be convenient to make an informal geometric description of the topological statement we just made and discuss in this context the difficulties that will show up to obtain metric estimates.

The sequence of intervals J_i is defined as the maximal parameter interval containing f satisfying two properties: the dynamical interval I_i has a continuation (recall that the boundary of I_i is preperiodic, so the meaning of continuation is quite clear), and the first return map to this continuation has always the same combinatorics. Since it has the same combinatorics, the partition $C_i^{\underline{d}}$ also has a continuation along J_i.

Let us represent in two dimensions those continuations. Let $\mathcal{I}_i = \cup_{g \in J_i} \{g\} \times I_i[g]$ represent the "moving phase space" of R_i. It is a topological rectangle, its boundary consists of four analytic curves, the top and bottom (continuations of the boundary points of I_i) and the laterals (the limits of the continuations of I_i as the parameter converges to the boundary of J_i). Similarly, the continuations of each interval $C_i^{\underline{d}}$ form a strip $\mathcal{C}_i^{\underline{d}}$ inside $\mathcal{I}_i$. The resulting decomposition of $\mathcal{I}_i$ looks like a flag with countable many strips. The top and bottom boundaries of those strips (and the strips themselves) are horizontal in the sense that they connect one lateral if $\mathcal{I}_i$ to the other.

Remark 3.1. The boundaries of the strips are more formally described as forming a lamination in the topological rectangle $\mathcal{I}_i$, whose leaves are codimension-one, and indeed real analytic graphs over the first coordinate. In this sense, the present geometric description can be seen as motivation for §4: in the complex setting, the theory of codimension-one laminations is the same as the theory of holomorphic motions, described in §4.4, and which is the basis of the actual phase-parameter analysis.

Let us now look at the verticals $\{g\} \times I_i[g]$. They are all transversal to the strips of the flag, so we can consider the "horizontal" holonomy map between any two such verticals. If we fix one vertical as the phase-space of f while we vary the other, the resulting family of holonomy maps is exactly $H_i[g]$ as defined above.

Consider now the motion of $R_i[g](0)$ (the critical value of the first return map to the continuation of I_i) inside J_i, which we can represent by its graph $\mathcal{D} = \cup_{g \in J_i} \{g\} \times \{R_i[g](0)\}$. It is a diagonal to $\mathcal{I}_i$ in the sense that it connects a corner of the rectangle $\mathcal{I}_i$ to the opposite corner. In other words, if we vary continuously the quadratic map g (inside a slightly bigger parameter window then J_i), we see the window J_i appear when $g^{v_i}(0)$ enters $I_i[g]$ from one side and disappear when $g^{v_i}(0)$ escapes from the other side (where v_i is such that $R_i|_{I_i^0} = f^{v_i}$).

The main content of the Topological phase-parameter relation is that the motion of the critical value is not only a diagonal to $\mathcal{I}_i$ but to the flag: it cuts each strip exactly once in a monotonic way with respect to the partition. Thus, the diagonal motion of the critical point is transverse in a certain sense to the horizontal motion of the partition of the phase space (strips). The phase-parameter map is just the composition of two maps: the holonomy map between two transversals to the flag (from the "vertical" phase space of f to the diagonal $\mathcal{D}$) followed by projection on the first coordinate (from $\mathcal{D}$ to J_i).

Remark 3.2. One big advantage of complex analysis is that "transversality can be detected for topological reasons". So, while the statement that the critical point goes from the bottom to the top of $\mathcal{I}_i$ does not imply that it is transverse to all horizontal strips, the corresponding implication holds for the complex analogous of those statements. This is a consequence of the Argument Principle, see §4.4.9.

Let us now pay attention to the geometric format of those strips. The set $\cup_{g \in J_i} \{g\} \times R_i[g](I_i^0[g])$ is a topological triangle formed by the diagonal $\mathcal{D}$, one of the laterals of $\mathcal{I}_i$ (which we will call the right lateral[3]) and either the top or

[3]It is possible to prove that it is indeed located at the right side (with the usual ordering of the real line).

bottom of $\mathcal{I}_i$. In particular, the strip $\mathcal{I}_i^0$ is not a rectangle, but a triangle: the left side of $\mathcal{I}_i^0$ degenerates into a point. By their dynamical definition, all strips $C_i^{\underline{d}}$ also share the same property: the $C_i^{\underline{d}}[g]$ are collapsing as g converges to the left boundary of J_i. In particular the partition of the phase space of f must be metrically very different from the partition of the phase space of some g close to the left boundary of J_i.

This shows that it is not reasonable to expect the phase-parameter map $\Xi_i|_{K_i}$ to be very regular: if it was true that the phase-parameter relation is always regular, then the phase partitions of f and g would have to be metrically similar (since a correspondence between both partitions can be obtained as composition the phase-parameter relation for f and the inverse of the phase-parameter relation for g).

Let us now consider the decomposition of $\mathcal{I}_i$ in strips $\mathcal{I}_i^j$ (the continuations of I_i^j). This new flag is rougher than the previous one: each of its strips $\mathcal{I}_i^j$, $j \neq 0$ can be obtained as the (closure of the) union of $\mathcal{C}_i^{\underline{d}}$ where $\underline{d}$ starts with j. However, the strips are nicer: they are indeed rectangles if $j \neq 0$, though the "niceness" gets weaker and weaker as we get closer to the central strip. This suggests one way to obtain a regular map from Ξ_i: work with the rougher partition I_i^j outside of a certain small neighborhood of the critical strip (this neighborhood will be introduced in the next section, it will be called the gape interval). This procedure will indeed have the desired effect in the sense that we will be able to prove that for simple maps f the restriction of Ξ_i to $I_i \setminus \cup I_i^j$ has good regularity outside of the gape interval (this is PhPa2 in the Phase-Parameter relation below).

The resulting estimate does not say anything about what happens inside the rough partition by J_i^j. To do so, we consider the finer flag (whose strips are the $\mathcal{C}_i^{\underline{d}}$) intersected with the rectangles $\mathcal{Q}_i^j = \cup_{g \in J_i^j} \{g\} \times I_i^j[g]$ (those rectangles cover the diagonal $\mathcal{D}$ formed by the motion of the critical value). While the strips degenerate near the left boundary point of J_i, they intersect each $\mathcal{Q}_i^j$ in a nice rectangle (or the empty set). It will be indeed possible to prove that the phase-parameter map restricted to those rectangles is quite regular in the sense that if f is a simple map such that $f \in J_i^j$ (that is, $(f, R_n(0)) \in \mathcal{Q}_i^j$), the restriction of Ξ_i to I_i^j has good regularity (this is PhPa1 in the Phase-Parameter relation below).

3.2.2. *Gape interval.* If $i > 1$, we define the gape interval $\tilde{I}_{i+1}$ as follows.

We have that $R_i|_{I_{i+1}} = L_{i-1} \circ R_{i-1} = R_{i-1}^{\underline{d}} \circ R_{i-1}$ for some $\underline{d}$, so that $I_{i+1} = (R_{i-1}|_{I_i})^{-1}(C_{i-1}^{\underline{d}})$. We define the gape interval $\tilde{I}_{i+1} = (R_{i-1}|_{I_i})^{-1}(I_{i-1}^{\underline{d}})$.

We notice that for each I_i^j, the gape interval $\tilde{I}_{i+1}$ either contains or is disjoint from I_i^j.

3.2.3. *The phase-parameter relation.* As we discussed before, the dynamical information contained in Ξ_i is entirely given by $\Xi_i|_{K_i}$: a map obtained by Ξ_i by modification inside a $C_i^{\underline{d}}$ window has still the same properties. Therefore it makes sense to ask about the regularity of $\Xi_i|_{K_i}$. As we anticipated before we must erase some information to obtain good results.

Let f be a simple map and let τ_i be such that $R_i(0) \in I_i^{\tau_i}$. We define two Cantor sets, $K_i^\tau = K_i \cap I_i^{\tau_i}$ which contains refined information restricted to the $I_i^{\tau_i}$ window and $\tilde{K}_i = I_i \setminus \cup I_i^j \setminus \tilde{I}_{i+1}$, which contains global information, at the cost of erasing information inside each I_i^j window and in $\tilde{I}_{i+1}$.

Theorem 3.3 (Phase-Parameter relation)**.** *Let f be a simple map. For all $\gamma > 1$ there exists i_0 such that for all $i > i_0$ we have*

PhPa1: $\Xi_i|_{K_i^\tau}$ *is γ-qs,*
PhPa2: $\Xi_i|_{\tilde{K}_i}$ *is γ-qs,*
PhPh1: $H_i[g]|_{K_i}$ *is γ-qs if $g \in J_i^{\tau_i}$,*
PhPh2: *the map $H_i[g]|_{\tilde{K}_i}$ is γ-qs if $g \in J_i$.*

The proof of this Theorem will be given in §5, but involves mainly complex dynamics, so we now turn to discuss the complex analogous of the principal nest description of unimodal maps.

4. Complex dynamics

4.1. **Notation.** A Jordan curve T is a subset of $\mathbb{C}$ homeomorphic to a circle. A Jordan disk is a bounded open subset U of $\mathbb{C}$ such that ∂U is a Jordan curve.

We let $\mathcal{D}_r = \{z \in \mathbb{C} | |z| < r\}$ and we let $\mathbb{D} = \mathbb{D}_1$. If $r > 1$, let $A_r = \{z \in \mathbb{C} | 1 < |z| < r\}$.

An annulus A is a subset of $\mathbb{C}$ such that there exists a conformal map from A to some A_r. In this case, r is uniquely defined and we denote the moduli of A as $\mathrm{mod}(A) = \ln(r)$.

If U is a Jordan disk, we say that U is the filling of ∂U. A subset $X \subset U$ is said to be bounded by ∂U.

A graph of a continuous map $\phi : \Lambda \to \mathbb{C}$ is the set of all $(z, \phi(z)) \in \mathbb{C}^2$, $z \in \Lambda$.

Let $\mathbf{0} : \mathbb{C} \to \mathbb{C}^2$ be defined by $\mathbf{0}(z) = (z, 0)$.

Let $\pi_1, \pi_2 : \mathbb{C}^2 \to \mathbb{C}$ be the coordinate projections. Given a set $\mathcal{X} \subset \mathbb{C}^2$ we denote its fibers $X[z] = \pi_2(\mathcal{X} \cap \pi_1^{-1}(z))$.

A fiberwise map $\mathcal{F} : \mathcal{X} \to \mathbb{C}^2$ is a map such that $\pi_1 \circ \mathcal{F} = \pi_1$. We denote its fibers $F[z] : X[z] \to \mathbb{C}$ such that $\mathcal{F}(z, w) = (z, F[z](w))$.

4.2. **Quasiconformal maps.** Let $U \subset \mathbb{C}$ be a domain. A map $h : U \to \mathbb{C}$ is K-quasiconformal (K-qc) if it is a homeomorphism onto its image and for any annulus $A \subset U$, $\mathrm{mod}(A)/K \leq \mathrm{mod}(h(A)) \leq K\,\mathrm{mod}(A)$. The minimum such K is called the dilatation $\mathrm{Dil}(h)$ of h.

Let $h : X \to \mathbb{C}$ be a homeomorphism and let $C, \epsilon > 0$. An extension $H : U \to \mathbb{C}$ of h to a Jordan disk U is (C, ϵ)-qc if there exists an annulus $A \subset U$ with $\mathrm{mod}(A) > C$ such that X is contained in the bounded component of the complement of A and H is $1 + \epsilon$-qc. The following fact is well known:

Lemma 4.1. *For all $\gamma > 1$ there exists $C, \epsilon > 0$ such that if $h : X \to \mathbb{R}$ has a (C, ϵ)-qc extension $H : U \to \mathbb{C}$ which is real-symmetric then h is γ-qs.*

4.3. **Complex maps.**

4.3.1. *R-puzzle and R-maps.* A R-puzzle P is a family of Jordan disks (U, U^j) such that $\overline{U^j}$ are disjoint, $\overline{U^j} \subset U$ for all j and $0 \in U^0$.

If $P = (U, U_j)$ is a R-puzzle, a R-map (return type map) in P is a holomorphic map $R : \cup U^j \to U$, surjective in each component, such that for $j \neq 0$, $R|_{U^j}$ extends to a homeomorphism onto $\overline{U}$ and $R|_{U^0}$ extends to a double covering map onto $\overline{U}$ ramified at 0.

4.3.2. *From R-maps to L-maps.* Given a R-map R we induce an L-map (landing type map) as follows.

We define $U^{\underline{d}}$, $\underline{d} \in \Omega$ by induction on $|\underline{d}|$: if $\underline{d}$ is empty, we let $U^{\underline{d}} = U$, otherwise, if $\underline{d} = (j_1, ..., j_m)$, we let $U^{\underline{d}} = (R|_{U^{j_1}})^{-1}(U^{\sigma^-(\underline{d})})$.

Let $R^{\underline{d}} = R^{|\underline{d}|}|_{U^{\underline{d}}}$, and let $W^{\underline{d}} = (R^{\underline{d}})^{-1}(U^0)$.

The L-map associated to R is defined as $L(R) : \cup W^{\underline{d}} \to U^0$, $L(R)|_{W^{\underline{d}}} = R^{\underline{d}}$. Notice that $L(R)|_{W^{\underline{d}}}$ extends to a homeomorphism onto $\overline{U}$.

4.3.3. *Renormalization: from L-maps to R-maps.* Given a R-map R such that $R(0) \in W^{\underline{d}}$ we can define the (generalized in the sense of Lyubich) renormalization $N(R)$ by $N(R) = L(R) \circ R$ where defined in U^0: its domain is the R-puzzle (V, V^j) such that $V = U^0$ and the V^j are connected components of $(R|_{U^0})^{-1}(\cup W^{\underline{d}})$.

4.3.4. *Truncation and gape renormalization.* The $\underline{d}$ truncation of $L(R)$ is defined by $L^{\underline{d}}(R) = L(R)$ outside $U^{\underline{d}}$ and as $L^{\underline{d}}(R) = R^{\underline{d}}$ in $U^{\underline{d}}$.

Given a R-map R such that $R(0) \in U_0^{\underline{d}}$ we can defined the gape renormalization $G^{\underline{d}}(R)$ by $G^{\underline{d}}(R) = L^{\underline{d}}(R) \circ R$ where defined in U^0.

Notice that either $G^{\underline{d}}(R) = R|_{U^0}$ (if $\underline{d}$ is empty) or the domain of $G^{\underline{d}}(R)$ consists of countably many Jordan disks (with disjoint closures) where $G^{\underline{d}}(R)$ coincides with $N(R)$ and acts as a diffeomorphism onto U^0 and one central Jordan disk (containing 0) where $G^{\underline{d}}(R)$ acts as a double covering onto U.

4.3.5. *Complex extensions of unimodal maps.* The complex maps just introduced are designed to model complexifications of first return maps to intervals I_i of the principal nest. Indeed, let us fix a map $f \in \mathcal{F}_\kappa$ and let $U_1 \subset \mathbb{C}$ be a real-symmetric Jordan disk such that $U_1 \cap \mathbb{R} = \operatorname{int} I_1$. Under certain conditions on U_1 (which partially emulate the nice condition for unimodal maps), the first return map $R_1 : \cup I_1^j \to I_1$ extends to a puzzle map (still denoted R_1) from $\cup U_1^j$ to U_1. In this case, if we define inductively $R_{i+1} = N(R_i)$, $i \geq 1$, it is easy to see that the real trace of R_{i+1} coincides with the first return map from $\cup I_{i+1}^j$ to I_{i+1}, and the real trace of $L(R_{i+1})$ coincides with $L_{i+1} : \cup C_{i+1}^{\underline{d}} \to I_{i+1}^0$.

We did not introduce a unimodal analogous to gape renormalization. However, notice that if $R_i(0) \in C_i^{\underline{d}}$ then the central component of the domain of $G^{\underline{d}}(R_i)$ is a real-symmetric Jordan disk $(R_i|_{U_i^0})^{-1}(U_i^{\underline{d}})$ whose real trace is $\tilde{I}_{i+2}$.

There are several ways to construct the domain U_1 as above (for instance, Lyubich in [L2] uses the Yoccoz puzzle, but there are easier ways). Thus we can state:

Theorem 4.2. *Let $f \in \mathcal{F}_\kappa$. Then there exists a real-symmetric R-map R_1 such that the real trace of R_1 is the first return map (under f) to I_1.*

Once we have the complexification of the first return map to I_1, the complexifications of the first return maps to I_i, $i > 1$ are obtained by the renormalization procedure we just described.

The geometric estimate in the following theorem is (a special case of) one of the main results of the theory of quadratic maps, Theorem III of [L2].

Theorem 4.3. *Let $f \in \mathcal{F}_\kappa$. Let R_1 be as in Theorem 4.2 and let $R_{i+1} = N(R_i)$, $i \geq 1$ (in particular, the real trace of R_i coincides with the first return map to I_i). Then, if f is simple,* $\operatorname{mod}(U_i \setminus \overline{U_i^0})$ *grows at least linearly fast.*

Remark 4.1. In Theorem III of [L2], a much more general result is proved.

(1) Non-real quadratic maps are also considered.
(2) A geometric estimate is valid for more general maps then simple maps: if $n_k - 1$ counts the non-central levels of the principal nest, then $\operatorname{mod}(U_{n_k} \setminus \overline{U^0_{n_k}})$ grows at least linearly fast.
(3) The rate of linear growth is independent of the real quadratic map considered (a consequence of complex bounds).
(4) For complex maps, the rate of linear growth depends on finite combinatorial information.

4.4. **Holomorphic motions.**

4.4.1. *Preliminaries.* There are several ways to look at holomorphic motions. The way we describe them tries to make no commitment to a base point.

A holomorphic motion h over a domain Λ is a family of holomorphic maps defined on Λ whose graphs (called leaves of h) do not intersect. The support of h is the set $\mathcal{X} \subset \mathbb{C}^2$ which is the union of the leaves of h.

We have naturally associated maps $h[z] : \mathcal{X} \to X[z]$, $z \in \Lambda$ defined by $h[z](x, y) = w$ if (z, w) and (x, y) belong to the same leaf.

The transition (or holonomy) maps $h[z, w] : X[z] \to X[w]$, $z, w \in \Lambda$, are defined by $h[z, w](x) = h[w](z, x)$.

Given a holomorphic motion h over a domain Λ, a holomorphic motion h' over a domain $\Lambda' \subset \Lambda$ whose leaves are contained in leaves of h is called a restriction of h. If h is a restriction of h' we also say that h' is an extension of h.

Let $K : [0, 1) \to \mathbb{R}$ be defined by $K(r) = (1 + \rho)/(1 - \rho)$ where $0 \leq \rho < 1$ is such that the hyperbolic distance in $\mathbb{D}$ between 0 and ρ is r.

λ-Lemma ([MSS], [BR]) *Let h be a holomorphic motion over a hyperbolic domain Λ and let $z, w \in \Lambda$. Then $h[z, w]$ extends to a quasiconformal map of $\mathbb{C}$ with dilatation bounded by $K(r)$, where r is the hyperbolic distance between z and w in Λ.*

A completion of a holomorphic motion means an extension of h to the whole complex plane: $X[z] = \mathbb{C}$ for all $z \in \Lambda$.

Extension Lemma ([Sl]) *Any holomorphic motion over a simply connected domain can be completed.*

From now on, we will always assume that Λ is a Jordan disk.

4.4.2. *Meaning and notation warning.* Holomorphic motions are a convenient way to work with the dependence of the phase dynamics on the parameter. The analogous of the intervals $I_i[g]$ varying in intervals J_i (which we defined for real quadratic maps) will be domains $U[z]$ varying in domains Λ.

We will use the following conventions. Instead of talking about the sets $X[z]$, fixing some $z \in \Lambda$, we will say that h is the motion of X over Λ, where X is to be thought of as a set which depends on the point $z \in \Lambda$. In other words, we usually drop the brackets from the notation.

We will also use the following notation for restrictions of holomorphic motions: if $Y \subset X$, we denote $\mathcal{Y} \subset \mathcal{X}$ as the union of leaves through Y.

4.4.3. *Symmetry.* For any $n \in \mathbb{N}$, we let $\mathrm{conj} : \mathbb{C}^n \to \mathbb{C}^n$ denote the conjugacy $\mathrm{conj}(z_1, ..., z_n) = (\overline{z}_1, ..., \overline{z}_n)$.

A set $X \subset \mathbb{C}^n$ is called real-symmetric if $\mathrm{conj}(X) = X$. Let $\Lambda \subset \mathbb{C}$ be a real-symmetric domain. A holomorphic motion h over Λ is called real-symmetric if the image of any leaf by conj is also a leaf.

The systems we are interested on are real, so they naturally possess symmetry. In many cases, we will consider a real-symmetric holomorphic motion associated to the system, which will need to be completed using the Extension Lemma. The Extension Lemma adds ambiguity on the procedure, since the extension is not unique. In particular, this could lead to loss of symmetry. In order to avoid this problem, we will choose a little bit more carefully our extensions.

Symmetry assumption. *Extensions of real-symmetric motions will always be taken real-symmetric.*

Of course, to apply this Symmetry assumption we need the following:

Real Extension Lemma. *Any real-symmetric holomorphic motion can be completed to a real-symmetric holomorphic motion.*

This version of the Extension Lemma can be proved in the same way as the non-symmetric one. The reader can check for instance that in the proof of Douady [D] of the Extension Lemma there exists only one step which could lead to loss of symmetry, and thus needs to be looked more carefully in order to obtain the Real Extension Lemma: in Proposition 1 we should ask that the (not uniquely defined) diffeomorphism F is chosen real-symmetric (the proof is the same).

4.4.4. *Dilatation.* If $h = h_U$ is a holomorphic motion of an open set, we define $\mathrm{Dil}(h)$ as the supremum of the dilatations of the maps $h[z, w]$.

4.4.5. *Proper motions.* A proper motion of a set X over Λ is a holomorphic motion of X over Λ such that the map $\mathbf{h}[z] : \Lambda \times X[z] \to \mathcal{X}$ defined by $\mathbf{h}[z](w, x) = (w, h[z](w, x))$ has an extension to $\overline{\Lambda} \times X[z]$ which is a homeomorphism.

4.4.6. *Tubes and tube maps.* An equipped tube h_T is a holomorphic motion of a Jordan curve T. Its support is called a tube.

Remark 4.2. To each tube there is always an equipped tube associated (by definition). This association is unique however, since it is easy to see that if h_1 and h_2 are holomorphic motions supported on a set $\mathcal{X}$ with empty interior then $h_1 = h_2$ (consider a completion h of h_1 and conclude that $h|_X = h_2$).

We say that an equipped tube is proper if it is a proper motion. Its support is called a proper tube.

The filling of a tube is the set $\mathcal{U} \subset \Lambda \times \mathbb{C}$ such that $U[z]$ is the filling of $T[z]$, $z \in \Lambda$.

A motion of a set X is said to be bounded by the tube $\mathcal{T}$ if $\mathcal{X}$ is contained in the interior of the tube.

A motion of a set X is said to be well bounded by the proper tube $\mathcal{T}$ if it is bounded by $\mathcal{T}$ and the closure of each leaf of the motion does not intersect the closure of $\mathcal{T}$.

A special motion is a holomorphic motion $h = h_{X \cup T}$ such that $h|_T$ is an equipped proper tube and $h|_X$ is well bounded by $\mathcal{T}$.

Sometimes the following obvious criteria is useful. Let $h = h_{X_1 \cup X_2 \cup T}$ be such that $h|_{X_1 \cup T}$ is special. Suppose that for any $x \in X_2$ there is a compact set K such that $x \in K$ and $\partial K \subset X_1$ (in applications usually K is the closure of a Jordan disk). Then h is special.

If $\mathcal{T}$ is a tube over Λ, and $\mathcal{U}$ is its filling, a fiberwise map $\mathcal{F} : \mathcal{U} \to \mathbb{C}^2$ is called a tube map if it admits a continuous extension to $\overline{\mathcal{U}}$.

Let $h = h_X$ be a holomorphic motion and let $T \subset X$ be such that $h|_T$ is an equipped tube and $\mathcal{F}$ be a tube map on $\mathcal{U}$ (the filling of $\mathcal{T}$) such that $\mathcal{F}(\mathcal{T}) \subset \mathcal{X}$. We say that $(h, \mathcal{F})$ is equivariant on $\mathcal{T}$ if the image of any leaf of $h|_T$ is a leaf of h.

4.4.7. *Tube pullback.* Let us now describe a way of defining new holomorphic motions by conformal pullback of another holomorphic motion.

Let $\mathcal{F} : \mathcal{V} \to \mathbb{C}^2$ be a tube map such that $\mathcal{F}(\partial \mathcal{V}) = \partial \mathcal{U}$, where $\mathcal{U}$ is the filling of a tube over Λ and let h be a holomorphic motion supported on $\overline{\mathcal{U}} \cap \pi_1^{-1}(\Lambda)$.

Let Γ be a (parameter) open set such that $\overline{\Gamma} \subset \Lambda$ and W be a (phase) open set which moves holomorphically by h over Λ and such that $\overline{W} \subset U$. Assume that W contains critical values of $\mathcal{F}|_{\mathcal{V} \cap \pi_1^{-1}(\Gamma)}$, that is, if $\lambda \in \Gamma$, $z \in V[\lambda]$ and $DF[\lambda](z) = 0$ then $F[\lambda](z) \in W[\lambda]$.

Let us consider a leaf of h through $z \in U \setminus \overline{W}$, and let us denote by $\mathcal{E}(z)$ its preimage by $\mathcal{F}$ intersected with $\pi_1^{-1}(\Gamma)$. Each connected component of $\mathcal{E}(z)$ is a graph over Γ, moreover, $\overline{\mathcal{E}}(z) \subset \mathcal{U}$. So the set of connected components of $\mathcal{E}(z)$, $z \in U \setminus \overline{W}$ is a holomorphic motion over Γ.

We define a new holomorphic motion over Γ, called the lift of h by $(\mathcal{F}, \Gamma, W)$, as an extension to the closure of V of the holomorphic motion whose leaves are the connected components of $\mathcal{E}(z)$, $z \in U \setminus \overline{W}$ (the lift is not uniquely defined). It is clear that this holomorphic motion is a special motion of V over Γ and its dilatation over $F^{-1}(U \setminus \overline{W})$ is bounded by $K(r)$ where r is the hyperbolic diameter of Γ in Λ.

Notice that if $\lambda \in \Gamma$, $F[\lambda]^{-1}(U[\lambda] \setminus \overline{W[\lambda]})$ is a neighborhood of $\partial V[\lambda]$ in $\overline{V[\lambda]}$. In particular, if $z_n \in V$ converges to $z \in \partial V$, the leaf (of the lift of h by $(\mathcal{F}, \Gamma, W)$) through z is the limit of the leaves through z_n. By continuity, the image by $\mathcal{F}$ of the leaves through z_n converge to the image of the leaf through z. For this reason, there is a certain equivariance in the sense that the image by $\mathcal{F}$ of a leaf (of the lift of h by $(\mathcal{F}, \Gamma, W)$) is a leaf of h contained in $\partial \mathcal{U}$ (intersected with $\pi_1^{-1}(\Gamma)$), even though the map $\mathcal{F}$ is only continuous on $\partial \mathcal{V}$.

4.4.8. *Diagonal.* Let $\mathcal{T}$ be a proper tube and h the equipped proper tube supported on $\mathcal{T}$. A diagonal of $\mathcal{T}$ is a holomorphic map $\Psi : \Lambda \to \mathbb{C}^2$ with the following properties.

(1) $\Psi(\Lambda)$ is contained in the filling of $\mathcal{T}$,
(2) $\pi_1 \circ \Psi = \pi_1$,
(3) Ψ extends continuously to $\overline{\Lambda}$,
(4) $h[\lambda] \circ \Psi|_{\partial \Lambda}$ has degree one onto $T[\lambda]$.

4.4.9. *Phase-parameter holonomy maps.* Let $h = h_{X \cup T}$ be a special motion and let Φ be a diagonal of $h|_T$.

It is a consequence of the Argument Principle (see [L3]) that the leaves of $h|_X$ intersect $\Phi(\Lambda)$ in a unique point (with multiplicity one).

From this we can define a map $\chi[\lambda] : X[\lambda] \to \Lambda$ such that $\chi[\lambda](z) = w$ if (λ, z) and $\Phi(w)$ belong to the same leaf of h.

It is clear that each $\chi[\lambda]$ is a homeomorphism onto its image, moreover, if $U \subset X$ is open, $\chi[\lambda]|_{U[\lambda]}$ is locally quasiconformal, and if $\mathrm{Dil}(h|_U) < \infty$ then $\chi[\lambda]|_{U[\lambda]}$ is globally quasiconformal with dilatation bounded by $\mathrm{Dil}(h|_U)$.

We will say that χ is the holonomy family associated to the pair (h, Φ).

4.5. **Families of complex maps.**

4.5.1. *Families of R-maps.* A R-family is a pair $(\mathcal{R}, h)$, where $\mathcal{R}$ is a holomorphic map $\mathcal{R} = \mathcal{R}_{\mathcal{U}} : \cup \mathcal{U}^j \to \mathcal{U}$ and h is a holomorphic motion $h = h_{\overline{U}}$ with the following properties.

1. $P = (U, U^j)$ is a R-puzzle,
2. $h|_{\partial U \cup \cup_j \partial U_j}$ is special,
3. For every j, $\mathcal{R}|_{\mathcal{U}^j}$ is a tube map,
4. For any $\lambda \in \Lambda$, $R[\lambda]$ is a R-map in $P[\lambda]$,
5. $(\mathcal{R}, h)$ is equivariant in ∂U^j.

If additionally $\mathcal{R} \circ \mathbf{0}$ is a diagonal to h, we say that the $\mathcal{R}$ is full.

Remark 4.3. It is easy to see using Remark 4.2 that property (5), equivariance, follows from the others.

4.5.2. *From R-families to L-families.* Given a R-family $\mathcal{R}$ with motion $h = h_{\overline{U}}$ we induce a family of L-maps as follows.

We first define tubes $\mathcal{U}^{\underline{d}}$ inductively on $|\underline{d}|$: if $\underline{d} \in \Omega$ with $\underline{d} = (j_1, ..., j_m)$ we take $\mathcal{U}^{\underline{d}} = \mathcal{U}$ if $|\underline{d}| = 0$, and we let $\mathcal{U}^{\underline{d}} = (\mathcal{R}|_{\mathcal{U}^{j_1}})^{-1}(\mathcal{U}^{\sigma^-(\underline{d})})$ otherwise.

We define $\mathcal{R}^{\underline{d}} = \mathcal{R}^{|\underline{d}|}|_{\mathcal{U}^{\underline{d}}}$ and construct tubes $\mathcal{W}^{\underline{d}} = (\mathcal{R}^{\underline{d}})^{-1}(\mathcal{U}^0)$.

We define $L(\mathcal{R}) : \cup \mathcal{W}^{\underline{d}} \to \mathcal{U}^0$ by $L(\mathcal{R})|_{\mathcal{W}^{\underline{d}}} = \mathcal{R}^{\underline{d}}$. Notice that the L-maps which are associated with the fibers of $\mathcal{R}$ coincide with the fibers of $L(\mathcal{R})$.

We define a holomorphic motion $L(h)$ in the following way. The leaf through $z \in \partial U$ is the leaf of h through z. If there is a smallest $U^{\underline{d}}$ such that $z \in U^{\underline{d}}$, we let the leaf through z be the preimage by $\mathcal{R}^{\underline{d}}$ of the leaf through $\mathcal{R}^{\underline{d}}(z)$. We finally extend it to $\overline{U}$ using the Extension Lemma.

The L-family associated to $(\mathcal{R}, h)$ is a pair $(L(\mathcal{R}), L(h))$. It has some properties similar to R-maps.

1. $L(h)|_{\partial U \cup \cup_j \partial U^j} = h$ and so $L(h)|_{\partial U \cup \cup_j \overline{U^j}}$ is special,
2. $L(\mathcal{R})|_{\mathcal{W}^{\underline{d}}}$ is a tube map,
3. $(L(\mathcal{R}), L(h))$ is equivariant in $\partial W^{\underline{d}}$.

4.5.3. *Parameter partition.* Let $(\mathcal{R}, h)$ be a full R-family. Since $L(h)|_{U \cup \cup_j \overline{U^j}}$ is special, we can consider the holonomy family of the pair $(L(h)|_{U \cup \cup_j \overline{U^j}}, \mathcal{R}(\mathbf{0}))$, which we denote by χ.

We can now use χ to partition Λ: let $\Lambda^{\underline{d}} = \chi(U^{\underline{d}})$ and let $\Gamma^{\underline{d}} = \chi(W^{\underline{d}})$.

4.5.4. *Family renormalization.* Let $(\mathcal{R}, h)$ be a full R-family. The $\underline{d}$ renormalization of $(\mathcal{R}, h)$ is the R-family $(N^{\underline{d}}(\mathcal{R}), N^{\underline{d}}(h))$ over $\Gamma^{\underline{d}}$ defined as follows. We take $N^{\underline{d}}(h)$ as the lift of $L(h)$ by $(\mathcal{R}|_{\mathcal{U}_0}, \Gamma^{\underline{d}}, W^{\underline{d}})$ where defined.

It is clear that $(N^{\underline{d}}(\mathcal{R}), N^{\underline{d}}(h))$ is full, and its fibers are renormalizations of the fibers of $(\mathcal{R}, h)$. Moreover, $N^{\underline{d}}(h)$ is a special motion.

4.5.5. *Truncation and gape renormalization.* Let $(\mathcal{R}, h)$ be a full R-family and let $\underline{d} \in \Omega$. We define the $\underline{d}$ truncation of $L(\mathcal{R})$ as $L^{\underline{d}}(\mathcal{R}) = L(\mathcal{R})$ outside $\mathcal{U}^{\underline{d}}$ and $L^{\underline{d}}(\mathcal{R}) = \mathcal{R}^{\underline{d}}$ in $\mathcal{U}^{\underline{d}}$. Let $G^{\underline{d}}(\mathcal{R}) = L^{\underline{d}}(\mathcal{R}) \circ \mathcal{R}|_{\mathcal{U}_0 \cap \pi_1^{-1}(\Lambda^{\underline{d}})}$ where defined.

If $\underline{d}$ is empty, let $G^{\underline{d}}(h) = L(h)$. Otherwise, let $G^{\underline{d}}(h)$ be a holomorphic motion of U over $\Lambda^{\underline{d}}$, which coincides with $L(h)$ on $U \setminus \overline{U^0}$ and coincides with the lift of $L(h)$ by $(\mathcal{R}|_{\mathcal{U}_0}, \Lambda^{\underline{d}}, U^{\underline{d}})$ on U_0.

The $\underline{d}$ gape renormalization of $(\mathcal{R}, h)$ is the pair $(G^{\underline{d}}(\mathcal{R}), G^{\underline{d}}(h))$. It is clear that the fibers of $G^{\underline{d}}(\mathcal{R})$ are $\underline{d}$ gape renormalizations of the fibers of $\mathcal{R}$.

Observe that if h is a special motion, then $G^{\underline{d}}(h)$ is also a special motion and that $G^{\underline{d}}(\mathcal{R}) \circ \mathbf{0}$ is a diagonal to it.

Notice that $G^{\underline{d}}(\mathcal{R})(\lambda, 0) = N^{\underline{d}}(\mathcal{R})(\lambda, 0)$ for $\lambda \in \Gamma^{\underline{d}}$. Moreover, the holomorphic motion $G^{\underline{d}}(h)$ is an extension (both in phase as in parameter) of $N^{\underline{d}}(h)|_A$, where $A = \overline{U^0} \setminus (R|_{U^0})^{-1}(U^{\underline{d}})$.

4.6. **Chains.** Assume now that we are given a full R-family, which we will denote $\mathcal{R}_1$ (over some domain Λ_1, with motion h_1) together with a parameter $\lambda \in \mathbb{R}$. If λ belongs to some renormalization domain (that is, there exists $\underline{d}_1$ such that $\lambda \in \Lambda_1^{\underline{d}_1}$), let $\mathcal{R}_2 = N_1^{\underline{d}}(\mathcal{R}_1)$ (over Λ_2). Assume we can continue this process constructing $\mathcal{R}_{i+1} = N_i^{\underline{d}_i}$, $n \geq 1$. Then the sequence $\mathcal{R}_i$ (over Λ_i) will be called the $\mathcal{R}$-chain over λ.

4.6.1. *Notation.* The holomorphic motion associated to $\mathcal{R}_i$ is denoted h_i (so that $h_{i+1} = N^{\underline{d}_i}(h_i)$),

To simplify the notation for the gape renormalization, we let $G^{\underline{d}_i}(\mathcal{R}_i) = G(\mathcal{R}_i)$ and $G^{\underline{d}_i}(h_i) = G(h_i)$.

Notice that the sequence $\underline{d}_i$ defined as above satisfies $\Lambda_{i+1} = \Gamma_i^{\underline{d}_i}$. We denote $\tilde{\Lambda}_{i+1} = \Lambda_i^{\underline{d}_i}$.

We denote $\tilde{U}_{i+2} = (R_i|_{U_{i+1}})^{-1}(U_i^{\underline{d}_i})$ over $\tilde{\Lambda}_{i+1}$. Notice that $\tilde{U}_{i+2}$ moves holomorphically by $G(h_i)$.

Notice that for any j, either $\overline{U_{i+1}^j} \subset \tilde{U}_{i+2}$ or $\overline{U_{i+1}^j} \cap \overline{\tilde{U}_{i+2}} = \emptyset$. The first case happens in particular for $j = 0$. Notice also that $h_{i+1}|_{\overline{U_{i+1}} \setminus \tilde{U}_{i+2}}$ agrees with $G(h_i)$ over Λ_{i+1}.

4.6.2. *Holonomy maps.* Notice that for $i > 1$, the holomorphic motion h_i is special (since it is obtained by renormalization and so coincides with $N^{\underline{d}_{i-1}}(h_{i-1})$). In particular, we can consider the holonomy family associated to $(h_i, \mathcal{R}_i \circ \mathbf{0})$, which we denote by $\chi_i^0 : U_i \to \Lambda_i$.

For $i > 1$, $L(h_i)$ is also special, let $\chi_i : U_i \to \Lambda_i$ be the holonomy family associated to $(L(h_i), \mathcal{R}_i \circ \mathbf{0})$.

For $i > 2$, $G(h_{i-1})$ is also special, let $\tilde{\chi}_i : U_{i-1} \to \tilde{\Lambda}_i$ be the holonomy family of the pair $(G(h_{i-1}), G(\mathcal{R}_{i-1}) \circ \mathbf{0})$.

Notice that for $\lambda \in \Lambda_i$, $\chi_i[\lambda]|_{\overline{U_i} \setminus \cup U_i^j}$ coincides with $\chi_i^0[\lambda]$.

Notice that for $\lambda \in \Lambda_i$, then $\tilde{\chi}_i[\lambda]|_{\overline{U_i} \setminus \tilde{U}_{i+1}}$ coincides with $\chi_i^0[\lambda]$.

4.6.3. *Real chains.* A fiberwise map $\mathcal{F} : \mathcal{X} \to \mathbb{C}^2$ is real-symmetric if $\mathcal{X}$ is real-symmetric and $\mathcal{F} \circ \mathrm{conj} = \mathrm{conj} \circ \mathcal{F}$.

We will say that a chain $\{\mathcal{R}_i\}$ over a parameter $\lambda \in \mathbb{R}$ is real-symmetric if each $\mathcal{R}_i$ and each underlying holomorphic motion h_i is real-symmetric.

Because of the Symmetry assumption, a chain $\{\mathcal{R}_i\}$ over a parameter $\lambda \in \mathbb{R}$ is real-symmetric provided the first step data $\mathcal{R}_1$ and h_1 is real-symmetric.

In this case, all objects related to the chain are real-symmetric, including the holonomy families χ_i^0, χ_i and $\tilde{\chi}_i$ (where we say that a holonomy family $\chi[\lambda] : U[\lambda] \to \Lambda$ is real-symmetric if $\overline{\chi[\lambda](z)} = \chi[\overline{\lambda}](\overline{z})$, in particular, for real λ, $\chi[\lambda]$ is a real-symmetric function).

Remark 4.4. If $\mathcal{X}$ has empty interior and supports a holomorphic motion, then this holomorphic motion is unique by Remark 4.2. In particular, if $\mathcal{X}$ is real-symmetric, this holomorphic motion is real-symmetric.

4.7. **Complex extension of the real quadratic family and parapuzzle geometry.** The complex families we introduced are designed to model families of the complex return maps which appear in Theorems 4.2 and 4.3. In [L3], Lyubich proves (a stronger version of) the following result:

Theorem 4.4. *Let $\lambda_0 \in \mathbb{R}$ be a parameter corresponding to a quadratic map $f = f_{\lambda_0} \in \mathcal{F}_\kappa$. Then there exists a real-symmetric R-family $\mathcal{R}_1$ over a domain Λ_1 such that $\Lambda_1 \cap \mathbb{R} = J_1$ and such that if $\lambda \in J_1$ then the real trace of $R_1[\lambda]$ is the first return map to $I_1[f_\lambda]$.*

Remark 4.5. The family $\mathcal{R}_1$ above is constructed in §3.4 of [L3], using combinatorial properties of the Mandelbrot set which we did not discuss.

Remark 4.6. In particular, if $\lambda_0 \in \mathbb{R}$ and $f_{\lambda_0} \in \mathcal{F}_\kappa$, then we can consider the chain $\mathcal{R}_i$ over λ_0 with first step $\mathcal{R}_1$ as in Theorem 4.4.

In the setting of the Topological Phase-Parameter relation, we have $\Lambda_i \cap \mathbb{R} = J_i$ and so we have $\Xi_i|_{K_i} = \chi_i[\lambda_0]$, and for $\lambda \in J_i$, $H_i[f_\lambda]|_{K_i} = L(h_i)[\lambda_0, \lambda]$. In particular, this gives a (rather sophisticated) proof of the Topological Phase-Parameter relation, since χ_i and $L(h_i)$ are homeomorphisms.

Notice also that $\tilde{U}_{i+1} \cap \mathbb{R} = \tilde{I}_{i+1}$ and so $\Xi_i|_{\tilde{K}_i} = \tilde{\chi}_i[\lambda_0]$. We also have for $\lambda \in J_i$, $H_i[f_\lambda]|_{\tilde{K}_i} = G(h_{i-1})[\lambda_0, \lambda]$.

We are now ready to state the following result of Lyubich on the geometry of the domains Λ_i. This estimate is (a special case of) Theorem A in [L3].

Theorem 4.5. *Let $\lambda_0 \in \mathbb{R}$ be a parameter corresponding to a simple quadratic map $f = f_{\lambda_0} \in \mathcal{F}_\kappa$. Let $\mathcal{R}_1$ be as in Theorem 4.4, and let $\mathcal{R}_i$ be the real chain over λ_0 (with first step $\mathcal{R}_1$). Then* $\mathrm{mod}(\Lambda_i \setminus \overline{\Lambda_{i+1}})$ *grows at least linearly.*

Moreover, we still have (from Theorem 4.3) that $\mathrm{mod}(U_i[\lambda_0] \setminus \overline{U_i^0[\lambda_0]})$ *grows linearly.*

Remark 4.7. The statement of [L3] looks initially different from Theorem 4.5. Indeed, Lyubich considers only a subsequence of our sequence of parapuzzle pieces Λ_i, namely his subsequence consists of the Λ_{n_i} (which he denotes Δ^i) where $\underline{d}_{n_i-1}$ is not empty. In his notation, if $\underline{d}_{n_i}$ is empty, then our Λ_{n_i+1} is denoted Π^i. His statement is that both $\Delta^i \setminus \overline{\Delta^{i+1}}$ and $\Delta^i \setminus \overline{\Pi^i}$ have linearly big modulus. Since we consider only simple maps, after a finite number of steps, both sequences Λ_i and Δ^i differ only by a shift, so his statement implies ours.

Remark 4.8. Our construction of renormalization is directly based on §3.5, 3.6 and 3.7 of [L3]. Apart from differences of notation ([L3] does not use the name chain, for instance), the main difference is that we do not jump over central levels, as we mentioned in the previous remark, and that we introduce the concept of gape renormalization.

Remark 4.9. In Theorem A of [L3], a much more general result is proved.

(1) Non-real quadratic maps are also considered, but to state precise results it is needed to introduce quite a bit of the combinatorial theory of the Mandelbrot set.
(2) A geometric estimate is valid for more general maps then simple maps: if $n_k - 1$ counts the non-central levels of the principal nest, then $\mathrm{mod}(\Lambda_{n_k} \setminus \overline{\Lambda_{n_k+1}})$ grows at least linearly fast (see Remark 4.7). This version is used by Lyubich to show that non-simple parameters in $\mathcal{F}_\kappa$ have zero Lebesgue measure.
(3) The rate of linear growth is independent of the real quadratic map considered (a consequence of complex bounds).
(4) For complex maps, the rate of linear growth depends on finite combinatorial information.

5. The phase-parameter relation

Let us fix a simple map f_{λ_0}, and let $\mathcal{R}_i$ be the R-chain given by Theorem 4.5.

5.1. **Preliminary estimates.**

The following is a well known estimate of hyperbolic geometry.

Proposition 5.1. *Let V, W be Jordan disks such that $\overline{W} \subset V$. If $m = \mathrm{mod}(V \setminus \overline{W})$ is big, then the hyperbolic diameter of W in V is exponentially small in m.*

It immediately implies:

Proposition 5.2. *Let V, V_1 and V_2 be Jordan disks such that $\overline{V}_1, \overline{V}_2 \subset V$. If $m = \min\{\mathrm{mod}(V \setminus \overline{V_1}), \mathrm{mod}(V \setminus \overline{V_2})\}$ is big then there exists an annulus $A \subset V \setminus \overline{V_1 \cup V_2}$, non-homotopic to a constant on $V \setminus \overline{V_1}$ and whose modulus is linearly big in m.*

Proof. We may assume that $V = \mathbb{D}$ and $0 \in V_1$. By the previous proposition, the hyperbolic diameters of V_1 and V_2 in $\mathbb{D}$ are exponentially small in m. Since $0 \in V_1$, the Euclidean diameter r of V_1 is also exponentially small in m. Let $D = \mathbb{D}_{2r^{1/2}}$, $D' = \mathbb{D}_{r^{1/2}}$ be disks around 0 of Euclidean radius $2r^{1/2}$, $r^{1/2}$ respectively. Then both annulus $A = \mathbb{D} \setminus \overline{D}$ and $A' = D' \setminus \overline{V_1}$ surround V_1 and have modulus of order $-\ln(r)$, and so linearly big in m. Notice that the hyperbolic distance between ∂D and $\partial D'$ is of order $r^{1/2}$, so is bigger than the hyperbolic diameter of V_2. In particular, one of the annulus A and A' does not intersect V_2. $\square$

We will also need the following easy fact:

Proposition 5.3. *Let U, U' be Jordan disks and $f : U \to U'$ be a double covering of U' ramified at 0. Let $V' \subset U'$ be a Jordan disk, and let V be a component of $f^{-1}(V')$. Then $\mathrm{mod}(U \setminus \overline{V}) = \mathrm{mod}(U' \setminus \overline{V'})/2$ if $0 \in V$ and $\mathrm{mod}(U \setminus \overline{V}) \geq \mathrm{mod}(U' \setminus \overline{V'})/3$ if $0 \notin V$.*

5.1.1. *Finding space.*

Lemma 5.4. *For all $\epsilon > 0$ there exists i_0 such that if $i > i_0$ the dilatation of $\chi_i^0|_{U_i \setminus \overline{U_i^0}}$ is less then $1 + \epsilon$.*

Proof. Notice that the dilatation of $\chi_i^0|_{U_i \setminus \overline{U_i^0}}$ can be bounded by $\mathrm{Dil}(h_i|_{U_i \setminus \overline{U_i^0}})$. On the other hand, h_i is obtained by lift of $L(h_{i-1})$ by $(\mathcal{R}_{i-1}|_{\mathcal{U}_i}, \Gamma_{i-1}^{\underline{d}_{i-1}}, W_{i-1}^{\underline{d}_{i-1}})$, and $(\mathcal{R}_{i-1}|_{\mathcal{U}_i})^{-1}(U_{i-1} \setminus \overline{W_{i-1}^{\underline{d}_{i-1}}}) = U_i \setminus \overline{U_i^0}$, so $\mathrm{Dil}(h_i|_{U_i \setminus \overline{U_i^0}}) \le K(r)$ where r is the hyperbolic diameter of Λ_i in Λ_{i-1}. The conclusion follows from Theorem 4.5 which gives growth of $\mathrm{mod}(\Lambda_{i-1} \setminus \overline{\Lambda_i})$, together with Proposition 5.1. $\square$

Remark 5.1. Lemma 5.4 implies that $\inf_{\lambda \in \Lambda_i} \mathrm{mod}(U_i[\lambda] \setminus \overline{U_i^0[\lambda]})$ is asymptotically the same as $\mathrm{mod}(U_i[\lambda_0] \setminus \overline{U_i^0[\lambda_0]})$ which is linearly big.

In particular, since $\mathrm{mod}(U_i^{\underline{d}} \setminus \overline{W_i^{\underline{d}}}) = \mathrm{mod}(U_i \setminus \overline{U_i^0})$, $\inf_{\lambda \in \Lambda_i} \mathrm{mod}(U_i^{\underline{d}}[\lambda] \setminus \overline{W_i^{\underline{d}}[\lambda]})$ is also linearly big (independently of $\underline{d}$).

Notice that if $\underline{d}$ is such that $R_{i-1}(U_i^j) = W_{i-1}^{\underline{d}}$ then by Proposition 5.3, $\mathrm{mod}(U_i \setminus \overline{U_i^j}) \ge \mathrm{mod}(U_{i-1} \setminus \overline{W_{i-1}^{\underline{d}}})/3$, so we conclude that $\inf_{\lambda \in \Lambda_i} \mathrm{mod}(U_i[\lambda] \setminus \overline{U_i^j[\lambda]})$ is linearly big (independently of j).

Lemma 5.5. *In this setting,* $\inf_j \mathrm{mod}(\Lambda_i \setminus \overline{\Lambda_i^j})$ *grows at least linearly fast on i.*

Proof. Fix some Λ_i^j. By Remark 5.1, both $\mathrm{mod}(U_i \setminus \overline{U_i^j})$ and $\mathrm{mod}(U_i \setminus \overline{U_i^0})$ are linearly big, so by Proposition 5.2 there exists an annulus A such that $A \subset U_i \setminus \overline{U_i^j \cup U_i^0}$ is not homotopic to a constant on $U_i \setminus \overline{U_i^j}$ and such that $A[\lambda_0]$ has linearly big moduli.

By Lemma 5.4, $\mathrm{mod}(\chi_i^0(A))$ is linearly big, and since $\mathrm{mod}(\Lambda_i \setminus \overline{\Lambda_i^j}) \ge \mathrm{mod}(\chi_i^0(A))$ the result follows. $\square$

Lemma 5.6. *In this setting,* $\inf_{\underline{d}} \mathrm{mod}(\Lambda_i^{\underline{d}} \setminus \overline{\Gamma_i^{\underline{d}}})$ *grows at least linearly fast on i.*

Proof. Fix some $\Lambda_i^{\underline{d}}$. If $\underline{d}$ is empty, the result is contained in Lemma 5.5. Otherwise there exists j with $U_i^{\underline{d}} \subset U_i^j$. By Remark 5.1, $\mathrm{mod}(U_i^{\underline{d}}[\lambda] \setminus \overline{W_i^{\underline{d}}[\lambda]})$ is linearly big independently of $\lambda \in \Lambda_i$.

By Lemma 5.5, the hyperbolic diameter of Λ_i^j in Λ_i is small, so for $\lambda \in \Lambda_i^{\underline{d}}$ the dilatation of $\chi_i[\lambda]|_{U_i^{\underline{d}}}$ is close to 1. This implies that $\mathrm{mod}(\Lambda_i^{\underline{d}} \setminus \overline{\Gamma_i^{\underline{d}}}) = \mathrm{mod}(\chi_i(U_i^{\underline{d}} \setminus \overline{W_i^{\underline{d}}}))$ is linearly big. $\square$

Lemma 5.7. *Let $h_{\overline{U}}$ be a special motion over Λ, Φ be a diagonal, and χ be the holonomy family. Let $W \subset U$ be a Jordan disk such that* $\mathrm{mod}(\Lambda \setminus \chi(\overline{W})) = m$. *Then there exists a Jordan disk $V \subset U$ such that* $\mathrm{mod}(\chi(V) \setminus \chi(W)) = m/2$ *and such that $\chi[\lambda]|_{V[\lambda]}$, $\lambda \in \chi(V)$, has dilatation bounded by $1 + \epsilon(m)$, where $\epsilon(m) \to 0$ exponentially fast as $m \to \infty$.*

Proof. Let Υ be a Jordan disk such that $\mathrm{mod}(\Lambda \setminus \overline{\Upsilon}) = \mathrm{mod}(\Upsilon \setminus \chi(\overline{W})) = m/2$. Consider $V = \chi^{-1}(\Upsilon)$. By the λ-Lemma, $\mathrm{Dil}(\chi[\lambda]|_{V[\lambda]}) \le K(r) = 1 + O(r)$, $\lambda \in \Upsilon$, where r is the hyperbolic diameter of Υ on Λ. By Proposition 5.1, the hyperbolic diameter of Υ is exponentially small on $m/2$. $\square$

5.2. **The phase-parameter estimates.**

Lemma 5.8. *For all $C, \epsilon > 0$ there exists i_0 such that for all $i > i_0$, $j \in \mathbb{Z}$, and for all $\lambda \in \Lambda_i^j \cap \mathbb{R}$, $\chi_i[\lambda]|_{U_i^j[\lambda]}$ has a (C, ϵ)-qc extension which is real-symmetric.*

Proof. The extension claimed is just a convenient restriction of $\chi_i[\lambda]|_{U_i}$. By Lemma 5.5, if i is sufficiently big, the moduli of $\Lambda_i \setminus \overline{\Lambda_i^j}$ is linearly big. By Lemma 5.7 it follows that if i is sufficiently big we can find a Jordan disk $V_i \subset U_i$ such that $\operatorname{mod}(V_i[\lambda] \setminus \overline{U_i^j[\lambda]})$ is bigger than C and $\chi_i[\lambda]|_{V_i[\lambda]}$ has dilatation less than $1+\epsilon$. $\square$

Lemma 5.9. *For all $C, \epsilon > 0$ there exists i_0 such that for all $i > i_0$ and for all $\lambda \in \Lambda_i \cap \mathbb{R}$, $\tilde{\chi}_i[\lambda]|_{U_i[\lambda]}$ has a (C,ϵ)-qc extension which is real-symmetric.*

Proof. The idea is the same as before: the extension claimed is just a convenient restriction of $\tilde{\chi}_i[\lambda]|_{U_{i-1}[\lambda]}$. By Lemma 5.6, the modulus of $\tilde{\chi}_i(U_{i-1} \setminus \overline{U_i}) = \Lambda_{i-1}^{\underline{d}_{i-1}} \setminus \Gamma_{i-1}^{\underline{d}_{i-1}}$ is linearly big. By Lemma 5.7, for i sufficiently big there exists a Jordan disk $V_i \subset U_{i-1}$ such that $\operatorname{mod}(U_{i-1}[\lambda] \setminus \overline{V_i[\lambda]})$ is bigger than C and $\tilde{\chi}_i[\lambda]|_{V_i[\lambda]}$ has dilatation less than $1+\epsilon$. $\square$

Remark 5.2. Estimates for the holonomy map between phase spaces also follow. Using the estimates on moduli from Lemmas 5.5 and 5.6, the Real Extension Lemma and the λ-Lemma, we can easily deduce that for all $\epsilon > 0$ there exists i_0 such that if $i > i_0$ we have:

(1) For $j \in \mathbb{Z}$, for $\lambda_1, \lambda_2 \in \Lambda_i^j \cap \mathbb{R}$, the phase holonomy map $L(h_i)[\lambda_1, \lambda_2]$ extends to the holonomy map between λ_1 and λ_2 of the completion of $L(h_i)$, which is a quasiconformal real-symmetric map of the whole plane. Its dilatation is smaller then $1+\epsilon$ due to Lemma 5.5 and the λ-Lemma.
(2) For $\lambda_1, \lambda_2 \in \Lambda_i \cap \mathbb{R}$, the phase holonomy map $G(h_{i-1})[\lambda_1, \lambda_2]$ extends to the holonomy map between λ_1 and λ_2 of the completion of $G(h_{i-1})$, which is a quasiconformal real-symmetric map of the whole plane. Its dilatation is smaller then $1+\epsilon$ due to Lemma 5.6 and the λ-Lemma.

5.2.1. *Proof of the Phase-Parameter relation.* Let $f = f_{\lambda_0} \in \mathcal{F}_\kappa$ as above. By Lemma 4.1, it is enough to show that for i sufficiently big the required phase-parameter and phase-phase maps have big (real-symmetric) extensions with small dilatation.

According to the discussion of Remark 4.6, $\Xi_i|_{K_i} = \chi_i[\lambda_0]$, so Lemma 5.8 imply PhPa1. In the same way, $\Xi_i|_{\tilde{K}_i} = \tilde{\chi}_i[\lambda_0]$, so Lemma 5.9 imply PhPa2.

Still according to Remark 4.6, if $f_\lambda \in J_i$, $H_i[f_\lambda]|_{K_i}$ coincides with the holonomy map $L(h_i)[\lambda_0, \lambda]$, so the first item of Remark 5.2 implies PhPh1. Moreover, $H_i[f_\lambda]|_{\tilde{K}_i}$ coincides with the holonomy map $G(h_{i-1})[\lambda_0, \lambda]$, so the second item of Remark 5.2 implies PhPh2.

6. Measure and capacities

6.1. **Quasisymmetric maps.** If $X \subset \mathbb{R}$ is measurable, let us denote $|X|$ its Lebesgue measure. Let us explicit the metric properties of γ-qs maps we will use.

To each γ, there exists a constant $k \geq 1$ such that for all $f \in QS(\gamma)$, for all $J \subset I$ intervals,

$$\frac{1}{k}\left(\frac{|J|}{|I|}\right)^k \leq \frac{|f(J)|}{|f(I)|} \leq \left(\frac{k|J|}{|I|}\right)^{1/k}.$$

Furthermore $\lim_{\gamma \to 1} k(\gamma) = 1$. So for each $\epsilon > 0$ there exists $\gamma > 1$ such that $k(2\gamma - 1) < 1 + \epsilon/5$. From now on, once a given γ close to 1 is chosen, ϵ will always denote a small number with this property.

6.2. **Capacities and trees.** The γ-capacity of a set X in an interval I is defined as follows:

$$p_\gamma(X|I) = \sup_{h \in QS(\gamma)} \frac{|h(X \cap I)|}{|h(I)|}.$$

This geometric quantity is well adapted to our context, since it is well behaved under tree decompositions of sets. In other words, if I^j are disjoint subintervals of I and $X \subset \cup I^j$ then

$$p_\gamma(X|I) \leq p_\gamma(\cup_j I_j|I) \sup_j p_\gamma(X|I^j).$$

6.3. **A measure theoretical lemma.** Our usual procedure consists in picking a class X of maps which we show to have full measure among non-regular maps and then for each map in this class we proceed to describe what happens in the principal nest, and trying to use this information to show that a subset Y of X still has full measure. An example of this parameter exclusion process is done by Lyubich in [L3] where he shows using a probabilistic argument that the class of simple maps has full measure in $\mathcal{F}$.

Let us now describe our usual argument (based on the argument of Lyubich which in turn is a variation of the Borel-Cantelli Lemma). Assume at some point we know how to prove that almost every simple map belongs to a certain set X. Let Q_n be a (bad) property that a map may have (usually some anomalous statistical parameter related to the n-th stage of the principle nest). Suppose we prove that if $f \in X$ then the probability that a map in $J_n(f)$ has the property Q_n is bounded by $q_n(f)$ which is shown to be summable for all $f \in X$. We then conclude that almost every map does not have property Q_n for n big enough.

Sometimes we also apply the same argument, proving instead that $q_n(f)$ is summable where $q_n(f)$ is the probability that a map in $J_n^{\tau_n}(f)$ has property Q_n, (recall that τ_n is such that $f \in J_n^{\tau_n}(f)$).

In other words, we apply the following general result.

Lemma 6.1. *Let $X \subset \mathbb{R}$ be a measurable set such that for each $x \in X$ is defined a sequence $D_n(x)$ of nested intervals converging to x such that for all $x_1, x_2 \in X$ and any n, $D_n(x_1)$ is either equal or disjoint to $D_n(x_2)$. Let Q_n be measurable subsets of $\mathbb{R}$ and $q_n(x) = |Q_n \cap D_n(x)|/|D_n(x)|$. Let Y be the set of x in X which belong to finitely many Q_n. If $\sum q_n(x)$ is finite for almost any $x \in X$ then $|Y| = |X|$.*

Proof. Let $Y_n = \{x \in X | \sum_{k=n}^{\infty} q_k(x) < 1/2\}$. It is clear that $Y_n \subset Y_{n+1}$ and $|\cup Y_n| = |X|$.

Let $Z_n = \{x \in Y_n | |Y_n \cap D_m(x)|/|D_m(x)| > 1/2, m \geq n\}$. It is clear that $Z_n \subset Z_{n+1}$ and $|\cup Z_n| = |X|$.

Let $T_n^m = \cup_{x \in Z_n} D_m(x)$. Let $K_n^m = T_n^m \cap Q_m$. Of course

$$|K_n^m| \leq \int_{T_n^m} q_m \leq 2 \int_{Y_n} q_m.$$

And of course

$$\sum_{m \geq n} \int_{Y_n} q_m \leq \frac{1}{2} |Y_n|.$$

This shows that $\sum_{m \geq n} |K_n^m| \leq |Y_n|$, so almost every point in Z_n belong to finitely many K_n^m. We conclude then that almost every point in X belong to finitely many Q_m. $\square$

The following obvious reformulation will be often convenient

Lemma 6.2. *In the same context as above, assume that we are given sequences $Q_{n,m}$, $m \geq n$ of measurable sets and let Y_n be the set of x belonging to at most finitely many $Q_{n,m}$. Let $q_{n,m}(x) = |Q_{n,m}(x) \cap D_m(x)|/|D_m(x)|$. Let $n_0(x)$ be such that $\sum_{m=n}^{\infty} q_{n,m}(x) < \infty$ for $n \geq n_0(x)$. Then for almost every $x \in X$, $x \in Y_n$ for $n \geq n_0(x)$.*

In practice, we will estimate the capacity of sets in the phase space: that is, given a map f we will obtain subsets $\tilde{Q}_n[f]$ in the phase space, corresponding to bad branches of return or landing maps. We will then show that for some $\gamma > 1$ we have $\sum p_\gamma(\tilde{Q}_n[f]|I_n[f]) < \infty$ or $\sum p_\gamma(\tilde{Q}_n[f]|I_n^{\tau_n}[f]) < \infty$. We will then use PhPa1, PhPa2 and the measure-theoretical Lemma to conclude that with total probability among non-regular maps, for all n sufficiently big, $R_n(0)$ does not belong to a bad set.

From now on when we prove that almost every non-regular map has some property, we will just say that with total probability (without specifying) such property holds.

(To be strictly formal, we have fixed the renormalization level κ (in particular to define the sequence J_n without ambiguity), so applications of the measure theoretical argument will actually be used to conclude that for almost every parameter in $\mathcal{F}_\kappa$ a given property holds. Since almost every non-regular map belongs to some $\mathcal{F}_k$, this is equivalent to the statement regarding almost every non-regular parameter.)

7. Basic ideas of the statistical analysis

Before proceeding for the statistical analysis, let us explain the basic use that we will make of the phase-parameter relation. We will illustrate our ideas with an informal presentation of one of our first results in the next section.

For a map $f \in \Delta_\kappa$ (recall that, as always, we work in a fixed level κ of renormalization), let us associate a sequence of "statistical parameters" in some way. A good example of statistical parameter is s_n, which denotes the number of times the critical point 0 returns to I_n before the first return I_{n+1}. Each of the points of the sequence $R_n(0)$,...,$R_n^{s_n}(0)$ can be located anywhere inside I_n. Pretending that the distribution of those points is indeed uniform with respect to Lebesgue measure, we may expect that s_n is near c_n^{-1}, where $c_n = |I_{n+1}|/|I_n|$.

Let us try to make this rigorous. Consider the set of points $A_k \subset I_n$ which iterate exactly k times in I_n before entering I_{n+1}. Then most points $x \in I_n$ belong to some A_k with k in a neighborhood of c_n^{-1}. By most, we mean that the complementary event has small probability, say q_n, for some summable sequence q_n. This neighborhood has to be computed precisely using a statistical argument, in this case if we choose the neighborhood $c_n^{-1+2\epsilon} < k < c_n^{-1-\epsilon}$), we obtain the sequence $q_n < c_n^\epsilon$ which is indeed summable for all simple maps f by [L1].

If the phase-parameter relation was Lipschitz, we would now argue as follows: the probability of a parameter be such that $R_n(0) \in A_k$ with k out of the "good neighborhood" of values of k is also summable (since we only multiply those probabilities by the Lipschitz constant) and so, by the measure-theoretical argument of Lemma 6.1, for almost every parameter this only happens a finite number of times.

Unfortunately, the phase-parameter relation is not Lipschitz. To make the above argument work, we must have better control of the size of the "bad set" of points which we want the critical value $R_n(0)$ to not fall into. In order to do so, in the statistical analysis of the sets A_k, we control instead the quasisymmetric capacity of the complement of points falling in the good neighborhood. This makes the analysis sometimes much more difficult: capacities are not probabilities (since they are not additive): in fact we can have two disjoint sets with capacity close to 1. This will usually introduce some error that was not present in the naive analysis: this is the ϵ in the range of exponents present above. If we were not forced to deal with capacities, we could get much finer estimates.

Incidentally, to keep the error low, making ϵ close to 0, we need to use capacities with constant γ close to 1. Fortunately, our phase-parameter relation has a constant converging to 1, which will allow us to partially get rid of this error.

Coming back to our problem, we see that we should concentrate in proving that for almost every parameter, certain bad sets have summable γ-qs capacities for some constant γ independent of n (but which can depend on f).

There is one final detail to make this idea work in this case: there are two phase-parameter statements, and we should use the right one. More precisely, there will be situations where we are analyzing some sets which are union of I_n^j (return sets), and sometimes union of $C_n^{\underline{d}}$ (landing sets). In the first case, we should use the PhPa2 and in the second the PhPa1. Notice that our phase-parameter relations only allow us to "move the critical point" inside I_n with respect to the partition by I_n^j, to do the same with respect to the partition by $C_n^{\underline{d}}$, we must restrict ourselves to $I_n^{\tau_n}$. In all cases, however, the bad sets considered should be either union of I_n^j or $C_n^{\underline{d}}$.

For our specific example, the A_k are union of $C_n^{\underline{d}}$, we must use PhPa1. In particular we have to study the capacity of a bad set inside $I_n^{\tau_n}$. Here is the estimate that we should go after (see Lemma 8.2 for a more precise statement):

Lemma 7.1. *For almost every parameter, for every $\epsilon > 0$, there exists γ such that $p_\gamma(X_n|I_n^{\tau_n})$ is summable, where X_n is the set of points $x \in I_n$ which enter I_{n+1} either before $c_n^{-1+\epsilon}$ or after $c_n^{-1-\epsilon}$ returns to I_n.*

We are now in position to use PhPa1 to make the corresponding parameter estimate: using the measure-theoretic argument, we get that with total probability

$$\lim_{n\to\infty} \frac{\ln s_n}{\ln c_n^{-1}} = 1.$$

This is the content of Lemma 8.3.

This particular estimate we choose to describe in this section is extremely important for the analysis to follow: we can use s_n to estimate c_{n+1} directly. This last lemma implies (see Lemma 8.4) that c_n decays torrentially to 0 for typical parameters. For general simple maps, the best information is given by [L1]: c_n decays exponentially (this was actually used to obtain summability of q_n in the above argument). This improvement from exponential to torrential should give the reader an idea of the power of this kind of statistical analysis.

7.1. **How to estimate hyperbolicity.**

7.1.1. *Distribution of the hyperbolicity random variable.* Let us now explain how we will relate information on statistical parameters of typical non-regular parameters

f in order to conclude the Collet-Eckmann condition. We make several simplifications, in particular we don't discuss here the difficulty involved in working with capacities instead of probabilities.

Let us start by thinking of hyperbolicity at a given level n as a random variable $\lambda_n(j)$ (introduced in §11.2) which associates to each non-central branch of R_n its average expansion, that is, if $R_n|_{I_n^j} = f^{r_n(j)}$ ($r_n(j)$ is the return time of $R_n|_{I_n^j}$), we let $\lambda_n(j) = \ln |Df^{r_n(j)}|/r_n(j)$ evaluated at some point $x \in I_n^j$, say, the point where $|Df^{r_n(j)}|$ is minimal. (the choice of the point in I_n^j turns out to be not very relevant due to our almost sure bounds on distortion, see Lemma 8.10).

Our tactic is to evaluate the evolution of the distribution of $\lambda_n(j)$ as n grows. The basic information we will use to start our analysis is the hyperbolicity estimate of Lemma 3.1, which, together with our distortion estimates, shows that $\lambda_n = \inf_j \lambda_n(j) > 0$ for n big enough. We then fix such a big level n_0 and the remaining of the analysis will be based on inductive statistical estimates for levels $n > n_0$.

Of course, nothing guarantees a priori that λ_n does not decay to 0. Indeed, it turns out that $\liminf_{n\to\infty} \lambda_n > 0$, but as a consequence of the Collet-Eckmann condition and our distortion estimates. But this is not what we will analyze: we will concentrate on showing that $\lambda_n(j) > \lambda_{n_0}/2$ outside of a "bad set" of torrentially small γ-qs capacity. The complementary set of hyperbolic branches will be called good.

To do so, we inductively describe branches of level $n+1$ as compositions of branches of level n. Assuming that most branches of level n are good, we consider branches of level $n+1$ which spend most of their time in good branches of level n. They inherit hyperbolicity from good branches of level n, so they are themselves good of level $n+1$. To make this idea work we should also have additionally a condition of "not too close returns" to avoid drastic reduction of derivative due to the critical point.

The fact that most branches of level n were good (quantitatively: branches which are not good have capacity bounded by some small q_n) should reflect on the fact that most branches of level $n+1$ spend a small proportion of their time (less than $6q_n$) on branches which are not good, and so most branches of level $n+1$ are also good (capacity of the complement is a small q_{n+1}): indeed the notion of most should improve from level to level (so that $q_{n+1} \ll q_n$). This reflects the tendency of averages of random variables to concentrate around the expected value with exponentially small errors (Law of Large numbers and Law of Large deviations). Those laws give better results if we average over a larger number of random variables. In particular, those statistical laws are very effective in our case, since the number of random variables that we average will be torrential in n: our arguments will typically lead to estimates as $\ln q_{n+1}^{-1} > q_n^{-1+\epsilon}$ (torrential decay of q_n).

In practice, we will obtain good branches in a more systematic way. We extract from the above crude arguments a couple of features that should allow us to show that some branch is good. Those features define what we call a very good branch:

(1) for very good branches we can control the distance of the branch to 0 (to avoid drastic loss of derivative);
(2) the definition of very good branches has an inductive component: it must be a composition of many branches, most of which are themselves very good of the previous level (with the hope of propagating hyperbolicity inductively);

(3) the distribution of return times of branches of the previous level taking part in a very good branch has a controlled "concentration around the average".

Let us explain the third item above: to compute the hyperbolicity of a branch j of level $n+1$, which is a composition of several branches $j_1, ..., j_m$ of level n we are essentially estimating

$$\frac{\sum r_n(j_i)\lambda_n(j_i)}{\sum r_n(j_i)},$$

the rate of the total expansion over the total time of the branch. The second item assures us that many branches j_i are very good, but this does not mean that their total time is a reasonable part of the total time $r_{n+1}(j)$ of the branch j. This only holds if we can guarantee some concentration (in distribution) of the values of $r_n(j)$.

With those definitions we can prove that very good branches are good, but to show that very good branches are "most branches", we need to understand the distribution of the return time random variable $r_n(j)$ that we describe later.

Let us remark that our statistical work does not finish with the proof that good (hyperbolic) branches are most branches: this is enough to control hyperbolicity only at full returns. To estimate hyperbolicity at any moment of some orbit, we must use good branches as building blocks of hyperbolicity of some special branches of landing maps (cool landings). Branches which are not very good are sparse inside truncated cool landings, so that if we follow a piece of orbit of a point inside a cool landing, we still have enough hyperbolic blocks to estimate their hyperbolicity.

After all those estimates, we use the phase-parameter relation to move the critical value into cool landings, and obtain exponential growth of derivative of the critical value (with rate bounded from below by $\lambda_{n_0}/2$).

7.1.2. *The return time random variable.* As remarked above, to study the hyperbolicity random variable $\lambda_n(j)$, we must first estimate (in §10) the distribution of the return time random variable $r_n(j)$.

Intuitively, the "expectation" of $r_n(j)$ should be concentrated in a neighborhood of c_{n-1}^{-1}: pretending that iterates $f^k(x)$ are random points in I, we expect to wait about $|I|/|I_n|$ to get back to I_n. But $|I_n|/|I| = c_{n-1}c_{n-2}...c_1|I_1|/|I|$. Since c_n decays torrentially, we can estimate $|I_n|/|I|$ as $c_{n-1}^{1-\epsilon}$ (it is not worth to be more precise, since ϵ errors in the exponent will appear necessarily when considering capacities). Although this naive estimate turns out to be true, we of course don't try to follow this argument: we never try to iterate f itself, only return branches.

The basic information we use to start is again the hyperbolicity estimate of Lemma 3.1. This information gives us exponential tails for the distribution of $r_n(j)$ (the γ-qs capacity of $\{r_n(j) > k\}$ decays exponentially in k). Of course we have no information on the exponential rate: to control it we must again use an inductive argument which studies the propagation of the distribution of r_n from level to level. The idea again is that random variables add well and the relation between r_n and r_{n+1} is additive: if the branch j of level $n+1$ is the composition of branches j_i of level n, then $r_{n+1}(j)$ is the sum of $r_n(j_i)$.

Using that the transition from level to level involves adding a large number of random variables (torrential), we are able to give reasonable bounds for the decay of the tail of $r_n(j)$ for n big (this step is what we call a Large Deviation estimate). Once we control this tail, an estimate of the concentration of the distribution of return times becomes natural from the point of view of the Law of Large Numbers.

7.2. **Recurrence of the critical orbit.** After doing the preliminary work on the distribution of return times, the idea of the estimate on recurrence (in §12.2) is quite transparent. We first estimate the rate that a typical sequence $R_n^k(x)$ in I_n approaches 0, before falling into I_{n+1}. If the sequence $R_n^k(x)$ was random, then this recurrence would clearly be polynomial with exponent 1. The system of non-central branches is Markov with good estimates of distortion, so it is no surprise that $R_n^k(x)$ has the same recurrence properties, even if the system is not really random. We can then conclude that some inequality as

$$|R_n^k(x)| > |I_n| 2^{-n} k^{-1-\epsilon} \tag{7.1}$$

holds for most orbits (summable complement).

We must then relate the recurrence in terms of iterates of R_n to the recurrence in terms of iterates of f. Since in the Collet-Eckmann analysis we proved that (almost surely) the critical value belongs to a cool landing, it is enough to do the estimates inside a cool landing. But cool landings are formed by well distributed building blocks with good distribution of return times, so we can relate easily those two recurrence estimates.

To see that when we pass from the estimates in terms of iterations by R_n to iterations in term of f we still get polynomial recurrence, let us make a rough estimate which indicates that $R_n(0) = f^{v_n}(0)$ is at distance approximately v_n^{-1} of 0. Indeed $R_n(0)$ is inside I_n by definition, so $|R_n(0)| < c_{n-1}$. Using the phase-parameter relation, the critical orbit has controlled recurrence (in terms of (7.1)), thus we get $|R_n(0)| > 2^{-n}|I_n| > c_{n-1}^{1+\epsilon}$. On the other hand v_n (number of iterates of f before getting to I_n) is at least s_{n-1} (number of iterates of R_{n-1} before getting to I_n), thus as we saw before, $v_n > c_{n-1}^{-1+\epsilon}$. On the other hand, v_n is s_{n-1} times the average time of branches R_{n-1}: due to our estimates on the distribution of return times,

$$v_n < s_n c_{n-2}^{1-\epsilon} < c_{n-1}^{-1-\epsilon} c_{n-2}^{-1-\epsilon} < c_{n-1}^{-1-2\epsilon}$$

(0 is a "typical" point for the distribution of return times since it belongs to cool landings). Those estimates together give

$$1 - 4\epsilon < \frac{\ln |R_n(0)|}{\ln v_n} < 1 + 4\epsilon.$$

7.3. **Some technical details.** The statistical analysis described above is considerably complicated by the use of capacities: while traditional results of probability can be used as an inspiration for the proof (as outlined here), we can not actually use them. We also have to use statistical arguments which are adapted to tree decomposition of landings into returns: in particular, more sophisticated analytic estimates are substituted by more "bare-hands" techniques.

Following the details of the actual proof, the reader will notice that we work very often with a sequence of quasisymmetric constants which decrease from level to level but stays bounded away from 1. We don't work with a fixed capacity because, when adding random variables as above, some distortion is introduced. We can make the distortion small but not vanishing, and the distortion affects the constant of the next level: if we could make estimates of distribution using some constant γ_n, in the next level the estimates are in terms of a smaller constant γ_{n+1}. These ideas are introduced in §9.

Since the phase-parameter relation has two parts, our statistical analysis of the transition between two levels will very often involve two steps: one in order to move the critical value out of bad branches of the return map R_n, and another to move it inside a given branch of R_n outside of bad branches of the landing map L_n.

Fighting against the technical difficulties is the torrential decay of c_n. The concentration of statistical parameters related to level n are usually related to c_n or c_{n-1}, up to small exponential error. When statistical parameters of different levels interact, usually only one of them will determine the result. This is specially true since all our estimates include an ϵ error in the exponent. The reader should get used to estimates as "$c_n c_{n-1}$ is approximately c_n", in the sense that the rate of the logarithms of both quantities are actually close to 1 (compare the estimates in the end of the last section, specially relating s_n and v_n). Even if many proofs are quite technical, they are also quite robust due to this.

8. Statistics of the principal nest

8.1. **Decay of geometry.** Let as before $\tau_n \in \mathbb{Z}$ such that $R_n(0) \in I_n^{\tau_n}$.

An important parameter in our construction will be the scaling factor

$$c_n = \frac{|I_{n+1}|}{|I_n|}.$$

This variable of course changes inside each $J_n^{\tau_n}$ window, however, not by much. From PhPh1, for instance, we get that with total probability

$$\lim_{n\to\infty} \sup_{g_1,g_2 \in J_n^{\tau_n}} \frac{\ln(c_n[g_1])}{\ln(c_n[g_2])} = 1.$$

This variable is by far the most important on our analysis of the statistics of return maps. We will often consider other variables (say, return times): we will show that the distribution of those variables is concentrated near some average value. Our estimates will usually give a range of values near the average, and c_n will play an important role. Due (among other issues) to the variability of c_n inside the parameter windows, the ranges we select will depend on c_n up to an exponent (say $1-\epsilon$ to $1+\epsilon$), where ϵ is a small, but fixed, number. From the estimate we just obtained, for big n the variability (margin of error) of c_n will fall comfortably in such range, and we won't elaborate more.

A general estimate on the rates decay of c_n was obtained by Lyubich: he shows that (for a finitely renormalizable unimodal map with a recurrent critical point, c_{n_k} decays exponentially (on k), where $n_k - 1$ is the subsequence of non-central levels of f. For simple maps, the same is true with $n_k = k$, as there are only finitely many central cascades. Thus we can state:

Theorem 8.1 (see [L1])**.** *If f is a simple map then there exists $C > 0$, $\lambda < 1$ such that $c_n < C\lambda^n$.*

Let us use the following notation for the combinatorics of a point $x \in I_n$. If $x \in I_n^j$ we let $j^{(n)}(x) = j$ and if $x \in C_n^{\underline{d}}$ we let $\underline{d}^{(n)}(x) = \underline{d}$.

Lemma 8.2. *With total probability, for all n sufficiently big we have*

$$p_{2\gamma-1}(|\underline{d}^{(n)}(x)| \leq k | x \in I_n) < k c_n^{1-\epsilon/2}, \tag{8.1}$$

$$p_{2\gamma-1}(|\underline{d}^{(n)}(x)| \geq k | x \in I_n) < e^{-k c_n^{1+\epsilon/2}}. \tag{8.2}$$

We also have

$$p_{2\gamma-1}(|\underline{d}^{(n)}(x)| \leq k | x \in I_n^{\tau_n}) < k c_n^{1-\epsilon/2}, \tag{8.3}$$

$$p_{2\gamma-1}(|\underline{d}^{(n)}(x)| \geq k | x \in I_n^{\tau_n}) < e^{-k c_n^{1+\epsilon/2}}. \tag{8.4}$$

Proof. Let us compute the first two estimates.

Since I_n^0 is in the middle of I_n, we have as a simple consequence of the Real Schwarz Lemma (see [L1] and (8.7) in Lemma 8.5 below) that

$$\frac{c_n}{4} < \frac{|C_n^{\underline{d}}|}{|I_n^{\underline{d}}|} < 4c_n.$$

As a consequence

$$p_{2\gamma-1}(|\underline{d}^{(n)}(x)| = m | x \in I_n) < (4c_n)^{1-\epsilon/3}$$

and we get the estimate (8.1) summing up on $0 \leq m \leq k$.

For the same reason, we get that

$$\begin{aligned} &p_{2\gamma-1}(|\underline{d}^{(n)}(x)| > m | x \in I_n) \\ &\qquad < \left(1 - \left(\frac{c_n}{4}\right)^{1+\epsilon/3}\right) p_{2\gamma-1}(|\underline{d}^{(n)}(x)| \geq m | x \in I_n). \end{aligned}$$

This implies

$$p_{2\gamma-1}(|\underline{d}^{(n)}(x)| \geq m | x \in I_n) \leq \left(1 - \left(\frac{c_n}{4}\right)^{1+\epsilon/3}\right)^m.$$

Estimate (8.2) follows from

$$\begin{aligned} \left(1 - \left(\frac{c_n}{4}\right)^{1+\epsilon/3}\right)^k &< (1 - c_n^{1+\epsilon/2})^k \\ &< ((1 - c_n^{1+\epsilon/2})^{c_n^{-1-\epsilon/2}})^{k c_n^{1+\epsilon/2}} \\ &< e^{-k c_n^{1+\epsilon/2}}. \end{aligned}$$

The two remaining estimates are analogous. □

Let us now transfer this result (more precisely the second pair of estimates) to the parameter in each $J_n^{\tau_n}$ window using PhPa1. To do this notice that the measure of the complement of the set of parameters in $J_n^{\tau_n}$ such that $c_n^{-1+2\epsilon} < s_n < c_n^{-1-2\epsilon}$ can be estimated by $2c_n^{-\epsilon}$ for n big which is summable for all ϵ by Theorem 8.1. So we have:

Lemma 8.3. *With total probability,*

$$\lim_{n \to \infty} \frac{\ln(s_n)}{\ln(c_n^{-1})} = 1.$$

Remark 8.1. The parameter s_n influences the size of c_{n+1} in a determinant way. It is easy to see (using for instance the Real Schwarz Lemma, see [L1], see also item (8.8) in Lemma 8.5 below) that, in this setting, $\ln(c_{n+1}^{-1}) > K s_n$ for some constant K, which in general is universally bounded from below (real a priori bounds). (Due to decay of geometry, it turns out that $\ln(c_{n+1}^{-1})/s_n \to \infty$.)

As an easy consequence of the previous remark we get

Corollary 8.4. *With total probability,*

$$\liminf_{n\to\infty} \frac{\ln(\ln(c_{n+1}^{-1}))}{\ln(c_n^{-1})} \geq 1.$$

In particular, c_n decreases at least torrentially fast.

8.2. **Fine partitions.** We use Cantor sets K_n and $\tilde{K}_n$ to partition the phase space. In many circumstances we are directly concerned with intervals of this partition. However, sometimes we just want to exclude an interval of given size (usually a neighborhood of 0). This size does not usually correspond to a union of gaps, so we instead should consider in applications an interval which is union of gaps, with approximately the given size. The degree of relative approximation will always be torrentially good (in n), so we usually won't elaborate on this. In this section we just give some results which will imply that the partition induced by the Cantor sets are fine enough to allow torrentially good approximations.

The following lemma summarizes the situation. The proof is based on estimates of distortion using the Real Schwarz Lemma and the Koebe Principle (see [L1]) and is very simple, so we just sketch the proof.

Lemma 8.5. *The following estimates hold:*

$$\frac{|I_n^j|}{|I_n|} = O(\sqrt{c_{n-1}}), \tag{8.5}$$

$$\frac{|I_n^{\underline{d}}|}{|I_n^{\sigma^+(\underline{d})}|} = O(\sqrt{c_{n-1}}), \tag{8.6}$$

$$\frac{c_n}{4} < \frac{|C_n^{\underline{d}}|}{|I_n^{\underline{d}}|} < 4c_n, \tag{8.7}$$

$$\frac{|\tilde{I}_{n+1}|}{|I_n|} = O(e^{-s_{n-1}}). \tag{8.8}$$

Proof. (Sketch.) Since $R_n^{\underline{d}}$ has negative Schwarzian derivative, it immediately follows that the Koebe space[4] of $C_n^{\underline{d}}$ inside $I_n^{\underline{d}}$ has at least order c_n^{-1}.

It is easy to see that $R_{n-1}|_{I_n}$ can be written as $\phi \circ f$ where ϕ extends to a diffeomorphism onto I_{n-2} with negative Schwarzian derivative and thus with very small distortion. Since $R_{n-1}(I_n^j)$ is contained on some $C_{n-1}^{\underline{d}}$, we see that the Koebe space of I_n^j in I_n is at least of order $c_{n-1}^{-1/2}$ which implies (8.5).

Let us now consider an interval $I_n^{\underline{d}}$. Let I_n^j be such that $R_n^{\sigma^+(\underline{d})}(I_n^{\underline{d}}) = I_n^j$. We can pullback the Koebe space of I_n^j inside I_n by $R_n^{\sigma^+(\underline{d})}$, so (8.5) implies (8.6). Moreover, this shows by induction that the Koebe space of $I_n^{\underline{d}}$ inside I_n is at least of order $c_{n-1}^{-|\underline{d}|/2}$. Since $R_n(\tilde{I}_{n+1}) \subset I_n^{\underline{d}}$ with $|\underline{d}| = s_{n-1}$, the Koebe space of $\tilde{I}_{n+1}$ in I_n is at least $c_{n-1}^{-|\underline{d}|/4}$, which implies (8.8).

[4]The Koebe space of an interval T' inside an interval $T \supset T'$ is the minimum of $|L|/|T'|$ and $|R|/|T'|$ where L and R are the components of $T \setminus T'$. If the Koebe space of T' inside T is big, then the Koebe Principle states that a diffeomorphism T' which has an extension with negative Schwarzian derivative T has small distortion. In this case, it follows that the Koebe space of the preimage of T' inside the preimage of T is also big.

It is easy to see that $R_n|_{I_n^{\underline{d}}}$ can be written as $\phi \circ f \circ R_n^{\sigma^+(\underline{d})}$, where ϕ has small distortion. Due to (8.5), $R_n^{\sigma^+(\underline{d})}|_{I_n^{\underline{d}}}$ also has small distortion, so a direct computation with f (which is purely quadratic) gives (8.7). $\square$

In other words, distances in I_n can be measured with precision $\sqrt{c_{n-1}}|I_n|$ in the partition induced by $\tilde{K}_n$, due to (8.5) and (8.8) (since $e^{-s_{n-1}} \ll c_{n-1}$).

Distances can be measured much more precisely with respect to the partition induced by K_n, in fact we have good precision in each $I_n^{\underline{d}}$ scale. In other words, inside $I_n^{\underline{d}}$, the central gap $C_n^{\underline{d}}$ is of size $O(c_n|I_n^{\underline{d}}|)$ (by (8.7)) and the other gaps have size $O(\sqrt{c_{n-1}}|C_n^{\underline{d}}|)$ (by (8.6) and (8.7)).

Remark 8.2. We need to consider intervals which are union of gaps due to our phrasing of the phase-parameter relation, which only gives information about such gaps. However, this is not absolutely necessary, and we could have proceeded in a different way: our proof of the phase-parameter relation actually shows that there is a holonomy map with good qs estimates between phase and parameter intervals (and not only Cantor sets). While this map is not canonical, the fact that it is a holonomy map for a motion with good phase-phase estimates would allow our proofs to work.

8.3. **Initial estimates on distortion.** To deal with the distortion control we need some preliminary known results. We won't get too much in details here, those estimates are related to the estimates on gaps of Cantor sets and the Koebe Principle, and can be concluded easily.

Proposition 8.6. *For any j, if $R_n|_{I_n^j} = f^k$ then* $\operatorname{dist}(f^{k-1}|_{f(I_n^j)}) = 1 + O(c_{n-1})$.

For any $\underline{d}$, $\operatorname{dist}(R_n^{\sigma^+(\underline{d})}|_{I_n^{\underline{d}}}) = 1 + O(\sqrt{c_{n-1}})$.

We will use the following immediate consequence for the decomposition of certain branches.

Lemma 8.7. *With total probability,*

(1) $R_n|_{I_n^0} = \phi \circ f$ *where ϕ has torrentially small distortion,*

(2) $R_n^{\underline{d}} = \phi_2 \circ f \circ \phi_1$ *where ϕ_2 and ϕ_1 have torrentially small distortion and* $\phi_1 = R_n^{\sigma^+(\underline{d})}$.

8.4. **Estimating derivatives.**

Lemma 8.8. *Let w_n denote the relative distance in I_n of $R_n(0)$ to 0 and to the boundary of I_n, that is,*

$$w_n = \frac{d(R_n(0), \partial I_n \cup \{0\})}{|I_n|}$$

(where $d(x, X)$ denotes the usual Euclidean distance between a point and a set). With total probability,

$$\limsup_{n \to \infty} \frac{-\ln(w_n)}{\ln(n)} \leq 1.$$

In particular $R_n(0) \notin \tilde{I}_{n+1}$ for all n large enough.

Proof. This is a simple consequence of PhPa2, using that $n^{-1-\delta}$ is summable, for all $\delta > 0$. $\square$

From now on we suppose that f satisfies the conditions of the above lemma.

Lemma 8.9. *With total probability,*

$$\limsup_{n\to\infty} \frac{\sup_{j\neq 0} \ln(\operatorname{dist}(f|_{I_n^j}))}{\ln(n)} \leq 1/2.$$

Proof. Denote by $P_n^{\underline{d}}$ a $|C_n^{\underline{d}}|/n^{1+\delta}$ neighborhood of $C_n^{\underline{d}}$. Notice that the gaps of the Cantor sets K_n inside $I_n^{\underline{d}}$ which are different from $C_n^{\underline{d}}$ are torrentially (in n) smaller then $C_n^{\underline{d}}$, so we can take $P_n^{\underline{d}}$ as a union of gaps of K_n up to torrentially small error.

It is clear that if h is a γ-qs homeomorphism (γ close to 1) then

$$|h(P_n^{\underline{d}} \setminus C_n^{\underline{d}})| \leq n^{-1-\delta/2}|h(C_n^{\underline{d}})|$$

Notice that if $C_n^{\underline{d}}$ is contained in I_n^j with $j \neq \tau_n$, then $P_n^{\underline{d}}$ does not intersect $I_n^{\tau_n}$. Since the $C_n^{\underline{d}}$ are disjoint,

$$p_\gamma(I_n^{\tau_n} \cap \cup(P_n^{\underline{d}} \setminus C_n^{\underline{d}})|I_n^{\tau_n}) \leq n^{-1-\delta/2}$$

which is summable.

Transferring this estimate to the parameter using PhPa1 we see that with total probability, if n is sufficiently big, if $R_n(0)$ does not belong to $C_n^{\underline{d}}$ then $R_n(0)$ does not belong to $P_n^{\underline{d}}$ as well. In particular, if n is sufficiently big, the critical point 0 will never be in a $n^{-1/2-\delta/5}$ neighborhood of any I_{n+1}^j with $j \neq 0$ (the change from $n^{-1-\delta}$ to $n^{-1/2-\delta/5}$ is due to taking the inverse image by $R_n|_{I_{n+1}}$, which corresponds, up to torrentially small distortion, to taking a square root, and causes the division of the exponent by two). $\square$

Lemma 8.10. *With total probability,*

$$\limsup_{n\to\infty} \frac{\sup_{\underline{d}} \ln(\operatorname{dist}(R_n^{\underline{d}}))}{\ln(n)} \leq \frac{1}{2}. \tag{8.9}$$

In particular, for n big enough, $\sup_{|\underline{d}|\neq 0} \operatorname{dist}(R_n^{\underline{d}}) \leq 2^n$ and $|DR_n(x)| > 2$, $x \in \cup_{j\neq 0} I_n^j$.

Proof. By Lemma 8.7, Lemma 8.9 implies (8.9). If $j \neq 0$, by (8.5) of Lemma 8.5 we get that $|R_n(I_n^j)|/|I_n^j| = |I_n|/|I_n^j| > c_{n-1}^{-1/3}$, so $\operatorname{dist}(R_n|_{I_n^j}) \leq 2^n$ implies that for all $x \in I_n^j$, $|DR_n(x)| > c_{n-1}^{-1/3} 2^{-n} > 2$. $\square$

Remark 8.3. Lemma 8.9 has also an application for approximation of intervals. The result implies that if $I_n^j = (a,b)$ and $j \neq 0$, we have $1/2^n < b/a < 2^n$. As a consequence, for any symmetric (about 0) interval $I_{n+1} \subset X \subset I_n$, there exists a symmetric (about 0) interval $X \subset \tilde{X}$, which is union of I_n^j and such that $|\tilde{X}|/|X| < 2^n$ (approximation by union of $C_n^{\underline{d}}$, with $|\tilde{X}|/|X|$ torrentially close to 1, follows more easily from the discussion on fine partitions).

We will also need to estimate derivatives of iterates of f, and not only of return branches.

Lemma 8.11. *With total probability, if n is sufficiently big and if $x \in \cup_{j\neq 0} I_n^j$ and $R_n|_{I_n^j} = f^K$, then for $1 \leq k \leq K$, $|(Df^k(x))| > |x| c_{n-1}^3$.*

Proof. First notice that by Lemma 8.8 and Lemma 8.7, $R_n|_{I_n^0} = \phi \circ f$ with $|D\phi| > 1$, provided n is big enough (since ϕ has small distortion and there is a big macroscopic expansion from $f(I_n^0)$ to $R_n(I_n^0)$). Also, by Lemma 8.4, $|I_n|$ decay so fast that $\prod_{j=1}^n |I_n| > c_{n-1}^{3/2}$ for n big enough. Finally, by Lemma 8.10, for n big enough, $|DR_n(x)| > 1$ for $x \in I_n^j$, $j \neq 0$. Let n_0 be so big that if $n \geq n_0$, all the above properties hold.

From hyperbolicity of f restricted to the complement of I_{n_0} (from Lemma 3.1), there exists a constant $C > 0$ such that if $f^s(x) \notin I_{n_0}^0$, $r \leq s < k$ then $|Df^{k-r}(f^r(x))| > C$.

Let us now consider some $n \geq n_0$. If $k = K$, we have a full return and the result follows from Lemma 8.10.

Assume now $k < K$. Let us define $d(s)$, $0 \leq s \leq k$ such that $f^s(x) \in I_{d(s)} \setminus I_{d(s)}^0$ (if $f^s(x) \notin I_0$ we set $d(s) = -1$). Let $m(s) = \max_{s \leq t \leq k} d(t)$. Let us define a finite sequence $\{k_j\}_{j=0}^l$ as follows. We let $k_0 = 0$ and supposing $k_j < k$ we let $k_{j+1} = \max\{k_j < s \leq k | d(s) = m(s)\}$. Notice that $d(k_i) < n$ if $i \geq 1$, since otherwise $f^{k_i}(x) \in I_n$ so $k = k_i = K$ which contradicts our assumption.

The sequence $0 = k_0 < k_1 < ... < k_l = k$ satisfies $n = d(k_0) > d(k_1) > ... > d(k_l)$. Let θ be maximal with $d(k_\theta) \geq n_0$. We have of course

$$|Df^{k-k_\theta}(f^{k_\theta}(x))| > C|Df(f^{k_\theta}(x))|,$$

so if $\theta = 0$ then $Df^k(x) > |2Cx|$ and we are done.

Assume now $\theta > 0$. We have of course

$$|Df^{k-k_\theta}(f^{k_\theta}(x))| > C|Df^{k_\theta}(x)| > C|I_{d(k_\theta)+1}|$$

For $1 \leq j \leq \theta$, $f^{k_j - k_{j-1}}(x)$ is obtained by applying the central component of $R_{d(k_j)}$ followed by several non-central components of $R_{d(k_j)}$. Since $d(k_j) \geq n_0$, we can estimate

$$|Df^{k_j-k_{j-1}}(f^{k_{j-1}}(x))| > |DR_{d(k_j)}(f^{k_{j-1}}(x))| > |Df(f^{k_{j-1}}(x))|.$$

For $j = 1$, this argument gives $|Df^{k_1}(x)| \geq |Df(x)|$, while for $j > 1$ we can estimate

$$|Df^{k_j-k_{j-1}}(f^{k_{j-1}}(x))| > |Df(f^{k_{j-1}}(x))| > |I_{d(k_{j-1})+1}|.$$

Combining it all we get

$$\begin{aligned} |Df^k(x)| &= |Df^{k_1}(x)| \cdot |Df^{k-k_\theta}(f^{k_\theta}(x))| \prod_{j=2}^{\theta} |Df^{k_j-k_{j-1}}(f^{k_{j-1}}(x))| \\ &> |2x| \cdot C \cdot |I_{d(k_\theta)+1}| \prod_{j=2}^{\theta} |I_{d(k_{j-1})+1}| = |2Cx| \prod_{j=1}^{\theta} |I_{d(k_j)+1}| \\ &\geq |2Cx| \prod_{j=1}^{n} |I_n| > |x| c_{n-1}^3. \end{aligned}$$

□

9. Sequence of quasisymmetric constants and trees

9.1. **Preliminary estimates.** From now on, we will need to consider not only γ-capacities with some γ fixed, but different constants for different levels of the principal nest. To do so, we will make use of sequence of constants converging (decreasing) to a given value γ. We recall that γ is some constant very close to 1 such that $k(2\gamma-1)<1+\epsilon/5$, with ϵ very small.

We define the sequences $\rho_n=(n+1)/n$ and $\tilde{\rho}_n=(2n+3)/(2n+1)$, so that $\rho_n>\tilde{\rho}_n>\rho_{n+1}$ and $\lim\rho_n=1$. We define the sequence $\gamma_n=\gamma\rho_n$ and an intermediate sequence $\tilde{\gamma}_n=\gamma\tilde{\rho}_n$.

As we know, renormalization process has two phases, first R_n to L_n and then L_n to R_{n+1}. The following remarks shows why it is useful to consider the sequence of quasisymmetric constants due to losses related to distortion.

Remark 9.1. Let S be an interval contained in I^d_n. Using Lemma 8.7 we have $R^d_n|_S=\psi_2\circ f\circ\psi_1$, where the distortion of ψ_2 and ψ_1 are torrentially small and $\psi_1(S)$ is contained in some I^j_n, $j\neq 0$. If S is contained in I^0_n we may as well write $R_n|_S=\phi\circ f$, with $\operatorname{dist}(\phi)$ torrentially small.

In either case, if we decompose S in $2km$ intervals S_i of equal length, where k is the distortion of either $R^d_n|_S$ or $R_n|_S$ and m is subtorrentially big (say, $m<2^n$), the distortion obtained restricting to any interval S_i will be bounded by $1+1/m$.

Remark 9.2. Let us fix now γ such that the corresponding ϵ is small enough. We have the following estimate for the effect of the pullback of a subset of I_n by the central branch $R_n|_{I^0_n}$. With total probability, for all n sufficiently big, if $X\subset I_n$ satisfies

$$p_{\tilde{\gamma}_n}(X|I_n)<\delta\leq n^{-1000}$$

then

$$p_{\gamma_{n+1}}((R_n|_{I_{n+1}})^{-1}(X)|I_n)<\delta^{1/5}.$$

Indeed, let V be a $\delta^{1/4}|I_{n+1}|$ neighborhood of 0. Then $R_n|_{I_{n+1}\setminus V}$ has distortion bounded by $2\delta^{1/4}$.

Let $W\subset I_n$ be an interval of size $\lambda|I_n|$. Of course

$$p_{\tilde{\gamma}_n}(X\cap W|W)<\delta\lambda^{-1-\epsilon}.$$

Let us decompose each side of $I_{n+1}\setminus V$ as a union of $n^3\delta^{-1/4}$ intervals of equal length. Let W be such an interval. From Lemma 8.8, it is clear that the image of W covers at least $\delta^{1/2}n^{-4}|I_n|$. It is clear then that

$$p_{\tilde{\gamma}_n}(X\cap R_n(W)|R_n(W))<\delta^{(1-\epsilon)/2}n^{4+4\epsilon}.$$

So we conclude that (since the distortion of $R_n|_W$ is of order $1+n^{-3}$)

$$p_{\gamma_{n+1}}((R_n|_{I_{n+1}})^{-1}(X)\cap W|W)<\delta^{(1-\epsilon)/2}n^5$$

(we use the fact that the composition of a γ_{n+1}-qs map with a map with small distortion in $\tilde{\gamma}_n$-qs). Since

$$p_{\gamma_{n+1}}(V|I_{n+1})<(2\delta^{1/4})^{1-\epsilon},$$

we get the required estimate.

9.2. **More on trees.** Let us see an application of the above remarks.

Lemma 9.1. *With total probability, for all n sufficiently big*

$$p_{\tilde{\gamma}_n}((R^d_n)^{-1}(X)|I^d_n)<2^n p_{\gamma_n}(X|I_n).$$

Proof. Decompose $I_n^{\underline{d}}$ in $n^{\ln(n)}$ intervals of equal length, say, $\{W_i\}_{i=1}^{n^{\ln(n)}}$. Then by Lemma 8.10, $|R_n^{\underline{d}}(W_i)| > n^{-2\ln n}|I_n|$, so we get

$$p_{\gamma_n}(R_n^{\underline{d}}(W_i) \cap X | R_n^{\underline{d}}(W_i)) < n^{4\ln(n)} p_{\gamma_n}(X|I_n).$$

Applying Remark 9.1, we see that

$$p_{\tilde{\gamma}_n}((R_n^{\underline{d}})^{-1}(X) \cap W_i | W_i) < n^{4\ln(n)} p_{\gamma_n}(X|I_n),$$

(we use the fact that the composition of a $\tilde{\gamma}_n$-qs map with a map with small distortion in γ_n-qs) which implies the desired estimate. $\square$

By induction we get:

Lemma 9.2. *With total probability, for n is big enough, if $X_1, ..., X_m \subset \mathbb{Z} \setminus \{0\}$*

$$p_{\tilde{\gamma}_n}(\underline{d}^{(n)}(x) = (j_1, ..., j_m, ..., j_{|\underline{d}^{(n)}(x)|}), j_i \in X_i, 1 \le i \le m | x \in I_n)$$
$$\le 2^{mn} \prod_{i=1}^{m} p_{\gamma_n}(j^{(n)}(x) \in X_i | x \in I_n).$$

The following is an obvious variation of the previous lemma fixing the start of the sequence.

Lemma 9.3. *With total probability, for n is big enough, if $X_1, ..., X_m \subset \mathbb{Z} \setminus \{0\}$, and if $\underline{d} = (j_1, ..., j_k)$ we have*

$$p_{\tilde{\gamma}_n}(\underline{d}^{(n)}(x) = (j_1, ..., j_k, j_{k+1}, ..., j_{k+m}, ..., j_{|\underline{d}^{(n)}(x)|}), j_{i+k} \in X_i, 1 \le i \le m | x \in I_n^{\underline{d}})$$
$$\le 2^{mn} \prod_{i=1}^{m} p_{\gamma_n}(j^{(n)}(x) \in X_i | x \in I_n).$$

In particular, with $\underline{d} = (\tau_n)$,

$$p_{\tilde{\gamma}_n}(\underline{d}^{(n)}(x) = (\tau_n, j_1, ..., j_m, j_{m+1}, ..., j_{|\underline{d}^{(n)}(x)|}), j_i \in X_i, 1 \le i \le m | x \in I_n^{\tau_n})$$
$$\le 2^{mn} \prod_{i=1}^{m} p_{\gamma_n}(j^{(n)}(x) \in X_i | x \in I_n).$$

The last part of the above lemma will be often necessary in order to apply PhPa1.

Sometimes we are more interested in the case where the X_i are all equal.

Let $Q \subset \mathbb{Z} \setminus \{0\}$. Let $Q(m, k)$ denote the set of $\underline{d} = (j_1, ..., j_m)$ such that $\#\{i | j_i \in Q\} \cap \{1, ..., m\} \ge k$.

Define $q_n(m, k) = p_{\tilde{\gamma}_n}(\cup_{\underline{d} \in Q(m,k)} I_n^{\underline{d}} | I_n)$.

Let $q_n = p_{\gamma_n}(\cup_{j \in Q} I_n^j | I_n)$.

Lemma 9.4. *With total probability, for n large enough,*

$$q_n(m, k) \le \binom{m}{k} (2^n q_n)^k. \tag{9.1}$$

Proof. We have the following recursive estimates for $q_n(m, k)$:

(1) $q_n(1, 0) = 1$, $q_n(1, 1) \le q_n \le 2^n q_n$, and $q_n(m + 1, 0) \le 1$ for $m \ge 1$,
(2) $q_n(m + 1, k + 1) \le q_n(m, k + 1) + 2^n q_n q_n(m, k)$.

Indeed, (1) is completely obvious and if $(j_1, ..., j_{m+1}) \in Q(m + 1, k + 1)$ then either $(j_1, ..., j_m) \in Q(m, k + 1)$ or $(j_1, ..., j_m) \in Q(m, k)$ and $j_{m+1} \in Q$, so (2) follows from Lemma 9.1. It is clear that (1) and (2) imply by induction (9.1). $\square$

We recall that by Stirling Formula,

$$\binom{m}{qm} < \frac{m^{qm}}{(qm)!} < \left(\frac{3}{q}\right)^{qm}.$$

So we can get the following estimate. For $q \geq q_n$,

$$q_n(m, (6 \cdot 2^n)qm) < \left(\frac{1}{2}\right)^{(6 \cdot 2^n)qm}. \tag{9.2}$$

It is also used in the following form. If $q^{-1} > 6 \cdot 2^n$ (it is usually the case, since q will be torrentially small)

$$\sum_{k > q^{-2}} q_n(k, (6 \cdot 2^n)qk) < 2^{-n} q^{-1} \left(\frac{1}{2}\right)^{(6 \cdot 2^n)q^{-1}}. \tag{9.3}$$

This can be easily interpreted as a large deviation estimate in this context.

10. Estimates on time

Our aim in this section is to estimate the distribution of return times to I_n: they are concentrated around c_{n-1}^{-1} up to an exponent close to 1.

The basic estimate is a large deviation estimate which is proven in the next subsection (Corollary 10.5) and states that for $k \geq 1$ the set of branches with time larger then kc_n^{-4} has capacity less then e^{-k}.

10.1. **A Large Deviation lemma for times.** Let $r_n(j)$ be such that $R_n|_{I_n^j} = f^{r_n(j)}$. We will also use the notation $r_n(x) = r_n(j^{(n)}(x))$, the n-th return time of x (there should be no confusion for the reader, since we consistently use j for an integer index and x for a point in the phase space).

Let

$$A_n(k) = p_{\gamma_n}(r_n(x) \geq k | x \in I_n)$$

Since f restricted to the complement of I_{n+1} is hyperbolic, from Lemma 3.1, it is clear that $A_n(k)$ decays exponentially with k:

Lemma 10.1. *With total probability, for all $n > 0$, there exists $C > 0$, $\lambda > 1$ such that $A_n(k) < C\lambda^k$.*

Proof. Consider a Markov partition for $f|_{I \setminus I_{n+1}}$, that is, a finite union of intervals $M_1, ..., M_m$ such that $\cup M_i = I \setminus I_{n+1}$, $f|_{M_i}$ is a diffeomorphism and $f(M_i)$ is either the union of some M_j or the union of some M_j and I_{n+1}.

(To construct such Markov partition, let q be a periodic point in the forward orbit of ∂I_{n+1}, that is, $f^k(\partial I_{n+1}) = q$ where we choose k bigger than the period of q. Let K be the set of all x which never enter $\operatorname{int} I_{n+1}$ and such that $f^j(x) = q$ for some $j \leq k$, so that K is forward invariant. The Markov partition is just the set of connected components of $I \setminus (K \cup I_{n+1})$.)

It follows that if $f^j(x) \in \cup \operatorname{int} M_i$, $0 \leq j \leq k$ then there exists a unique interval $x \in M^k(x)$ such that $f^k|_{M^k(x)}$ is a diffeomorphism onto some M_j. By Lemma 3.1, $f^k|_{M^k(x)}$ has uniformly bounded distortion since f is C^2 (bounded distortion can also be obtained with the negative Schwarzian derivative).

By Lemma 3.1, the set of points $x \in I$ which never enter I_{n+1} has empty interior, so there exists $r > 0$ such that, for every M_j there exists $x \in M_j$ and $t_j < r$ with

$f^{t_j}(x) \in \operatorname{int} I_{n+1}$. It follows that there exists an interval $E_j \subset M_j$ such that $f^{t_j}(E_j) \subset \operatorname{int} I_{n+1}$.

Fixing some $M^k(x)$ with $f^k(M^k(x)) = M_j$, let $E^k(x) = (f^k|_{M^k(x)})^{-1}(E_j)$. By bounded distortion, it follows that $|E^k(x)|/|M^k(x)|$ is uniformly bounded from below independently of $M^k(x)$. In particular, $p_{2\gamma}(M^k(x) \setminus E^k(x)|M^k(x)) < \lambda$ for some constant $\lambda < 1$.

Let M^k be the union of the $M^k(x)$ and E^k be the union of the $E^k(x)$. Then $M^{k+r} \cap E^k = \emptyset$. In particular, $p_{2\gamma}(M^{(k+1)r}|I) < \lambda p_{2\gamma}(M^{kr}|I)$.

We conclude that $p_{2\gamma}(M^k|I_n) < C\lambda^{k/r}$ for some constant $C > 0$. If $k > r_n(0)$, then $M^k \cap I_n$ contains the set of points $x \in I_n$ such that $f^j(x) \notin I_n$, $1 \leq j \leq k$, that is, all points $x \in I_n$ with $r_n(x) > k$. Adjusting C and λ if necessary, we have $A_n(k) < C\lambda^k$. $\square$

Let ζ_n be the maximum $\zeta < c_{n-1}$ such that for all $k > \zeta^{-1}$ we have

$$A_n(k) \leq e^{-\zeta k}$$

and finally let $\alpha_n = \min_{1 \leq m \leq n} \zeta_m$.

Our main result in this section is to estimate α_n. We will show that with total probability, for n big we have $\alpha_{n+1} \geq c_n^4$. For this we will have to do a simultaneous estimate for landing times, which we define now.

Let $l_n(\underline{d})$ be such that $L_n|_{I_n^{\underline{d}}} = f^{l_n(\underline{d})}$. We will also use the notation $l_n(x) = l_n(\underline{d}^{(n)}(x))$.

Let

$$B_n(k) = p_{\tilde{\gamma}_n}(l_n(x) > k | x \in I_n).$$

Lemma 10.2. *If $k > c_n^{-3/2}\alpha_n^{-3/2}$ then*

$$B_n(k) < e^{-c_n^{3/2}\alpha_n^{3/2}k}. \tag{10.1}$$

Moreover,

$$p_{\tilde{\gamma}_n}(l_n(x) > k + r_n(x) | x \in I_n^{\tau_n}) < e^{-c_n^{-3/2}\alpha_n^{3/2}}. \tag{10.2}$$

Proof. We will just show (10.1), the proof of (10.2) being quite analogous.

Let $k > c_n^{-3/2}\alpha_n^{-3/2}$ be fixed. Let $m_0 = \alpha_n^{3/2}k$.

Notice that by Lemma 8.2

$$p_{\tilde{\gamma}_n}(|\underline{d}^{(n)}(x)| \geq m_0 | x \in I_n) \leq e^{-c_n^{5/4}\alpha_n^{3/2}k}.$$

Fix now $m < m_0$. Let us estimate

$$p_{\tilde{\gamma}_n}(|\underline{d}^{(n)}(x)| = m, l_n(x) > k | x \in I_n).$$

For each $\underline{d} = (j_1, ..., j_m)$ we can associate a sequence of m positive integers r_i such that $r_i \leq r_n(j_i)$ and $\sum r_i = k$. The average value of r_i is at least k/m so we conclude that

$$\sum_{r_i \geq k/2m} r_i > k/2.$$

Recall also that

$$\frac{k}{2m} > \frac{1}{(2\alpha_n^{3/2})} > \alpha_n^{-1}.$$

Given a sequence of m positive integers r_i as above we can do the following estimate using Lemma 9.2

$$\begin{aligned}
p_{\tilde{\gamma}_n}(\underline{d}^{(n)}(x) = (j_1, ..., j_m), r_n(j_i) > r_i | x \in I_n) \\
&\le 2^{mn} \prod_{j=1}^{m} p_{\gamma_n}(r_n(x) \ge r_j | x \in I_n) \\
&\le 2^{mn} \prod_{r_j \ge \alpha_n^{-1}} p_{\gamma_n}(r_n(x) \ge r_j | x \in I_n) \\
&\le 2^{mn} \prod_{r_j \ge k/2m} e^{-\alpha_n r_j} \\
&\le 2^{mn} e^{-\alpha_n k/2}.
\end{aligned}$$

The number of sequences of m positive integers r_i with sum k is

$$\begin{aligned}
\binom{k+m-1}{m-1} &\le \frac{1}{(m-1)!}(k+m-1)^{m-1} \\
&\le \frac{1}{m!}(k+m)^m \le \left(\frac{2ek}{m}\right)^m.
\end{aligned}$$

Notice that

$$\begin{aligned}
2^{mn}\left(\frac{2ek}{m}\right)^m &\le \left(\frac{2^{n+3}k}{m}\right)^{\frac{m}{k2^{n+3}}k2^{n+3}} \\
&\le \left(\frac{2^{n+3}k}{m_0}\right)^{\frac{m_0}{k2^{n+3}}k2^{n+3}} \qquad \text{(since } x^{1/x} \text{ is decreases for } x > e\text{)} \\
&\le \left(\frac{2^{n+3}}{\alpha_n^{3/2}}\right)^{m_0} \le e^{\alpha_n^{5/4}k}.
\end{aligned}$$

So we can finally estimate

$$\begin{aligned}
p_{\tilde{\gamma}_n}(|\underline{d}^{(n)}(x)| = m, l_n(x) \ge k | x \in I_n) &\le 2^{mn}\left(\frac{2ek}{m}\right)^m e^{-\alpha_n k/2} \\
&< e^{(\alpha_n^{1/4} - 1/2)\alpha_n k}.
\end{aligned}$$

Summing up on m we get

$$\begin{aligned}
p_{\tilde{\gamma}_n}(|\underline{d}^{(n)}(x)| < m_0, l_n(x) \ge k | x \in I_n) \\
&\le m_0 e^{(\alpha_n^{1/4} - 1/2)\alpha_n k} \\
&< e^{(2\alpha_n^{1/4} - 1/2)\alpha_n k} \qquad \left(\text{since} \frac{\ln(m_0)}{k} \le \frac{\ln(k)}{k} \le \alpha_n^{5/4}\right) \\
&\le e^{-\alpha_n k/3}.
\end{aligned}$$

As a direct consequence we get

$$B_n(k) < e^{-\alpha_n k/3} + e^{-c_n^{5/4}\alpha_n^{3/2}k} < e^{-c_n^{3/2}\alpha_n^{3/2}k}.$$

$\square$

Let $v_n = r_n(0)$ be the return time of the critical point.

Lemma 10.3. *With total probability, for n large enough,*

$$v_{n+1} < c_n^{-2}\alpha_n^{-2}/2.$$

Proof. By the definition of α_n and PhPa2, it follows that with total probability, for n large enough,

$$r_n(\tau_n) < c_{n-1}^{-1}\alpha_n^{-1}.$$

Recall that $\underline{d}^{(n)}(0)$ is such that $R_n(0) \in C_n^{\underline{d}^{(n)}(0)}$. Using Lemma 10.2, more precisely estimate (10.2), together with PhPa1, we get with total probability, for n large enough,

$$l_n(\underline{d}^{(n)}(0)) - r_n(\tau_n) < n\alpha_n^{-3/2}c_n^{-3/2},$$

and thus

$$v_{n+1} < v_n + c_{n-1}^{-1}\alpha_n^{-1} + n\alpha_n^{-3/2}c_n^{-3/2} < v_n + \alpha_n^{-2}c_n^{-2}/4.$$

Notice that α_n decreases monotonically, thus for n_0 big enough and for $n > n_0$,

$$v_n < v_{n_0} + \sum_{k=n_0}^{n-1} \alpha_k^{-2}c_k^{-2}/4 < v_{n_0} + \alpha_n^{-2}c_n^{-2}/3.$$

which for n big enough implies $v_{n+1} < c_n^{-2}\alpha_n^{-2}/2$. □

Lemma 10.4. *With total probability, for n large enough,*

$$\alpha_{n+1} \geq \min\{\alpha_n^4, c_n^4\}.$$

Proof. Let $k \geq \max\{\alpha_n^{-4}, c_n^{-4}\}$. From Lemma 10.3 one immediately sees that if $r_{n+1}(j) \geq k$ then $R_n(I_{n+1}^j))$ is contained on some $C_n^{\underline{d}}$ with $l_n(\underline{d}) \geq k/2 \geq \alpha_n^{-3/2}c_n^{-3/2}$.

Applying Lemma 10.2 we have $B_n(k/2) < e^{-\alpha_n^{3/2}c_n^{3/2}k/2}$.

Applying Remark 9.2 we get

$$A_{n+1}(k) < e^{-k\alpha_n^{3/2}c_n^{3/2}/200} < e^{-k\min\{\alpha_n^4, c_n^4\}}.$$

□

Since c_n decreases torrentially, we get

Corollary 10.5. *With total probability, for n large enough $\alpha_{n+1} \geq c_n^4$.*

Remark 10.1. In particular, using Lemma 10.3, for n big, $v_n < c_{n-1}^{-4}$.

10.2. **Consequences.**

Lemma 10.6. *With total probability, for all n sufficiently large we have*

$$p_{\tilde{\gamma}_n}(l_n(x) < c_n^{-1+\epsilon} | x \in I_n) < c_n^{\epsilon/2}, \tag{10.3}$$

$$p_{\tilde{\gamma}_n}(l_n(x) > c_n^{-1-5\epsilon/3} | x \in I_n) \leq e^{-c_n^{-\epsilon/4}}, \tag{10.4}$$

$$p_{\tilde{\gamma}_n}(l_n(x) - r_n(x) < c_n^{-1+\epsilon} | x \in I_n^{\tau_n}) < c_n^{\epsilon/2}, \tag{10.5}$$

$$p_{\tilde{\gamma}_n}(l_n(x) - r_n(x) > c_n^{-1-5\epsilon/3} | x \in I_n^{\tau_n}) \leq e^{-c_n^{-\epsilon/4}}. \tag{10.6}$$

Proof. We will concentrate on estimates (10.3) and (10.4), since (10.5) and (10.6) are analogous.

We have $l_n(\underline{d}) \geq |\underline{d}|$, and from Lemma 8.2

$$p_{\tilde{\gamma}_n}(|\underline{d}^{(n)}(x)| \leq c_n^{-1+\epsilon}|x \in I_n) \leq c_n^{\epsilon/2},$$

which implies (10.3).

On the other hand, by the same lemma,

$$p_{\tilde{\gamma}_n}(|\underline{d}^{(n)}(x)| \geq c_n^{-1-\epsilon}|x \in I_n) \leq e^{-c_n^{-\epsilon/2}}.$$

Defining

$$X_m = \bigcup_{\substack{\underline{d}=(j_1,\ldots,j_m),\\ r_n(j_m)>c_n^{-\epsilon/2}c_{n-1}^{-4}}} I_n^{\underline{d}}$$

we have

$$p_{\tilde{\gamma}_n}(X_m|I_n) \leq 2^n e^{-c_n^{-\epsilon/2}} < e^{-c_n^{-\epsilon/3}}.$$

Since

$$c_n^{-1-\epsilon}c_n^{-\epsilon/2}c_{n-1}^{-4} < c_n^{-1-5\epsilon/3},$$

we conclude that if x satisfies $l_n(x) > c_n^{-1-5\epsilon/3}$ and $|\underline{d}_n(x)| < c_n^{-1-\epsilon}$ then x belongs to some X_m with $1 \leq m \leq c_n^{-1-\epsilon}$. So we get

$$p_{\tilde{\gamma}_n}(l_n(x) > c_n^{-1-5\epsilon/3}|x \in I_n) \leq e^{-c_n^{-\epsilon/4}} \leq e^{-c_n^{-\epsilon/2}} + c_n^{-1-\epsilon}e^{-c_n^{-\epsilon/3}} < e^{-c_n^{-\epsilon/4}}$$

which implies (10.4). □

Corollary 10.7. *With total probability, for all n sufficiently large we have*

$$p_{\gamma_{n+1}}(r_{n+1}(x) < c_n^{-1+\epsilon}|x \in I_{n+1}) < c_n^{\epsilon/10}, \tag{10.7}$$

$$p_{\gamma_{n+1}}(r_{n+1}(x) > c_n^{-1-2\epsilon}|x \in I_{n+1}) \leq e^{-c_n^{-\epsilon/5}} \leq c_n^n. \tag{10.8}$$

Proof. Notice that $r_{n+1}(j) = v_n + l_n(\underline{d})$, where $R_n(I_{n+1}^j) \subset C_n^{\underline{d}}$. By Remark 10.1, we can estimate $v_n < c_{n-1}^{-4}$. The distribution of $r_{n+1}(j) - v_n$ can be then estimated by the distribution of $l_n(\underline{d})$ from Lemma 10.6, with a slight loss given by Remark 9.2. □

Using PhPa2 we get

Lemma 10.8. *With total probability, for all n sufficiently big*

$$\lim_{n\to\infty} \frac{\ln(r_n(\tau_n))}{\ln(c_{n-1}^{-1})} = 1. \tag{10.9}$$

Corollary 10.9. *With total probability, for all n sufficiently large we have*

$$p_{\tilde{\gamma}_n}(l_n(x) < c_n^{-1+\epsilon}|x \in I_n^{\tau_n}) \leq c_n^{\epsilon/10}, \tag{10.10}$$

$$p_{\tilde{\gamma}_n}(l_n(x) > c_n^{-1-11\epsilon/6}|x \in I_n^{\tau_n}) \leq e^{-c_n^{-\epsilon/5}}. \tag{10.11}$$

Proof. Just use Lemma 10.8 together with estimates (10.5) and (10.6) of Lemma 10.6. □

Corollary 10.10. *With total probability,*

$$\lim_{n\to\infty} \frac{\ln(v_{n+1})}{\ln(c_n^{-1})} = 1.$$

Proof. Notice that $v_{n+1} = v_n + l_n(\underline{d})$ where $R_n(0) \in C_n^{\underline{d}}$. Using Corollary 10.9 and PhPa1 we get $c_n^{-1+\epsilon} < l_n(\underline{d}) < c_n^{-1-11\epsilon/6}$. By Remark 10.1, $v_n < c_{n-1}^{-4}$, so $c_n^{-1+\epsilon} < v_{n+1} < c_n^{-1-2\epsilon}$. Letting ϵ go to 0 we get the result. □

Remark 10.2. Using Lemma 8.8, we see that $|R_n(I_{n+1})| > 2^{-n}|I_n|$. Since $|Df(x)| < 4$, $x \in I$, it follows that $|DR_n(x)| < 4^{v_n}$, $x \in I_{n+1}$. In particular, Corollary 10.10 implies that with total probability, for all $\epsilon > 0$, for all n big enough,

$$2^{-n}c_n^{-1} < \frac{|R_n(I_{n+1})|}{|I_{n+1}|} < 4^{v_n} < 4^{c_{n-1}^{-1-\epsilon}},$$

so $\ln(c_n^{-1}) < c_{n-1}^{-1-2\epsilon}$. This implies together with Corollary 8.4 that

$$\lim_{n\to\infty} \frac{\ln(\ln(c_n^{-1}))}{\ln(c_{n-1}^{-1})} = 1,$$

so c_n^{-1} grows torrentially (and not faster).

11. Dealing with hyperbolicity

In this section we show by an inductive process that the great majority of branches are reasonably hyperbolic. In order to do that, in the following subsection, we define some classes of branches with 'good' distribution of times and which are not too close to the critical point. The definition of 'good' distribution of times has an inductive component: they are composition of many 'good' branches of the previous level. The fact that most branches are good is related to the validity of some kind of Law of Large Numbers estimate.

11.1. Some kinds of branches and landings.

11.1.1. *Standard landings.* Let us define the set of standard landings of level n, $LS(n) \subset \Omega$ as the set of all $\underline{d} = (j_1, ..., j_m)$ satisfying the following.

LS1: (m not too small or large) $c_n^{-1/2} < m < c_n^{-1-2\epsilon}$,

LS2: (No very large times) $r_n(j_i) < c_{n-1}^{-14}$ for all i.

LS3: (Short times are sparse in large enough initial segments) For $c_{n-1}^{-2} \leq k \leq m$

$$\#\{r_n(j_i) < c_{n-1}^{-1+2\epsilon}\} \cap \{1, ..., k\} < (6 \cdot 2^n) c_{n-1}^{\epsilon/10} k,$$

LS4: (Large times are sparse in large enough initial segments) For $c_n^{-1/n} \leq k \leq m$

$$\#\{r_n(j_i) > c_{n-1}^{-1-2\epsilon}\} \cap \{1, ..., k\} < (6 \cdot 2^n) e^{-c_{n-1}^{\epsilon/5}} k.$$

Lemma 11.1. *With total probability, for all n sufficiently big we have*

$$p_{\tilde{\gamma}_n}(\underline{d}^{(n)}(x) \notin LS(n) | x \in I_n) < c_n^{1/3}/2, \tag{11.1}$$

$$p_{\tilde{\gamma}_n}(\underline{d}^{(n)}(x) \notin LS(n) | x \in I_n^{\tau_n}) < c_n^{1/3}/2. \tag{11.2}$$

Proof. Let us start with estimate (11.1) (on I_n). Let us estimate the complement of the set of landings which violate each item of the definition.

(LS1) This was already estimated before (see Lemma 8.2), an upper bound is $c_n^{1/3}/3$ (using ϵ small).

(LS2) By Corollary 10.5 the γ_n-capacity of $\{r_n(x) > c_{n-1}^{-14}\}$ is at most $e^{-c_{n-1}^{-10}} \ll c_n^3$. Using Lemma 9.1, we see that the $\tilde{\gamma}_n$-capacity of the set of $\underline{d} = (j_1, ..., j_m)$ with $r_n(j_i) > c_n^{-14}$ for some $i \leq c_n^{-1-2\epsilon}$ (in particular for some $i \leq m$ if m is as in LS1) is bounded by $2^n c_n^{-1-2\epsilon} c_n^3 \ll c_n$.

(LS3) This is a large deviation estimate, so we follow the ideas of §9.2, particularly estimate (9.2). Put $q = (6 \cdot 2^n) c_{n-1}^{\epsilon/10}$. By estimate (10.7) of Corollary 10.7, we can estimate the violation of LS3 for each fixed $c_{n-1}^{-2} \leq k \leq c_n^{-1-2\epsilon}$ by

$$\left(\frac{1}{2}\right)^{qk} \leq \left(\frac{1}{2}\right)^{c_{n-1}^{-3/2}} \ll c_n^3.$$

Summing up over k (and using the estimate by above on m) we get the upper bound c_n.

(LS4) We use the same method of the previous item. Put $q = (6 \cdot 2^n) e^{-c_{n-1}^{-\epsilon/5}}$. By estimate (10.8) of Corollary 10.7, we can bound the probability of violation of LS4 for each $c_n^{-1/n} \leq k \leq c_n^{-1-2\epsilon}$ by

$$\left(\frac{1}{2}\right)^{qk} \ll c_n^3.$$

Summing up in k (and using the estimate by above on m) we get the upper bound c_n.

Adding the losses of the four items, we get the estimate (11.1). To get estimate (11.2) (on $I_n^{\tau_n}$), the only item which changes is the second (since we have to avoid to have problems already because of the first iterate having very large time, which would automatically give $LS(n)$ not intersecting $I_n^{\tau_n}$). This was taken care in Lemma 10.8. $\square$

11.1.2. *Very good returns and excellent landings.* Define the set of very good returns, $VG(n_0, n) \subset \mathbb{Z} \setminus \{0\}$, $n_0, n \in \mathbb{N}$, $n \geq n_0$ by induction as follows. We let $VG(n_0, n_0) = \mathbb{Z} \setminus \{0\}$ and supposing $VG(n_0, n)$ defined, define the set of excellent landings $LE(n_0, n) \subset LS(n)$ satisfying the following extra assumption.

LE: (Not very good moments are sparse in large enough initial segments) For all $c_{n-1}^{-2} < k \leq m$

$$\#\{i | j_i \notin VG(n_0, n)\} \cap \{1, ..., k\} < (6 \cdot 2^n) c_{n-1}^{1/20} k,$$

And we define $VG(n_0, n+1)$ as the set of j such that $R_n(I_{n+1}^j) = C_n^{\underline{d}}$ with $\underline{d} \in LE(n_0, n)$ and the extra condition.

VG: (distant from 0) The distance of I_{n+1}^j to 0 is bigger than $c_n^{1/3} |I_{n+1}|$.

Lemma 11.2. *With total probability, for all n_0 sufficiently big and all $n \geq n_0$, if*

$$p_{\gamma_n}(j^{(n)}(x) \notin VG(n_0, n) | x \in I_n) < c_{n-1}^{1/20} \tag{11.3}$$

then

$$p_{\tilde{\gamma}_n}(\underline{d}^{(n)}(x) \notin LE(n_0, n) | x \in I_n) < c_n^{1/3}, \tag{11.4}$$

$$p_{\tilde{\gamma}_n}(\underline{d}^{(n)}(x) \notin LE(n_0, n) | x \in I_n^{\tau_n}) < c_n^{1/3}. \tag{11.5}$$

Proof. We first use Lemma 11.1 to estimate the $\tilde{\gamma}_n$-capacity of branches not is $LS(n)$ by $c_n^{1/3}/2$.

Let $q = (6 \cdot 2^n) c_{n-1}^{1/20}$. Using the hypothesis and estimate (9.2) of §9.2 (see also the estimate of the complement of LS3 in Lemma 11.1) we first estimate the $\tilde{\gamma}_n$-capacity of the set of landings which violate LE for a specific value of k with $k \geq c_{n-1}^{-2}$ by $(1/2)^{qk}$ and then summing up on k we get

$$\sum_{k \geq c_{n-1}^{-2}} \left(\frac{1}{2}\right)^{qk} \ll c_n.$$

This argument works both for (11.4) (in I_n) and (11.5) (in $I_n^{\tau_n}$). ☐

Lemma 11.3. *With total probability, for all n_0 sufficiently big and for all $n \geq n_0$,*

$$p_{\gamma_n}(j^{(n)}(x) \notin VG(n_0, n) | x \in I_n) < c_{n-1}^{1/20}. \tag{11.6}$$

Proof. It is clear that with total probability, for n_0 sufficiently big and $n \geq n_0$, the set of branches I_n^j at distance at least $c_{n-1}^{1/3}|I_n|$ of 0 has γ_n-capacity bounded by $c_{n-1}^{1/8}$.

For $n = n_0$, (11.6) holds (since all branches are very good except the central one). Using Lemma 11.2, if (11.6) holds for n then (11.4) also holds for n. Pulling back estimate (11.4) by $R_n|_{I_{n+1}}$ (using Remark 9.2), we get (11.6) for $n+1$. The result follows by induction on n. ☐

Using PhPa2 we get (using the measure-theoretical argument of Lemma 6.2)

Lemma 11.4. *With total probability, for all n_0 big enough, for all n big enough, $\tau_n \in VG(n_0, n)$.*

Lemma 11.5. *With total probability, for all n_0 big enough and for all $n \geq n_0$, if $j \in VG(n_0, n+1)$ then*

$$\frac{1}{2} m c_{n-1}^{-1+2\epsilon} < r_{n+1}(j) < 2 m c_{n-1}^{-1-2\epsilon},$$

where as usual, m is such that $R_n(I_{n+1}^j) = C_n^{\underline{d}}$ and $\underline{d} = (j_1, ..., j_m)$.

Proof. Notice that $r_{n+1}(j) = v_n + \sum r_n(j_i)$. To estimate the total time $r_{n+1}(j)$ from below we use LS3 and get

$$\frac{1}{2} m c_{n-1}^{-1+2\epsilon} < (1 - 6 \cdot 2^n c_n^{\epsilon/10}) m c_{n-1}^{-1+2\epsilon} < r_{n+1}(j).$$

To estimate from above, we notice $v_n < c_{n-1}^{-4}$ and

$$\sum_{r_n(j_i) > c_{n-1}^{-1-2\epsilon}} r_n(j_i) < 6 \cdot 2^n c_{n-1}^{-14} e^{-c_{n-1}^{-\epsilon/5}} m < m,$$

so

$$r_{n+1}(j) < m c_{n-1}^{-1-2\epsilon} + m + c_{n-1}^{-4} < 2 m c_{n-1}^{-1-2\epsilon}.$$

☐

Remark 11.1. Using LS1 we get the estimate $c_n^{-1/2} < r_{n+1}(j) < c_n^{-1-3\epsilon}$.

Let $j \in VG(n_0, n+1)$. We can write $R_{n+1}|_{I^j_{n+1}} = f^{r_{n+1}(j)}$, that is, a big iterate of f. One may consider which proportion of those iterates belong to very good branches of the previous level. More generally, we can truncate the return R_{n+1}, that is, we may consider $k < r_{n+1}(j)$ and ask which proportion of iterates up to k belong to very good branches.

Lemma 11.6. *With total probability, for all n_0 big enough and for all $n \geq n_0$, the following holds.*

Let $j \in VG(n_0, n+1)$, as usual let $R_n(I^j_{n+1}) = I^{\underline{d}}_n$ and $\underline{d} = (j_1, ..., j_m)$. Let m_k be biggest possible with

$$v_n + \sum_{j=1}^{m_k} r_n(j_i) \leq k$$

(the amount of full returns to level n before time k) and let

$$\beta_k = \sum_{\substack{1 \leq i \leq m_k, \\ j_i \in VG(n_0,n)}} r_n(j_i).$$

(the total time spent in full returns to level n which are very good before time k) Then $1 - \beta_k/k < c_{n-1}^{1/100}$ if $k > c_n^{-2/n}$.

Proof. Let us estimate first the time i_k which is not spent on non-critical full returns:

$$i_k = k - \sum_{j=1}^{m_k} r_n(j_i).$$

This corresponds exactly to v_n plus some incomplete part of the return $j_{m_{k+1}}$. This part can be bounded by $c_{n-1}^{-4} + c_{n-1}^{-14}$ (use Corollary 10.10 to estimate v_n and LS2 to estimate the incomplete part).

Using LS2 we conclude now that

$$m_k > (k - c_{n-1}^{-4} - c_{n-1}^{-14}) c_{n-1}^{14} > c_n^{-1/n}$$

so m_k is not too small.

Let us now estimate the contribution h_k from full returns j_i with time higher than $c_n^{-1-2\epsilon}$. Since m_k is big, we can use LS4 to conclude that the number of such high time returns must be less than $c_{n-1}^n m_k$, so their total time is at most $c_{n-1}^{n-14} m_k$.

The non very good full returns on the other hand can be estimated by LE (given the estimate on m_k), they are at most $c_{n-1}^{1/21} m_k$. So we can estimate the total time l_k of non very good full returns with time less then $c_{n-1}^{-1-2\epsilon}$ by

$$c_{n-1}^{1/25} c_{n-1}^{-1-2\epsilon} m_k.$$

Since m_k is big, we can use LS3 to estimate the proportion of branches with not too small time, so we conclude that at most $c_{n-1}^{\epsilon/11} m_k$ branches are not very good or have time less than $c_{n-1}^{-1+2\epsilon}$, so β_k can be estimated from below as

$$(1 - c_{n-1}^{\epsilon/11}) c_{n-1}^{-1+2\epsilon} m_k.$$

It is easy to see then that $i_k/\beta_k \ll c_{n-1}^{1/100}$, $h_k/\beta_k \ll c_{n-1}^{1/100}$. If ϵ is small enough, we also have

$$l_k/\beta_k < 2c_{n-1}^{1/25-4\epsilon} < c_{n-1}^{1/90}.$$

So $(i_k + h_k + l_k)/\beta_k$ is less then $c_{n-1}^{1/100}$. Since $i_k + h_k + l_k + \beta_k = k$ we have $1 - \beta_k/k < (i_k + h_k + l_k)/\beta_k$. □

11.1.3. *Cool landings.* Let us define the set of cool landings $LC(n_0, n) \subset \Omega$, $n_0, n \in \mathbb{N}$, $n \geq n_0$ as the set of all $\underline{d} = (j_1, ..., j_m)$ in $LE(n_0, n)$ satisfying

LC1: (Starts very good) $j_i \in VG(n_0, n)$, $1 \leq i \leq c_{n-1}^{-1/30}$,

LC2: (Short times are sparse in large enough initial segments) For $c_{n-1}^{-\epsilon/5} \leq k \leq m$

$$\#\{r_n(j_i) < c_{n-1}^{-1+2\epsilon}\} \cap \{1, ..., k\} < (6 \cdot 2^n) c_{n-1}^{\epsilon/10} k,$$

LC3: (Not very good moments are sparse in large enough initial segments) For all $c_{n-1}^{-1/30} < k \leq m$

$$\#\{i | j_i \notin VG(n_0, n)\} \cap \{1, ..., k\} < (6 \cdot 2^n) c_{n-1}^{1/60} k,$$

LC4: (Large times are sparse in large enough initial segments) For $c_{n-1}^{-200} \leq k \leq m$

$$\#\{r_n(j_i) > c_{n-1}^{-1-2\epsilon}\} \cap \{1, ..., k\} < (6 \cdot 2^n) c_{n-1}^{100} k,$$

LC5: (Starts with no large times) $r_n(j_i) < c_{n-1}^{-1-2\epsilon}$, $1 \leq i \leq e^{c_{n-1}^{-\epsilon/5}/2}$.

Notice that LC4 and LC5 overlap, since $c_{n-1}^{-200} < e^{c_{n-1}^{-\epsilon/5}/2}$ as do LC1 and LC3. From this we can conclude that we can control the proportion of large times or non very good times in all moments (and not only for large enough initial segments).

Lemma 11.7. *With total probability, for all n_0 sufficiently big and all $n \geq n_0$,*

$$p_{\tilde{\gamma}_n}(\underline{d}^{(n)}(x) \notin LC(n_0, n) | x \in I_n) < c_{n-1}^{1/100} \tag{11.7}$$

and for all n big enough

$$p_{\tilde{\gamma}_n}(\underline{d}^{(n)}(x) \notin LC(n_0, n) | x \in I_n^{\tau_n}) < c_{n-1}^{1/100}. \tag{11.8}$$

Proof. We follow the ideas of the proof of Lemma 11.1. Let us start with estimate (11.7). Notice that by Lemmas 11.3 and 11.2 we can estimate the $\tilde{\gamma}_n$-capacity of the complement of excellent landings by $c_n^{1/3}$. The computations below indicate what is lost going from excellent to cool due to each item of the definition:

(LC1) This is a direct estimate analogous to LS2. By Lemma 11.3, the γ_n-capacity of the complement of very good branches is bounded by $c_{n-1}^{1/20}$, so an upper bound for the $\tilde{\gamma}_n$-capacity of the set of landings which do not start with $c_{n-1}^{-1/30}$ very good branches is given by

$$2^n c_{n-1}^{1/20} c_{n-1}^{-1/30} \ll c_{n-1}^{1/100}.$$

(LC2) This is essentially the same large deviation estimate of LS3. We put $q = (6 \cdot 2^n) c_{n-1}^{\epsilon/10}$. By estimate (10.7) of Corollary 10.7, the $\tilde{\gamma}_n$-capacity of the set of landings violating LC2 for a specific value of k is bounded by $(1/2)^{qk}$, and summing up on k (see also estimate (9.3)) we get the upper bound

$$\sum_{k \geq c_{n-1}^{-\epsilon/5}} \left(\frac{1}{2}\right)^{(6 \cdot 2^n) c_{n-1}^{\epsilon/10} k} \leq (2^{-n} c_{n-1}^{-\epsilon/10}) \left(\frac{1}{2}\right)^{(6 \cdot 2^n) c_{n-1}^{-\epsilon/10}} \ll c_{n-1}^{1/100}.$$

(LC3) This is analogous to the previous item, we set $q=(6\cdot 2^n)c_{n-1}^{1/60}$ and using Lemma 11.3 we get an upper bound

$$\sum_{k\geq c_{n-1}^{-1/30}}\left(\frac{1}{2}\right)^{(6\cdot 2^n)c_{n-1}^{1/60}k}\leq(2^{-n}c_{n-1}^{-1/60})\left(\frac{1}{2}\right)^{2^n c_{n-1}^{-1/60}}\ll c_{n-1}^{1/100}.$$

(LC4) As before, we set $q=(6\cdot 2^n)c_{n-1}^{100}$ and using estimate (10.8) of Corollary 10.7 we get

$$\sum_{k\geq c_{n-1}^{-200}}\left(\frac{1}{2}\right)^{(6\cdot 2^n)c_{n-1}^{100}k}\leq(2^{-n}c_{n-1}^{-100})\left(\frac{1}{2}\right)^{2^n c_{n-1}^{-100}}\ll c_{n-1}^{1/100}.$$

(LC5) This is a direct estimate as LC1, using estimate (10.8) of Corollary 10.7 we get

$$2^n e^{-c_{n-1}^{\epsilon/5}}e^{c_{n-1}^{-\epsilon/5}/2}\ll c_{n-1}^{1/100}.$$

Putting those together, we obtain (11.7). For (11.8), we do the same and use Lemmas 11.4 and 10.8 to avoid problems for landing immediately in a not very good branch or with too small or large time. ☐

Transferring the result to the parameter, using PhPa1, we get (using the measure-theoretical argument of Lemma 6.2)

Lemma 11.8. *With total probability, for all n_0 big enough, for all n big enough we have $R_n(0)\in C_n^{\underline{d}}$ with $\underline{d}\in LC(n_0,n)$.*

11.2. Hyperbolicity.

11.2.1. *Preliminaries.* For $j\neq 0$, we define

$$\lambda_n(j)=\inf_{x\in I_n^j}\frac{\ln|DR_n(x)|}{r_n(j)}.$$

And $\lambda_n=\inf_{j\neq 0}\lambda_n(j)$. As a consequence of the exponential estimate on distortion for returns (which competes with torrential expansion from the decay of geometry), together with hyperbolicity of f in the complement of I_n^0 we immediately have the following

Lemma 11.9. *With total probability, for all n sufficiently big, $\lambda_n>0$.*

Proof. By Lemma 3.1, there exists a constant $\tilde{\lambda}_n>0$ such that each periodic orbit p of f whose orbit is entirely contained in the complement of I_{n+1} must satisfy $\ln|Df^m(p)|>\tilde{\lambda}_n m$, where m is the period of p. On the other hand, each non-central branch $R_n|_{I_n^j}$ has a fixed point. By Lemma 8.10, $\sup\operatorname{dist}(R_n|_{I_n^j})\leq 2^n$ and of course $\lim_{j\to\infty}r_n(j)=\infty$, so we have

$$\liminf_{n\to\infty}\lambda_n(j)\geq\tilde{\lambda}_n.$$

On the other hand, for any $j\neq 0$, $\lambda_n(j)>0$ by Lemma 8.10, so $\lambda_n>0$. ☐

11.2.2. *Good branches.* The "minimum hyperbolicity" $\liminf \lambda_n$ of the parameters we will obtain will in fact be positive, as it follows from one of the properties of Collet-Eckmann parameters (uniform hyperbolicity on periodic orbits, see [NS]), together with our estimates on distortion.

However, as we described before, our strategy is not to show that the minimum hyperbolicity is positive, but that the typical value of $\lambda_n(j)$ stays bounded away from 0 as n grows (and is in fact bigger than $\lambda_{n_0}/2$ for $n > n_0$ big). Since we also have to estimate the hyperbolicity of truncated branches it will be convenient to introduce a new class of branches with good hyperbolic properties.

We define the set of good returns $G(n_0,n) \subset \mathbb{Z} \setminus \{0\}$, $n_0, n \in \mathbb{N}$, $n \geq n_0$ as the set of all j such that

G1: (hyperbolic return)
$$\lambda_n(j) \geq \lambda_{n_0} \frac{1+2^{n_0-n}}{2},$$

G2: (hyperbolicity in truncated return) for $c_{n-1}^{-3/(n-1)} \leq k \leq r_n(j)$ we have
$$\inf_{x \in I_n^j} \frac{\ln |Df^k(x)|}{k} \geq \lambda_{n_0} \frac{1+2^{n_0-n+1/2}}{2} - c_{n-1}^{2/(n-1)}.$$

Notice that since c_n decreases torrentially, for n sufficiently big G2 implies that if j is good then for $c_{n-1}^{-3/(n-1)} \leq k \leq r_n(j)$ we have
$$\inf_{x \in I_n^j} \frac{\ln |Df^k(x)|}{k} \geq \lambda_{n_0} \frac{1+2^{n_0-n}}{2}.$$

Lemma 11.10. *With total probability, for all n_0 big enough and for all $n > n_0$, $VG(n_0,n) \subset G(n_0,n)$.*

Proof. Let us prove that if G1 is satisfied for all $j \in VG(n_0,n)$, then $VG(n_0,n+1) \subset G(n_0,n+1)$. Let us fix such a j. Notice that by definition of λ_{n_0} the hypothesis is satisfied for n_0. Let
$$a_k = \inf_{x \in I_{n+1}^j} \frac{\ln |Df^k(x)|}{k},$$
and let us assume that $k > c_n^{-3/n}$ (notice that if $k = r_{n+1}(j)$ this condition is automatically satisfied by Remark 11.1).

We let (as usual) $R_n(I_{n+1}^j) \subset C_n^{\underline{d}}$, $\underline{d} = (j_1, ..., j_m)$. Notice that by Remark 10.1, $v_n < c_{n-1}^{-4} < k$. Let us say that j_i was completed before k if $v_n + r_n(j_1) + ... + r_n(j_i) \leq k$. We let the queue be defined as
$$q_k = \inf_{x \in C_n^{\underline{d}}} \ln |Df^{k-r} \circ f^r(x)|$$
where $r = v_n + r_n(j_1) + ... + r_n(j_{m_k})$ with j_{m_k} the last complete return.

Let us show first that $|DR_n(x)| > 1$ if $x \in I_{n+1}^j$. Indeed, by Lemma 8.7, $DR_n|_{I_{n+1}} = \phi \circ f$, where ϕ has small distortion, so by Lemma 8.8,
$$|D\phi(x)| > \frac{|R_n(I_{n+1})|}{2|f(I_{n+1})|} > \frac{2^{-n}|I_n|}{|I_{n+1}|^2},$$
while by VG, $|Df(x)| = |2x| > c_n^{1/3}|I_{n+1}|$, so $|DR_n(x)| > c_n^{-1/2}$.

By Lemma 8.10, any complete return before k produces some expansion, (that is, the absolute value of the derivative of such return is at least 1). On the other hand, $-q_k$ can be bounded from above by $-\ln(c_n c_{n-1}^5)$ using Lemma 8.11. We have

$$-\frac{q_k}{k} < \frac{-\ln(c_n c_{n-1}^5)}{c_n^{-3/n}} \ll c_n^{2/n}.$$

Now we use Lemma 11.6 and get

$$\begin{aligned} a_k &> \frac{\beta_k}{k} \frac{\lambda_{n_0}(1+2^{n_0-n})}{2} - \frac{-q_k}{k} \\ &\geq \frac{\lambda_{n_0}(1+2^{n_0-n-1/2})}{2} - \frac{-q_k}{k} \end{aligned}$$

which gives G2. If $k = r_{n+1}(j)$ then $q_k = 0$, which gives G1. □

11.2.3. *Hyperbolicity in cool landings.*

Lemma 11.11. *With total probability, if n_0 is sufficiently big, for all n sufficiently big, if $\underline{d} \in LC(n_0, n)$ then for all $c_{n-1}^{-4/(n-1)} < k \leq l_n(\underline{d})$,*

$$\inf_{x \in C_n^{\underline{d}}} \frac{\ln|Df^k(x)|}{k} \geq \frac{\lambda_{n_0}}{2}.$$

Proof. Fix such $\underline{d} \in LC(n_0, n+1)$, and let as usual $\underline{d} = (j_1, ..., j_m)$. Let

$$a_k = \inf_{x \in C_n^{\underline{d}}} \frac{\ln|Df^k(x)|}{k}.$$

Analogously to Lemma 11.6, we define m_k number of full returns before k, so that m_k is the biggest integer such that

$$\sum_{i=1}^{m_k} r_n(j_i) \leq k.$$

We define

$$\beta_k = \sum_{\substack{1 \leq i \leq m_k, \\ j_i \in VG(n_0, n+1)}} r_n(j_i),$$

(counting the time up to k spent in complete very good returns) and

$$i_k = k - \sum_{i=1}^{m_k} r_n(j_i).$$

(counting the time in the incomplete return at k).

Let us now consider two cases: either all iterates are part of very good returns (that is, all j_i, $1 \leq i \leq m_k$ are very good and if $i_k > 0$ then j_{m_k+1} is also very good), or some iterates are not part of very good returns.

Case 1 (All iterates are part of very good returns). Since full good returns are very hyperbolic by G1 and very good returns are good, we just have to worry with possibly losing hyperbolicity in the incomplete time. To control this, we introduce the queue

$$q_k = \inf_{x \in C_{n-1}^{\underline{d}}} \ln|Df^{i_k} \circ f^{k-i_k}(x)|.$$

We have $-q_k < -\ln(c_{n-1}^{1/3} c_{n-1}^5)$ by Lemma 8.11 and VG, using that the incomplete time is in the middle of a very good branch. Let us split again in two cases: i_k big or otherwise.

Subcase 1a ($i_k > c_{n-1}^{-4/(n-1)}$). If the incomplete time is big, we can use G2 to estimate the hyperbolicity of the incomplete time (which is part of a very good return): $q_k/i_k > \lambda_{n_0}/2$. We have

$$a_k > \lambda_{n_0}\frac{(1+2^{n-n_0})}{2}\cdot\frac{k-i_k}{k}+\frac{q_k}{i_k}\cdot\frac{i_k}{k}>\frac{\lambda_{n_0}}{2}.$$

Subcase 1b ($i_k < c_{n-1}^{-4/(n-1)}$). If the incomplete time is not big, we can not use G2 to estimate q_k, but in this case i_k is much less than k: since $k > c_{n-1}^{-4/(n-1)}$, at least one return was completed ($m_k \geq 1$), and since it must be very good we conclude that $k > c_{n-1}^{-1/2}$ by Remark 11.1, so

$$a_k > \lambda_{n_0}\frac{(1+2^{n-n_0})}{2}\cdot\frac{k-i_k}{k}-\frac{-q_k}{k}>\frac{\lambda_{n_0}}{2}.$$

Case 2 (Some iterates are not part of a very good return). By LC1, $m_k > c_{n-1}^{-1/30}$. Notice that by LC2, if $m_k > c_{n-1}^{-\epsilon/5}$ then

$$k-i_k > c_{n-1}^{-1+2\epsilon}m_k/2.$$

So it follows that $m_k > c_{n-1}^{-1/30}$ implies that $k > c_{n-1}^{-35/34}$ (using small ϵ).

For the incomplete time we have $-q_k < -\ln(c_n c_{n-1}^5) < c_{n-1}^{-1-\epsilon}$, so $-q_k/k < c_{n-1}^{1/100}$.

Arguing as in Lemma 11.6, we split $k-\beta_k-i_k$ (time of full returns which are not very good) in part relative to returns with high time h_k (more than $c_{n-1}^{-1-2\epsilon}$) and in part relative to returns with low time l_k (less than $c_{n-1}^{-1-2\epsilon}$). Using LC4 and LC5 to bound the number of returns with high time, and using LS2 to bound their time, we get

$$h_k < c_{n-1}^{-14}(6\cdot 2^n)c_{n-1}^{100}m_k,$$

and using LC1 and LC3 we have

$$l_k < c_{n-1}^{-1-2\epsilon}(6\cdot 2^n)c_{n-1}^{1/60}m_k < c_{n-1}^{-79/80}m_k,$$

provided ϵ is small enough.

Since $k > c_{n-1}^{-1+2\epsilon}m_k/2$ we have

$$\frac{h_k+l_k}{k} < 4c_{n-1}^{1/85},$$

provided ϵ is small enough.

Now if $i_k < c_{n-1}^{-1-2\epsilon}$ then $i_k/k < c_{n-1}^{1/80}$ (using ϵ small), and if $i_k > c_{n-1}^{-1-2\epsilon}$ then by LC5, $m_k \geq e^{c_{n-1}^{-\epsilon/5}} > c_{n-1}^{-n}$, so by LS2, $i_k/k < i_k/m_k < c_{n-1}^{-14}/c_{n-1}^{-n}$ as well. So in both cases $i_k/k < c_{n-1}^{1/80}$.

From our estimates on i_k and on h_k and l_k we have $1-(\beta_k/k) < c_{n-1}^{1/90}$. Now very good returns are very hyperbolic, and full returns (even not very good ones) always give derivative at least 1 from Lemma 8.10, so we have the estimate

$$a_k > \lambda_{n_0}\frac{(1+2^{n-n_0})}{2}\cdot\frac{\beta_k}{k}-\frac{-q_k}{k}>\frac{\lambda_{n_0}}{2}.$$

□

12. Main Theorems

12.1. **Proof of Theorem A.** We must show that with total probability, f is Collet-Eckmann. We will use the estimates on hyperbolicity of cool landings to show that if the critical point always falls in a cool landing then there is uniform control of the hyperbolicity along the critical orbit.

Let

$$a_k = \frac{\ln |Df^k(f(0))|}{k}$$

and $e_n = a_{v_n - 1}$.

It is easy to see that if n_0 is big enough such that both Lemmas 11.8 and 11.11 are valid, we obtain for n large enough that

$$e_{n+1} \geq e_n \frac{v_n - 1}{v_{n+1} - 1} + \frac{\lambda_{n_0}}{2} \cdot \frac{v_{n+1} - v_n}{v_{n+1} - 1}$$

and so

$$\liminf_{n \to \infty} e_n \geq \frac{\lambda_{n_0}}{2}. \tag{12.1}$$

Let now $v_n - 1 < k < v_{n+1} - 1$. Define $q_k = \ln |Df^{k - v_n}(f^{v_n}(0))|$.

Assume first that $k \leq v_n + c_{n-1}^{-4/(n-1)}$. From LC1 we know that τ_n is very good, so by LS1 we have $r_n(\tau_n) > c_{n-1}^{-1/2}$, so k is in the middle of this branch (that is, $v_n \leq k \leq v_n + r_n(\tau_n) - 1$). Using that $|R_n(0)| > |I_n|/2^n$ (see Lemma 8.8), we get by Lemma 8.11 that

$$-q_k < -\ln(2^{-n} c_{n-1} c_{n-1}^5) < c_{n-2}^{-1-\epsilon}.$$

Since $v_n > c_{n-1}^{-1+\epsilon}$ (by Lemma 10.10) we have

$$a_k \geq e_n \frac{v_n - 1}{k} - \frac{-q_k}{k} > \left(1 - \frac{1}{2^n}\right) e_n - \frac{1}{2^n}. \tag{12.2}$$

If $k > v_n + c_{n-1}^{-4/(n-1)}$, using Lemma 11.11 we get

$$a_k \geq e_n \frac{v_n - 1}{k} + \frac{\lambda_{n_0}}{2} \cdot \frac{k - v_n + 1}{k}. \tag{12.3}$$

It is clear that estimates (12.1), (12.2) and (12.3) imply that $\liminf_{k \to \infty} a_k \geq \lambda_{n_0}/2$ and so f is Collet-Eckmann.

12.2. **Proof of Theorem B.** We must obtain, with total probability, upper and lower (polynomial) bounds for the recurrence of the critical orbit. It will be easier to first study the recurrence with respect to iterates of return branches, and then estimate the total time of those iterates.

12.2.1. *Recurrence in terms of return branches.* The principle is very simple: for the essentially Markov process generated by iteration of the non-central branches of R_n, almost every point (in the qs sense) has polynomial recurrence before falling in I_{n+1}.

Lemma 12.1. *With total probability, for n big enough and for $1 \leq i \leq c_{n-1}^{-2}$,*

$$\frac{\ln |R_n^i(0)|}{\ln(c_{n-1})} < (1 + 4\epsilon) \left(1 + \frac{\ln(i)}{\ln(c_{n-1}^{-1})}\right).$$

Proof. Notice that due to torrential (and monotonic) decay of c_n, we can estimate $|I_n| = c_{n-1}^{1+\delta_n}$, with δ_n decaying torrentially fast.

From Lemma 8.8, we have

$$\frac{\ln |R_n(0)|}{\ln c_{n-1}} < \frac{\ln(2^{-n}|I_n|)}{\ln c_{n-1}} < 1 + 4\epsilon$$

and the result follows for $i = 1$.

For $1 \leq j \leq 2\epsilon^{-1}$, let $X_j \subset I_n$ be a $c_{n-1}^{(1+2\epsilon)(1+j\epsilon)}$ neighborhood of 0. For n big, we can estimate (due to the relation between $|I_n|$ and c_{n-1})

$$\frac{|X_j|}{|I_n|} < \frac{c_{n-1}^{(1+2\epsilon)(1+j\epsilon)}}{c_{n-1}^{1+2\epsilon}} = c_{n-1}^{j\epsilon(1+2\epsilon)}$$

(we of course consider X_j a union of $C_n^{\underline{d}}$, so that its size is near the required size, the precision is high enough for our purposes due to Remark 8.3).

We have to make sure that the critical point does not land in some X_j for $c_{n-1}^{(1-j)\epsilon} < i \leq c_{n-1}^{-j\epsilon}$. This requirement can be translated on $R_n(0)$ not belonging to a certain set $Y_j \subset I_n$ such that

$$Y_j = \bigcup_{c_{n-1}^{(1-j)\epsilon} \leq |\underline{d}| < c_{n-1}^{-j\epsilon}} (R_n^{\underline{d}})^{-1}(X_j).$$

It is clear that

$$p_\gamma(I_n^{\tau_n} \cap Y_j | I_n^{\tau_n}) \leq c_{n-1}^{-j\epsilon} c_{n-1}^{(1+\epsilon)j\epsilon} < c_{n-1}^{\epsilon^2}$$

and

$$p_\gamma(I_n^{\tau_n} \cap \bigcup_{j=1}^{2\epsilon^{-1}} Y_j | I_n^{\tau_n}) < 2\epsilon^{-1} c_{n-1}^{\epsilon^2}.$$

Applying PhPa1, the probability that for some $1 \leq j \leq 2\epsilon^{-1}$ and $c_{n-1}^{(1-j)\epsilon} < i \leq c_{n-1}^{-j\epsilon}$ we have $|R_n^i(0)| < c_{n-1}^{(1+2\epsilon)(1+j\epsilon)}$ is at most $2\epsilon^{-1} c_{n-1}^{\epsilon^2}$, which is summable. In particular, with total probability, for j and i as above, we have

$$\begin{aligned}\frac{\ln |R_n^i(0)|}{\ln(c_{n-1})} &< (1+2\epsilon)(1+j\epsilon) \\ &< (1+4\epsilon)(1+(j-1)\epsilon) < (1+4\epsilon)\left(1 + \frac{\ln(i)}{\ln(c_{n-1}^{-1})}\right).\end{aligned}$$

∎

Lemma 12.2. *With total probability, for n big enough and for $c_{n-1}^{-1-\epsilon} < i \leq s_n$,*

$$\frac{\ln |R_n^i(0)|}{\ln(c_{n-1})} < (1+4\epsilon)\left(1 + \frac{\ln(i)}{\ln(c_{n-1}^{-1})}\right).$$

Proof. The argument is the same as for the previous lemma, but the decomposition has a slightly different geometry. Let

$$x_j = c_{n-1}^{(1+2\epsilon)(1+(1+\epsilon)^{j+1})},$$

so that

$$\frac{x_j}{|I_n|} < \frac{c_{n-1}^{(1+2\epsilon)(1+(1+\epsilon)^{j+1})}}{c_{n-1}^{1+\epsilon}} < c_{n-1}^{(1+2\epsilon)(1+\epsilon)^{j+1}}.$$

Let K be biggest with $x_K > c_n^{1-\epsilon}$. For $0 \le j \le K$, let $X_j \subset I_n$ be a x_j neighborhood of 0 (approximated as union of $C_n^{\underline{d}}$, notice that $x_j > c_n^{1-\epsilon} \gg |I_{n+1}|$ for $0 \le j \le K$, so the approximation is good enough for our purposes due to Remark 8.3). Let $Y_j \subset I_n$ be such that

$$Y_j = \bigcup_{c_{n-1}^{-(1+\epsilon)^j} \le |\underline{d}| < c_{n-1}^{-(1+\epsilon)^{j+1}}} (R_n^{\underline{d}})^{-1}(X_j).$$

It is clear that

$$p_\gamma(I_n^{\tau_n} \cap Y_j | I_n^{\tau_n}) \le c_{n-1}^{-(1+\epsilon)^{j+1}} c_{n-1}^{(1+\epsilon)^{j+2}} < c_{n-1}^{\epsilon(1+j\epsilon)}$$

and

$$p_\gamma(I_n^{\tau_n} \cap \bigcup_{j=0}^{K} Y_j | I_n^{\tau_n}) < \sum_{j=0}^{\infty} c_{n-1}^{\epsilon(1+j\epsilon)} = \frac{c_{n-1}^{\epsilon}}{1 - c_{n-1}^{\epsilon^2}} < c_{n-1}^{\epsilon/2}.$$

Applying PhPa1, the probability that for some $0 \le j \le K$ and

$$c_{n-1}^{-(1+\epsilon)^j} < i \le c^{-(1+\epsilon)^{j+1}}$$

we have

$$|R_n^i(0)| < c_{n-1}^{(1+2\epsilon)(1+(1+\epsilon)^{j+1})}$$

is at most $c_{n-1}^{\epsilon/2}$, which is summable. In particular, with total probability, for j and i as above, we have

$$\begin{aligned}\frac{\ln |R_n^i(0)|}{\ln(c_{n-1})} &< (1+2\epsilon)(1+(1+\epsilon)^{j+1}) \\ &< (1+4\epsilon)(1+(1+\epsilon)^j) < (1+4\epsilon)\left(1 + \frac{\ln(i)}{\ln(c_{n-1}^{-1})}\right).\end{aligned}$$

This covers the range $c_{n-1}^{-1} < i \le c_{n-1}^{-(1+\epsilon)^{K+1}}$. For $c_{n-1}^{-(1+\epsilon)^{K+1}} < i \le s_n$, notice that $R_n^i(0) \notin I_{n+1}$, so

$$\begin{aligned}\frac{\ln |R_n^i(0)|}{\ln c_{n-1}} &< \frac{\ln(|I_{n+1}|/2)}{\ln c_{n-1}} \\ &< \frac{1+4\epsilon}{1+2\epsilon} \cdot \frac{\ln c_n^{1-\epsilon}}{\ln c_{n-1}} \\ &\le \frac{1+4\epsilon}{1+2\epsilon} \cdot \frac{\ln x_{K+1}}{\ln c_{n-1}} && \text{(by definition of } K\text{)} \\ &\le (1+4\epsilon)(1+(1+\epsilon)^{K+1}) \\ &\le (1+4\epsilon)\left(1 + \frac{\ln(i)}{\ln(c_{n-1}^{-1})}\right).\end{aligned}$$

□

Both cases are summarized below:

Corollary 12.3. *With total probability, for n big enough and for $1 \le i \le s_n$,*

$$\frac{\ln |R_n^i(0)|}{\ln(c_{n-1})} < (1+4\epsilon)\left(1 + \frac{\ln(i)}{\ln(c_{n-1}^{-1})}\right).$$

12.2.2. *Total time of full returns.* We must now relate the return times in terms of R_n to the return times in terms of f.

For $1 \leq i \leq s_n$, let k_i be such that $R_n^i(0) = f^{k_i}(0)$.

Lemma 12.4. *With total probability, for n big enough and for $c_{n-1}^{-\epsilon} < i < s_n$,*

$$\frac{\ln(k_i)}{\ln(c_{n-1}^{-1})} > (1-3\epsilon)\left(1+\frac{\ln(i)}{\ln(c_{n-1}^{-1})}\right).$$

Proof. By Lemma 11.8, $R_n(0)$ belongs to a cool landing, so using LC2 (which allows to estimate the average of return times over a large initial segment of cool landings) we get

$$\frac{k_i}{i-1} > c_{n-1}^{-1+3\epsilon}.$$

This immediately gives

$$\frac{\ln(k_i)}{\ln(c_{n-1}^{-1})} > (1-3\epsilon)+\frac{\ln(i-1)}{\ln(c_{n-1}^{-1})} > (1-3\epsilon)\left(1+\frac{\ln(i)}{\ln(c_{n-1}^{-1})}\right)$$

□

Using that $v_n > c_{n-1}^{-1+\epsilon}$ (from Corollary 10.10) and that $k_i \geq v_n$ we get for $1 \leq i \leq c_{n-1}^{-\epsilon}$

$$\frac{\ln(k_i)}{\ln(c_{n-1}^{-1})} > \frac{\ln(v_n)}{\ln(c_{n-1}^{-1})} > \frac{\ln(c_{n-1}^{-1+\epsilon})}{\ln(c_{n-1}^{-1})} > (1-3\epsilon)(1+\epsilon) \geq (1-3\epsilon)\left(1+\frac{\ln(i)}{\ln(c_{n-1}^{-1})}\right).$$

Together with Lemma 12.4, this gives

Corollary 12.5. *With total probability, for n big enough and for $1 \leq i \leq s_n$,*

$$\frac{\ln(k_i)}{\ln(c_{n-1}^{-1})} > (1-3\epsilon)\left(1+\frac{\ln(i)}{\ln(c_{n-1}^{-1})}\right).$$

12.2.3. *Upper and lower bounds.* Notice that $|R_n(0)| = |f^{v_n}(0)| < c_{n-1}$, so using Lemma 10.10 we get

$$\limsup_{n\to\infty} \frac{-\ln|f^n(0)|}{\ln(n)} \geq \limsup_{n\to\infty} \frac{-\ln|f^{v_n}(0)|}{\ln(v_n)} \geq \limsup_{n\to\infty} \frac{-\ln(c_{n-1})}{\ln(v_n)} \geq 1.$$

Let now $v_n \leq k < v_{n+1}$. If $|f^k(0)| < k^{-1-10\epsilon}$ then Lemma 10.10 implies that $f^k(0) \in I_n$ and so $k = k_i$ for some i. It follows from Corollaries 12.3 and 12.5 that

$$|f^{k_i}(0)| > k_i^{-1-10\epsilon}.$$

Varying ϵ we get

$$\limsup_{n\to\infty} \frac{-\ln|f^n(0)|}{\ln(n)} \leq 1.$$

References

[A] A. Avila. Bifurcations of unimodal maps: the topological and metric picture. Thesis IMPA (2001) (www.math.sunysb.edu/∼artur).

[ALM] A. Avila, M. Lyubich, W. de Melo. Regular or stochastic dynamics in real analytic families of unimodal maps. Preprint (www.math.sunysb.edu/∼artur). Submitted for publication.

[AM] A. Avila, C. G. Moreira. Statistical properties of unimodal maps: smooth families with negative Schwarzian derivative. Preprint (www.arXiv.org). Submitted for publication.

[BBM] V. Baladi, M. Benedicks, V. Maume. Almost sure rates of mixing for i.i.d. unimodal maps. Preprint (1999), to appear Ann. E.N.S.

[BV] V. Baladi, M. Viana. Strong stochastic stability and rate of mixing for unimodal maps. Ann. scient. Éc. Norm. Sup., v. 29 (1996), 483-517.

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

[BC2] M. Benedicks, L. Carleson. On dynamics of the Hénon map. Ann. Math., v. 133 (1991), 73-169.

[BR] L. Bers & H.L. Royden. Holomorphic families of injections. Acta Math., v. 157 (1986), 259-286.

[D] A. Douady. Prolongement de mouvements holomorphes (d'apres Slodkowski et autres). Asterisque, v. 227, 7-20.

[GS] J. Graczyk, G. Swiatek. Generic hyperbolicity in the logistic family. Ann. of Math., v. 146 (1997), 1-52.

[HK] F. Hofbauer, G. Keller. Quadratic maps without asymptotic measure. Comm. Math. Physics, v. 127 (1990), 319-337.

[J] M. Jacobson. Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. Comm. Math. Phys., v. 81 (1981), 39-88.

[Jo] S. D. Johnson. Singular measures without restrictive intervals. Comm. Math. Phys., 110 (1987), 185-190.

[KN] G. Keller, T. Nowicki. Spectral theory, zeta functions and the distribution of periodic points for Collet-Eckmann maps. Comm. Math. Phys., 149 (1992), 31-69.

[L1] M. Lyubich. Combinatorics, geometry and attractors of quasi-quadratic maps. Ann. Math, **140** (1994), 347-404.

[L2] M. Lyubich. Dynamics of quadratic polynomials, I-II. Acta Math., **178** (1997), 185-297.

[L3] M. Lyubich. Dynamics of quadratic polynomials, III. Parapuzzle and SBR measure. Asterisque, v. 261 (2000), 173 - 200.

[L4] M. Lyubich. Feigenbaum-Coullet-Tresser universality and Milnor's hairiness conjecture. Ann. of Math. (2) 149 (1999), no. 2, 319–420.

[L5] M. Lyubich. Almost every real quadratic map is either regular or stochastic. Preprint IMS at Stony Brook, # 1997/8. To appear in Ann. Math.

[MN] M. Martens, T. Nowicki. Invariant measures for Lebesgue typical quadratic maps. Asterisque, v. 261 (2000), 239 - 252.

[MSS] R. Mañé, P. Sad & D. Sullivan. On the dynamics of rational maps, Ann. scient. Ec. Norm. Sup., **16** (1983), 193-217.

[MvS] W. de Melo, S. van Strien. One-dimensional dynamics. Springer, 1993.

[NS] T. Nowicki, D. Sands. Non-uniform hyperbolicity and universal bounds for S-unimodal maps. Invent. Math. 132 (1998), no. 3, 633–680.

[Pa] J. Palis. A global view of dynamics and a Conjecture of the denseness of finitude of attractors. Asterisque, v. 261 (2000), 335 - 348.

[Sl] Z. Slodkowsky. Holomorphic motions and polynomial hulls. Proc. Amer. Math. Soc., **111** (1991), 347-355.

[Y] L.-S. Young. Decay of correlations for certain quadratic maps. Comm. Math. Phys., 146 (1992), 123-138.

Collège de France – 3 Rue d'Ulm, 75005 Paris – France.
E-mail address: avila@impa.br

IMPA – Estr. D. Castorina 110, 22460-320 Rio de Janeiro – Brazil.
E-mail address: gugu@impa.br