STATISTICAL PROPERTIES OF UNIMODAL MAPS: PHYSICAL MEASURES, PERIODIC ORBITS AND PATHOLOGICAL LAMINATIONS

ARTUR AVILA AND CARLOS GUSTAVO MOREIRA

Abstract. In this work, we relate the geometry of chaotic attractors of typical analytic unimodal maps to the behavior of the critical orbit. Our main result is an explicit formula relating the combinatorics of the critical orbit with the exponents of periodic orbits. This connection between topological and smooth invariants is obtained through an analysis of the physical measure. Since the exponents of periodic orbits form a complete set of smooth invariants in this setting, we have "typical geometric rigidity" of the dynamics of such chaotic attractors. This unexpected result implies that the lamination structure of spaces of analytic maps (obtained by the partition into topological conjugacy classes, see [ALM]) has an absolutely singular nature.

Contents

1. Introduction	1
2. Preliminaries	7
3. The formula	10
4. Phase-parameter relation and statistics of the principal nest	12
5. The critical orbit is typical	16
6. Regularity of the physical measure	31
7. Proof of the Main Estimate	37
8. Pathological laminations and other consequences	46
References	49

1. Introduction

A *unimodal map* is a smooth (at least C^2) map $f : I \to I$, where $I \subset \mathbb{R}$ is an interval, which has a unique critical point $c \in \operatorname{int} I$ which is a maximum. A unimodal map f is said to be *regular* if it is hyperbolic and if its critical point is non-degenerate and is not periodic or preperiodic. This definition is such that the set of regular maps coincide with the set of unimodal maps which are structurally stable, see [K2] Theorem B. The class of regular maps is open in the C^2 topology and dense in any smooth, and even analytic, topology.

The main examples of unimodal maps are quadratic maps $p_a(x) = a - x^2$, $-1/4 \leq a \leq 2$. Behind their innocent definition, the dynamics of quadratic maps reveals an intricate structure and has been subject of intense research in the past few decades.

Recently, several works have concentrated on investigating the dynamics of typical unimodal maps. The most natural notion of typical in this context is measure-theoretical: a dynamical property is said to be typical *in the quadratic family* if it is satisfied by p_a for Lebesgue almost every parameter

Date: June 10, 2003.

Partially supported by Faperj and CNPq, Brazil.

a. This notion easily extends to the (infinite-dimensional) setting of general unimodal maps: a property is typical if it corresponds to a full measure set of parameters in an *ample class of families* of unimodal maps[1].

The dynamics of regular maps is quite well understood. Due to the works of Jakobson and Benedicks-Carleson, non-regular unimodal maps correspond to a positive measure set of parameters in a large (C^2 open) set of parametrized families. In the works [L4], [AM1], the dynamics of typical non-regular quadratic maps was described in great detail from the *statistical point of view*. Those results were subsequently extended to typical analytic (and even smooth) unimodal maps in [ALM], [AM2] (in the quasiquadratic[2] case), and finally in all generality in [AM3]: a typical non-regular unimodal map f possess a unique *non-trivial chaotic attractor* A_f, which is a transitive finite union of intervals (where periodic orbits are dense). Moreover, this attractor is the support of an absolutely continuous invariant measure μ_f, with excellent stochastic properties (due, notably, to the Collet-Eckmann condition). The measure-theoretical dynamics of f can be described by μ_f and finitely many *trivial attractors* (hyperbolic periodic orbits). The attractor A_f of f can be defined also on topological grounds: it is simultaneously a metric and topological attractor in the sense of Milnor (see [L1]).

Our aim in this paper is to establish much finer geometric properties of the non-trivial attractor of a typical non-regular analytic unimodal map f. Roughly speaking, we will show how topological invariants of f (coded using the theory of Milnor-Thurston) can be used to determine (and actually compute) a complete set of smooth invariants of A_f.

In the proof of this connection between topological and smooth invariants, the physical measure μ_f will play an important role. One of our most important steps is to show how the information contained in the physical measure is enough to compute some geometric invariants of hyperbolic Cantor sets.

Our main theorem can be seen as a proof of "geometric rigidity" in the typical setting, which is rather unexpected and even looks paradoxical at first. Fortunately, it is possible to visualize this consequence using the results of [ALM]. The resulting rather amusing picture is related to some recently discovered examples of measure-theoretical pathological laminations (Katok's "Fubini Foiled" phenomena presented by Milnor [Mi], and the examples in [SW] and [RW]).

1.1. **Statement of the results.** In this work, the ample set of families we will consider for the definition of typical is very explicit: the set of non-trivial analytic families of unimodal maps, that is, families which contain a dense set of regular parameters. The set of non-trivial families is very large (its complement has infinite codimension). Moreover, among families of quasiquadratic maps (a C^3 open condition) it is much easier to check for non-triviality: it is enough to show existence of one regular parameter (which is a C^2 open condition). In particular, analytic families C^3 close to the quadratic family are non-trivial.

1.1.1. *Typical unimodal maps and their invariants: relation between topological and combinatorial invariants.* To each point $x \in I$, let us associate an infinite sequence (the *itinerary*) of 0s and 1s as follows. The k-th element is 0 if $f^k(x)$ is to the left of the critical point, and 1 otherwise. Itineraries are clearly invariant under topological conjugacy. The itinerary of the critical point of f is called the *kneading sequence* of f, and it is a particularly important invariant: the work of Milnor-Thurston shows that the kneading sequence determines the set of itineraries of all points $x \in I$.

The kneading sequence is actually an "essentially" complete topological invariant in the sense that it determines the topological conjugacy class up to some well understood obstructions corresponding

[1]This notion of typical is inspired by Kolmogorov.

[2]A C^3 unimodal map is said to be quasiquadratic if any C^3 perturbation is conjugate to a quadratic map.

to trivial dynamics. A simpler (and perhaps more basic, as it applies in all dimensions) example of topological invariant is the set of periodic orbits of the system, together with their periods. If p is a periodic point, its itinerary is clearly periodic.

To a periodic orbit p of period n we can associate its exponent $Df^n(p)$. This quantity is easily checked to be invariant by a diffeomorphic change of coordinates, thus providing the simplest example of a smooth invariant. By the work of Liv̌sic [Li], see also Shub-Sullivan [ShSu], in some circumstances (say, expanding maps of the circle) exponents of periodic orbits form a complete set of smooth invariants, in the sense that a topological conjugacy which preserves exponents is necessarily smooth. In the unimodal case, the same result holds due to the work of Martens-de Melo [MM], at least for the cases that appear in our considerations (non-trivial attractor of a typical non-regular unimodal map).

The main result of this paper relates the above combinatorial and smooth invariants for typical non-regular analytic unimodal maps.

Theorem A. *Let f_λ be a non-trivial analytic family of unimodal maps. Then, for almost every non-regular parameter λ, and for every periodic orbit p in the non-trivial attractor A_{f_λ}, the exponent of p is determined by an explicit combinatorial formula involving the kneading sequence of f and the itinerary of p.*

The formula goes as follows: let β be the kneading sequence of f and let α be the periodic part of the itinerary of a periodic point p in A_f. Let us consider the asymptotic frequency $r(\alpha^k,\beta)$ of α^k (k repetitions of α) inside β. Ignoring for a moment the problem of existence of this asymptotic frequency (which is part of Theorem B below), we obtain a non-increasing sequence of numbers between 0 and 1. It turns out that this sequence decreases to 0 geometrically at some precise rate (this is related to Theorem C below). The inverse of this rate is the absolute value of the exponent of p (the sign being given by $(-1)^s$ where s is the number of 1s in α).

1.1.2. *The critical orbit is typical.* Let us say that the asymptotic distribution of a point x is given by μ (or equivalently, x is in the basin of μ, or x is typical for μ) if μ is a probability measure and for any continuous function $\phi : I \to \mathbb{R}$

$$\lim \frac{1}{n}\sum_{k=0}^{n-1}\phi(f^k(x)) = \int \phi d\mu. \tag{1.1}$$

One important step of the proof of Theorem A is to analyze the asymptotic distribution of the critical orbit. The existence of an asymptotic limit for the distribution of the critical orbit is directly related to the existence of asymptotic frequencies $r(\alpha,\beta)$ of an arbitrary finite sequence α inside the kneading sequence β of f.

Theorem B. *Let f_λ be a non-trivial analytic family of quasiquadratic maps. Then, for almost every non-regular parameter λ, the critical point belongs to the basin of μ_{f_λ} (the absolutely continuous invariant measure of f_λ).*

In other words, for typical non-regular unimodal maps, the critical orbit is typical for the "correct" measure of the system. We are thus able to obtain the following consequence:

Corollary 1.1. *In the setting of Theorem B, one also has equality between the Lyapunov exponent of the critical value and the Lyapunov exponent of μ_{f_λ}.*

Recall that the Lyapunov exponent of a point x is defined as

$$\lambda(x) = \lim \frac{\ln|Df^n(x)|}{n} \tag{1.2}$$

provided the limit exists. The Lyapunov exponent of μ_f is given by the formula

$$\lambda(\mu_f) = \int \ln |Df| d\mu_f. \tag{1.3}$$

Some work is needed to go from Theorem B to Corollary 1.1, since $\ln |Df|$ is not continuous.

Previous progress in the direction of Theorem B was achieved (with very different techniques) by Benedicks-Carleson [BC], who proved typicality of the critical orbit for a *positive measure set* of parameters for the quadratic family.

1.1.3. *Regularity of the physical measure and hyperbolic sets.* In Theorem A we are interested in the exponents of (repelling) periodic orbits. More generally, one is led to ask about the geometry of hyperbolic subsets $K \subset I$ (say, a Cantor set).

In order to apply Theorem B to reconstruct the geometry of K from the kneading sequence of f, one is led to ask: is it possible to obtain sharp estimates for the asymptotic geometry of K from knowledge of the physical measure?

In order to do so, one should be able to relate asymptotically the physical measure of gaps (and unions of gaps) of K and their (Lebesgue) size. Thus, behind this problem is the issue of regularity of the physical measure μ_f.

It turns out that this problem is non-trivial: indeed, if one tries to estimate general intervals, and not just gaps of hyperbolic sets, one would get quite negative results. For instance, let us take f to be a quadratic map and let T be an interval of radius ϵ around the critical point. Then $\mu_f(T) = \mu_f(f(T))$, but $|T|$ is of order ϵ while $|f(T)|$ is of order ϵ^2. Thus, for general intervals, estimates of the physical measure might lead to errors of order 2 (when taking logarithms) on estimates of Lebesgue measure (and thus on the formula for exponents of periodic orbits). Connected to this fact is the following limitation on the regularity of μ_f: its density $d\mu_f$ is never in L^2.

So one is led to regularize the density $d\mu_f$ using the Cantor set K (or view $d\mu_f$ through K). Let us denote $d\mu_f^K$ the function which is constant in each gap T of K and takes the average value of $d\mu_f$ on T.

In other words, $d\mu_f^K$ is the expectation of $d\mu_f$ with respect to the sigma-algebra $\mathcal{B}(K)$ of the gaps of K. The sigma algebra $\mathcal{B}(K)$ gives us enough information to compute the exponent of periodic orbits p in A_f if, say, K is a Cantor set containing p (any periodic point $p \in A_f$ can be included in such a Cantor set).

Theorem C. *Let f_λ be a non-trivial analytic family of unimodal maps. For almost every non-regular parameter λ and any hyperbolic set $K \subset I$, we have $d\mu_{f_\lambda}^K \in L^p$, $1 \leq p < \infty$.*

One can see this estimate (together with Theorem B) as a generalization of Theorem A, since it allows to compute using μ_f (which, due to Theorem B can be computed combinatorially), fine asymptotics of general hyperbolic sets (of which periodic orbits are an example).

We should point out that the lack of regularity of μ_f comes from the critical point, and essentially distributes itself along the orbit of the critical value. In order to show that μ_f behaves well with respect to hyperbolic sets, we must show roughly that "the critical orbit distributes transversely with respect to K".

1.1.4. *Geometric rigidity, pathological laminations.* The main motivation for Theorem A is, as described before, the possibility to compute, from topological information, a complete set of smooth invariants. This may seem at first paradoxical, since exponents of periodic orbits *can* be varied without changing the topological class, and are thus responsible for moduli of flexibility, as opposed to rigidity (examples of geometrically rigid systems, as Diophantine irrational rotations usually do not have periodic orbits).

In order to visualize what is really happening, we must consider the partition of the space of unimodal maps into topological conjugacy classes. The results of [ALM] show that, in appropriate Banach spaces of analytic unimodal maps, the set of non-regular topological classes form a lamination with analytic leaves and quasisymmetric holonomy, at least almost everywhere[3].

For each topological class of unimodal maps, the formula for exponents of periodic orbits determines *at most one* "preferred" smooth structure on the non-trivial attractor[4]. In each non-regular topological class (of codimension one by [ALM]), the set of maps with the "correct" smooth structure is a tiny set (of infinite codimension, the parameters being precisely the exponents of periodic orbits, or even empty). However, the set of typical non-regular unimodal maps (satisfying the conclusion of Theorem A) intersects each topological class precisely at such a tiny set.

So "typical rigidity" has interesting consequences for the regularity of the lamination by topological classes: *the stratification of the set of typical non-regular analytic unimodal maps by topological classes is highly non-homogeneous, in the sense that it fails drastically to be absolutely continuous.* Indeed, that the lamination can not be absolutely continuous is easily checked since the phenomena we described imply the complete failure of Fubini's Theorem. (Although the setting is infinite dimensional, one can interpret those results in parametrized families with at least two parameters.)

1.1.5. *On universality and the holonomy method.* The results of [ALM] imply that the parameter space of the quadratic family do have a universal quasisymmetric structure (due to the holonomy of the lamination). Although quasisymmetric maps are not necessarily absolutely continuous, the metric universality was used in [ALM] and [AM2] to transfer certain strong measure-theoretical results (regular or stochastic dichotomy, Collet-Eckmann condition and polynomial recurrence of the critical orbit) from the quadratic family to other analytic families of (quasiquadratic) unimodal maps.

This so called *holonomy method*, consisting in the comparison between parameter spaces of different families had to be applied to estimates which are topological invariants. More seriously, the set of combinatorics concerned must have full measure *simultaneously* in all non-trivial families of unimodal maps.

The lack of absolute continuity of the lamination established now sets a limit to the metric universality of the parameter space of unimodal families (as the quadratic family). Our Theorem A is particularly interesting in this respect since it gives an example of a result which is definitely inaccessible by the holonomy method (which clearly can not be used to prove that the lamination itself is not absolutely continuous).

1.1.6. *Related matters.* Another consequence of our techniques is existence of a combinatorial formula for the Lyapunov exponent of typical non-regular unimodal maps. This exponent coincides with the one of the critical value by Corollary 1.1. This formula is quite simple, but is formulated in terms of the principal nest description of the combinatorics instead of itineraries, so we postpone its formulation to §8.2.

In view of Theorem A, it is natural to ask how to effectively relate the information about the exponents of periodic orbits to other properties of interest of a typical non-regular unimodal map.

[3]Almost everywhere here is indeed stronger than our notion of typical. More precisely, the set of non-regular topological classes has a lamination structure in an open set containing all Kupka-Smale maps (unimodal maps with a non-degenerate critical point and without non-hyperbolic periodic orbits). The complement of this open set is clearly contained in a countably union of codimension-one analytic varieties.

[4]For a general topological class, several things might go wrong, so that no smooth structure is determined. At the level of the formula, for instance, its defining limits might not exist. The non-trivial attractor may not exist. Even if both exist, the values for exponents thus obtained might not correspond to any smooth structure on the non-trivial attractor.

Although we will not investigate this problem in this paper, we would like to call attention to one situation where such a relation might be explicitly obtained.

It is common to organize periodic orbits in a *zeta function*. The general formula for a zeta function is

$$\zeta_\phi(z) = \exp\left(\sum_{n=1}^{\infty} \frac{z^n}{n} \sum_{p \in \mathrm{Fix}(f^n)} \prod_{k=0}^{n-1} \phi(f^k(p))\right) \tag{1.4}$$

where $\mathrm{Fix}(f^n)$ is the set of fixed *points* of f^n and ϕ is a weight function which is to be chosen according to the problem to be studied.

The relation of zeta functions and the thermodynamical formalism of hyperbolic dynamical systems is well developed. However it is reasonable to expect that this relation might also hold for certain non-uniformly hyperbolic unimodal maps, and in [KN] some results in this direction were obtained in the Collet-Eckmann case.

For the weight $\phi = |Df|^{-1}$, the zeta function can be written as

$$\zeta_{|Df|^{-1}}(z) = \exp\left(\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{z^{mn}}{m} \sum_{p \in \mathrm{Per}_n(f)} \frac{1}{|Df^n(p)|^m}\right) \tag{1.5}$$

where $\mathrm{Per}_n(f)$ is the set of periodic *orbits* of (prime) period n. Notice that in this case the zeta function only depends on the exponent of periodic orbits, so by Theorem A it can be expressed combinatorially for typical non-regular maps. This choice of the weight is particularly interesting: it is related to the physical measure μ_f, and the results of [KN] show that the poles of $\zeta_{|Df|^{-1}}$ can be sometimes related to parts of the spectrum of the Ruelle transfer operator, which encodes (in some cases precise) information about the rates of decay of correlations of the system (for certain classes of observables). It is a natural problem to show that the first pole of $\zeta_{|Df|^{-1}}$ gives indeed the exact rate of decay of correlations (for smooth enough observables) of typical non-renormalizable unimodal maps.

1.2. **Complex techniques.** The successful investigation of families of unimodal maps, specially the quadratic family, was heavily tied to the possibility of the intertwined use of real and complex techniques. Our results are based on the coupling of two main methods. For the analysis of the dynamics in phase space, we use a statistical description of the critical orbit. Techniques from complex dynamics are used to obtain the Phase-Parameter relation, which allows to compare the phase space and the parameter space of a non-trivial family. Those complex techniques are mainly based in the theory of Lyubich (which in turn uses ideas from several different fields).

The Phase-Parameter relation was proved in [AM1] in the case of the quadratic family, and in [AM3] in all generality. This last result can be directly used in our context and will allow us to concentrate mostly on the real dynamics of unimodal maps.

Remark 1.1. Although, as discussed in 1.1.5, the holonomy method of [ALM] can not be used for this work, the lamination structure of the partition into topological classes is still the key result from complex dynamics used in the proof of the Phase-Parameter relation in [AM3]. This is possible because the regularity of the holonomy map between two transversals is related to their distance. The original holonomy method has a global nature and corresponds to relating phase and parameter between a given non-trivial family and the quadratic family, introducing serious distortion and lack of sharpness in the estimates. To get over those limitations, one uses local holonomy maps (which are more regular) to relate phase and parameter of the same family.

1.3. **Outline.** In §2, we present some background on the dynamics of unimodal maps. In §3, we state precisely the formula for periodic orbits. We then prove Theorem A, assuming the validity of Theorems B and C.

In §4, we discuss the combinatorics of the principal nest and introduce our basic tool to make parameter estimates: the Phase-Parameter relation, which was proved in [AM1] (for the quadratic family) and in [AM3] in all generality. We then present some of the estimates obtained in [AM1].

In §5, we prove Theorem B. The proof is technical but has a clear strategy, which we describe in §5.1. In §6, we reduce Theorem C to the so called Main estimate, which we prove in §7. The proof of the Main estimate is the most technically involved part of this work.

2. Preliminaries

2.1. **Notation.** As usual, $\mathbb{N} = \{0, 1, 2, \dots\}$ stands for the set of natural numbers; $\mathbb{R}$ stands for the real line; $\mathbb{C}$ stands for the complex plane.

The Lebesgue measure of a set $X \subset \mathbb{R}$ will be denoted by $|X|$.

Given a diffeomorphism $\phi : J \to J'$ between two real intervals, its *distortion* or *non-linearity* is defined as

$$\operatorname{dist}(\phi) = \sup_{x,y \in J} \frac{|D\phi(x)|}{|D\phi(y)|}. \tag{2.1}$$

Its *Schwarzian derivative* is given by the formula:

$$S\phi = \frac{D^3\phi}{D\phi} - \frac{3}{2}\left(\frac{D^2\phi}{D\phi}\right)^2. \tag{2.2}$$

The condition of negative Schwarzian derivative plays an important role in one-dimensional dynamics. This condition is preserved under composition.

2.2. **Quasisymmetric maps.** A quasisymmetric map is a homeomorphism $h : \mathbb{R} \to \mathbb{R}$ such that there exists a constant k such that for any $x \in \mathbb{R}$, $a > 0$,

$$\frac{1}{k} < \frac{h(x+a) - h(x)}{h(x) - h(x-a)} < k. \tag{2.3}$$

Equivalently, h is quasisymmetric if it has a symmetric quasiconformal extension to the whole $\mathbb{C}$ (Ahlfors-Beurling). We say that h is γ-qs if there exists such an extension with dilatation bounded by γ. The quasisymmetric constant of a quasisymmetric map h is the infimum of the dilatations of all those extensions[5]. In particular, if h_1 is γ_1-qs and h_2 is γ_2-qs, $h_2 \circ h_1$ is $\gamma_1\gamma_2$-qs.

If $h : X \to \mathbb{R}$ is a monotonic map defined on $X \subset \mathbb{R}$, we will also say that h is γ-qs if it has a γ-qs extension to $\mathbb{R}$.

One of the main concepts we will need in our paper was introduced in [AM1]. The γ-qs capacity of a set $X \subset \mathbb{R}$ inside some interval $T \subset \mathbb{R}$ is defined as

$$p_\gamma(X|T) = \sup \frac{|h(X \cap T)|}{h(T)} \tag{2.4}$$

where $h : \mathbb{R} \to \mathbb{R}$ ranges over all γ-qs maps. An important property of γ-qs capacity is its behavior under tree decomposition: if $T^j \subset T$ are disjoint intervals and $X \subset \cup T_j$ then

$$p_\gamma(X|T) \le p_\gamma(\cup T^j|T) \sup_j p_\gamma(X|T^j). \tag{2.5}$$

[5]It is possible to work out upper bounds for the quasisymmetric constant in terms of the k in (2.3) and inversely.

2.3. **Unimodal maps.** We refer to the book of de Melo and van Strien [MS] for the general background in one-dimensional dynamics.

We will say that a smooth (at least C^2) map $f : I \to I$ of the interval $I = [-1, 1]$ is *unimodal* if $f(-1) = -1$, $f(x) = f(-x)$ and 0 is the only critical point of f and is non-degenerate, so that $D^2 f(0) \neq 0$. The introduction of normalization and symmetry in this definition is exclusively for the simplicity of the notation, and is no loss of generality, see also Appendix C of [ALM]. The assumption of non-degeneracy of the critical point is clearly typical.

Basic examples of unimodal maps are given by quadratic maps

$$q_\tau : I \to I, \quad q_\tau(x) = \tau - 1 - \tau x^2, \tag{2.6}$$

where $\tau \in [1/2, 2]$ is a real parameter.

Let $\mathbb{U}^k$, $k \geq 2$ be the space of C^k unimodal maps. We endow $\mathbb{U}^k$ with the C^k topology. A map $f \in \mathbb{U}^3$ is *quasiquadratic* if any nearby map $g \in \mathbb{U}^3$ is topologically conjugate to some quadratic map. We denote by $\mathbb{U} \subset \mathbb{U}^3$ the space of quasiquadratic maps. By the theory of Milnor-Thurston and Guckenheimer [MS], a map $f \in \mathbb{U}^3$ with negative Schwarzian derivative and $Df(-1) > 1$ is quasiquadratic, so quadratic maps q_τ, $\tau \in (1/2, 2]$ belong to $\mathbb{U}$.

A map $f \in \mathbb{U}^2$ is said to be *Kupka-Smale* if all periodic orbits are hyperbolic. It is said to be hyperbolic if it is Kupka-Smale and the critical point is attracted to a periodic attractor. It is said to be *regular* if it is hyperbolic and its critical point is not periodic or preperiodic. It is well known that regular maps are structurally stable.

In this paper, an analytic family of unimodal maps will be understood as a one-parameter family $\{f_\lambda \in \mathbb{U}^2\}_{\lambda \in \Lambda}$ (where $\Lambda \subset \mathbb{R}$ is an interval), such that the correspondence $(\lambda, x) \mapsto f_\lambda(x)$ is analytic. (The measure-theoretical description of analytic families in several parameters follows from the one-parameter case, see [AM3].)

An analytic family of unimodal maps is called *non-trivial* if regular parameters are dense. If all maps in the family are quasiquadratic, it can be shown that a family is non-trivial if it contains one regular parameter (see Theorem A of [ALM]).

2.4. **Renormalization.** Let $f \in \mathbb{U}^2$. A symmetric (about 0) interval $T \subset I$ is said to be *nice* if the iterates of ∂T never return to $\operatorname{int} T$. A nice interval $T \neq I$ is said to be a restrictive (or periodic) interval of period m for f if $f^m(T) \subset T$ and m is minimal with this property. In this case, the map $A \circ f^m \circ A^{-1} : I \to I$ is again unimodal for some affine homeomorphism $A : T \to I$ and is called a *renormalization*[6] of f. The map $f^m : T \to T$ will be called a prerenormalization of f.

We say that f is *infinitely renormalizable* if there exists arbitrarily small restrictive intervals, and we say it is *finitely renormalizable* otherwise.

Let $\mathcal{F} \subset \mathbb{U}^2$ be the class of Kupka-Smale finitely renormalizable maps whose critical point is recurrent, but not periodic.

The following result shows that when investigating typical properties of analytic unimodal maps, it is enough to deal with the quasiquadratic case.

Theorem 2.1 (Theorem B of [AM3]). *Let f_λ be a non-trivial analytic family of unimodal maps. Then for almost every non-regular parameter λ, f_λ has a renormalization which is quasiquadratic.*

It is easy to check that the conclusions of Theorems A, B, or C do not depend on considering a map or its renormalization. Due to this result, in the arguments to follow, we will concentrate on the description of quasiquadratic map and non-trivial analytic families of quasiquadratic maps.

[6]A more usual convention is to call $A \circ f^m \circ A^{-1}$ a unimodal restriction if $m = 1$, reserving the name renormalization for the case $m > 1$, but we won't make this distinction.

2.5. **Some metric properties.** The condition of negative Schwarzian derivative plays an important role when one needs to do distortion estimates. One of the main tools is the *Koebe Principle*:

Lemma 2.2 (Koebe Principle, see [MS], page 258). *Let $f : T \to \mathbb{R}$ be a diffeomorphism with non-negative Schwarzian derivative. If $T' \subset T$ and both components L and R of $T \setminus T'$ are bigger than $\epsilon|T'|$ then the distortion of $f|T'$ is bounded by $\frac{(1+\epsilon)^2}{\epsilon^2}$. In particular, we have* $\min\{|f(L)|, |f(R)|\} \geq \delta(\epsilon)\epsilon|f(T')|$*, where $\delta(\epsilon) > 0$ satisfies $\lim_{\epsilon\to\infty} \delta(\epsilon) > \frac{9}{100}$.*

Due to the recent results of Kozlovski, the condition of negative Schwarzian is not needed for application of the Koebe Principle (for unimodal maps in $\mathbb{U}^3$ which are Kupka-Smale), see Theorem B of [K1] for instance. We will thus apply the above Koebe Principle without further comments in our setting.

2.5.1. *Hyperbolicity.* It was shown by Mañé [MS] that (for one-dimensional maps of class C^2) the obstruction to uniform expansion lies in critical points and non-repelling periodic orbits. Since quasiquadratic maps in $\mathcal{F}$ do not have non-repelling periodic orbits, this implies:

Lemma 2.3. *Let $f \in \mathcal{F}$ be a quasiquadratic map, and let T be a nice interval. There exists constants $C > 0$, $\lambda > 1$ such that if $f^k(x) \in I \setminus T$, $k = 0, ..., m-1$ then $|Df^m(x)| > C\lambda^m$.*

Corollary 2.4. *Under the hypothesis of the previous lemma, if K is a compact invariant set which does not contain 0, then $f|K$ is uniformly expanding.*

2.6. **Physical measures.** Let μ be a probability measure which is invariant under the dynamics of f. The *basin* of μ is the set of points $x \in I$ such that

$$\lim \frac{1}{m} \sum_{k=0}^{m-1} \delta_{f^k(x)} = \mu \tag{2.7}$$

in the weak sense. We say that μ is a *physical measure* if the basin of μ has positive Lebesgue measure. A quasiquadratic map can have at most one physical measure [BL], which (if it exists) has always a basin of full Lebesgue measure. If f is hyperbolic, then the uniform distribution in the attracting periodic orbit is the physical measure of f. If f is stochastic, that is, it has an absolutely continuous invariant measure μ, then this measure is ergodic and, by Birkhoff's Theorem, it is a physical measure. Notice that there exist quadratic maps without any physical measure, see [MS], Chapter V, Section 5.

If f is stochastic, then it is finitely renormalizable. Let $f^k : T \to T$ be its last prerenormalization. It turns out that the support of μ is $A = T_0 \cup ... \cup T_{k-1}$ where $T_0 = [f^{2k}(0), f^k(0)]$ and $T_j = f^j(T_0)$. Notice that $f^k(T_0) = T_0$. We could have defined A topologically in this way without any reference to μ.

The set A has another remarkable property: it is the smallest compact subset of I such that

1. for almost every $x \in I$, $\omega(x) \subset A$;
2. for generic $x \in I$, $\omega(x) \subset A$.

Those two conditions mean exactly that A is the topological and metric attractor of f in the sense of Milnor.

Remark 2.1. All quasiquadratic unimodal maps have a unique topological and a unique metric attractor. Both concepts of attractor coincide by [L1].

A sufficient condition for f to be stochastic is the Collet-Eckmann condition: $|Df^n(f(0))|$ grows exponentially fast.

Theorem 2.5 (Corollary C of [AM3]). *Let f_λ be a non-trivial family of analytic unimodal maps. Then almost every non-regular parameter belongs to $\mathcal{F}$ and satisfies the Collet-Eckmann condition.*

We will need the following result of Keller about general stochastic unimodal maps:

Theorem 2.6 (see [MS], Theorem 3.2, Chapter V)**.** *Let f be a quasiquadratic stochastic map, and let μ be its physical measure. Then $d\mu$ is uniformly bounded from below on A.*

Remark 2.2. Keller's Theorem is stated in [MS] for maps with negative Schwarzian derivative. The result for quasiquadratic maps can be obtained with the same proof using the results of Kozlovski [K1].

Notice that while $d\mu$ is always bounded from below, it is definitely not bounded from above, and we will need to work a lot to obtain in Theorem C a reasonable estimate for $d\mu$. Notice also that our proof of Theorem C is not a general one for stochastic maps: we have to exclude lots of them. (It is easy to see that some exclusion has to be done, for instance, one must exclude stochastic maps with non-recurrent critical point.)

3. The formula

3.1. **Combinatorics.** Let us have a symbol space Σ with finitely many elements. A (finite or infinite) sequence of elements of Σ will be called a word. In the space $\Sigma^{\mathbb{N}}$ of infinite words, we let the shift operator σ act by $\sigma(\alpha_0\alpha_1...) = \alpha_1\alpha_2....$

Given a finite word α and $r \in \mathbb{N} \cup \{\infty\}$, we let α^r denote r repetitions of α.

A finite word α is said to be irreducible if $\alpha = \beta^r$ for some r implies $\alpha = \beta$.

If α is an infinite word which is periodic, there exists a unique irreducible word $\overline{\alpha}$ such that $\alpha = \overline{\alpha}^\infty$.

3.1.1. *Frequencies.* If $\alpha = \alpha_0...\alpha_{m-1}$ is a finite word and $\beta = \beta_0\beta_1...$ is an infinite word, we define the lower and upper frequencies of α in β in the natural way:

$$r^+(\alpha,\beta) = \limsup_{n\to\infty} \frac{1}{n} \#\{0 \le k \le n-1 | \alpha_i = \beta_{k+i}, 0 \le i \le m-1\}, \tag{3.1}$$

$$r^-(\alpha,\beta) = \liminf_{n\to\infty} \frac{1}{n} \#\{0 \le k \le n-1 | \alpha_i = \beta_{k+i}, 0 \le i \le m-1\}. \tag{3.2}$$

The frequency $r(\alpha,\beta)$ is defined as the common value of $r^+(\alpha,\beta)$ and $r^-(\alpha,\beta)$ if they coincide. We say that β is normal if, for any α, $r^+(\alpha,\beta) = r^-(\alpha,\beta)$.

3.1.2. *Geometric frequencies.* Let α be a finite word and β be a normal infinite word. Let us consider the non-increasing sequence $r(\alpha^k,\beta)$. We want to associate to α and β a quantity related to the decay of $r(\alpha^k,\beta)$. In the case of exponential decay, it is natural to define the upper and lower geometric frequencies:

$$\rho^+(\alpha,\beta) = \limsup_{n\to\infty} r(\alpha^n,\beta)^{1/n}, \tag{3.3}$$

$$\rho^-(\alpha,\beta) = \liminf_{n\to\infty} r(\alpha^n,\beta)^{1/n}. \tag{3.4}$$

The geometric frequency $\rho(\alpha,\beta)$ is the common value of $\rho^+(\alpha,\beta)$ and $\rho^-(\alpha,\beta)$ if they coincide. We say that β is geometrically normal if for any α, $\rho^+(\alpha,\beta)$ and $\rho^-(\alpha,\beta)$ coincide.

3.2. **Itineraries.** Let us associate to an unimodal map f some symbolic dynamics. We fix the symbol space $\Sigma = \{0, c, 1\}$. Let $\Theta : I \to \Sigma$ be defined by $\Theta|[-1,0) = 0$, $\Theta|(0,1] = 1$, and $\Theta(0) = c$.

The *itinerary* of a point $x \in I$ is the infinite word $\theta(x) = \theta_0\theta_1...$, where $\theta_k = \Theta(f^k(x))$.

The (discontinuous) map $\theta : I \to \Sigma^{\mathbb{N}}$ satisfy $\theta \circ f = \sigma \circ \theta$. It is clear that if p is a periodic point for f, then $\theta(p)$ is a periodic word for σ.

Given a word α, we let $I_\alpha \subset I$ be the set of points whose itinerary starts with α. Depending on α, I_α can be either an interval, a point or empty.

3.3. **Proof of Theorem A assuming Theorems B and C.** We will actually prove the following stronger:

Theorem 3.1. *Let f be a quasiquadratic unimodal map such that*

(1) *f is Collet-Eckmann and has an absolutely continuous invariant measure μ supported in a cycle of intervals A;*
(2) *0 belongs to the basin of μ;*
(3) *For any invariant hyperbolic set K, and any $1 \le p < \infty$, $d\mu_f^K \in L^p$.*

Then $\theta(0)$ is geometrically normal and for any $p \in A$ periodic (of period m),

$$|Df^m(p)| = \rho(\overline{\theta(p)}, \theta(0))^{-1}. \tag{3.5}$$

Moreover, for any α such that $\rho(\alpha, \theta(0)) > 0$, there exists a periodic orbit $p \in A$ such that $\theta(p) = \alpha^\infty$.

Proof. Let $\theta(0) = \theta_0\theta_1....$ Let $\alpha = \alpha_0...\alpha_{m-1}$ be an arbitrary finite word. Notice that $\theta_{k+j} = \alpha_j$, $0 \le j \le m-1$, if and only if $f^k(0) \in I_\alpha$, so by definition of basin of μ, $r(\alpha, \theta(0)) = \mu(I_\alpha)$. In particular, $\theta(0)$ is normal.

Let $p \in A$ be a periodic orbit, and let $\alpha = \overline{\theta(p)}$. By item 1, we conclude that p is repelling, and since f is quasiquadratic, $\cap I_{\alpha^k} = p$, and the length m of α is the period of p. Let $q, q' \in I_\alpha$ be periodic orbits in opposite sides of p, and let $q_k = (f^{km}|I_{\alpha^{k+1}})^{-1}(q)$ and $q'_k = (f^{km}|I_{\alpha^{k+1}})^{-1}(q')$. Let K be the hyperbolic set consisting of p, the forward orbit of q and q' and all q_k and q'_k. Let $T_k = [q'_k, q_k]$. It is easy to see that there exists $j > 0$ such that for all $k > j$,

$$T_{k+j} \subset I_{\alpha^k} \subset T_{k-j}. \tag{3.6}$$

In particular,

$$\rho^+(\alpha, \theta(0)) = \limsup_{n\to\infty} \mu(T_n)^{1/n}, \tag{3.7}$$

$$\rho^-(\alpha, \theta(0)) = \liminf_{n\to\infty} \mu(T_n)^{1/n}. \tag{3.8}$$

By Theorem 2.6, there exists a constant $C > 0$ such that $d\mu|A \ge C$. On the other hand, $T_k \subset A$ for k big enough, so $\mu(T_k) \ge C|[q'_k, q_k]|$. It is clear that

$$\lim_{n\to\infty} |T_n|^{1/n} = |Df^m(p)|^{-1}, \tag{3.9}$$

so $\rho^-(\alpha, \theta(0)) \ge |Df^m(p)|^{-1}$.

By item 3, for all $1 \le p < \infty$, there exists a constant C_p such that, for all $k \ge 0$,

$$\left(\int_{T_k} (d\mu^K)^p\right)^{1/p} \le C_p. \tag{3.10}$$

In particular, by the Hölder inequality,

$$\mu(T_k) = \int_{T_k} d\mu^K \le \left(\int_{T_k} (d\mu^K)^p\right)^{\frac{1}{p}} \left(\int_{T_k} 1^{\frac{p}{p-1}}\right)^{1-\frac{1}{p}} \le C_p|T_k|^{1-\frac{1}{p}}. \tag{3.11}$$

Taking $k \to \infty$ we get

$$\rho^+(\alpha, \theta(0)) \leq |Df^m(p)|^{-1+\frac{1}{p}}. \tag{3.12}$$

Since $1 \leq p < \infty$ is arbitrary, we get (3.5).

If α is an arbitrary finite word, then either $I_{\alpha^k} \cap A$ is eventually empty or $\cap I_{\alpha^k}$ is a repelling periodic orbit p in A. In the first case, obviously $\rho(\alpha, \theta(0)) = 0$. In the second case, by the previous discussion, $\rho(\alpha, \theta(0)) = |Df^m(p)|^{-1} > 0$, where m is the length of α. In particular, $\theta(0)$ is geometrically normal. □

Remark 3.1. Let us note that the Collet-Eckmann condition already implies a number of interesting properties (see [NS]). For instance, if f is a quasiquadratic Collet-Eckmann map, then there exists a constant $\lambda > 1$ such that if p is a periodic orbit of period n then $|Df^n(p)| \geq \lambda^n$.

4. Phase-parameter relation and statistics of the principal nest

In this section we will discuss the principal nest combinatorics, and then state the Phase-Parameter relation, which is our means to obtain parameter results based on phase estimates. We will then present some results on the statistics of the principal nest.

4.1. Principal nest combinatorics.

If $T \subset I$ is a nice interval, the domain of the first return map R_T to T consists of a (at most) countable union of intervals which we denote T^j. We reserve the index 0 for the component of 0: $0 \in T^0$, if 0 returns to T. From the nice condition, $R_T|T^j$ is a diffeomorphism if $0 \notin T^j$, and is an even map if $0 \in T^j$. The domain containing 0 will be called the central domain of R_T and will be denoted T^0. The return R_T is said to be central if $R_T(0) \in T^0$. If f is quasiquadratic with recurrent but not periodic critical point, the domain of the first return map is dense and its complement is a regular Cantor set.

Let $f \in \mathcal{F}$ be quasiquadratic, and let T be its smallest restrictive interval (of period m'). Define a sequence of nested nice intervals I_n by induction as follows. Let $I_0 = [-p, p]$ where p is the unique orientation reversing fixed point of $f^{m'} : T \to T$. Assuming I_n defined, let $R_n : I_n \to I_n$ be the first return map and $I_{n+1} = I_n^0$. Since f is finitely renormalizable, $\cap I_n = \{0\}$.

Let Ω be the set of all finite sequences of non-zero integers (possibly empty). For any element $\underline{d} \in \Omega$, $\underline{d} = (j_1, ..., j_m)$ we associate a branch $R_n^{\underline{d}}$ of R_n^m, whose domain is $I_n^{\underline{d}} = \{x \in I_n | R^k(x) \in I_n^{j_{k+1}}, 0 \leq k < m\}$.

Let $L_n : I_n \to I_n^0$ be the first landing map. The domain of L_n is the union of intervals $C_n^{\underline{d}} = (R_n^{\underline{d}})^{-1}(I_n^0)$.

4.2. Phase-Parameter relation.

We will now quickly define formally the Phase-Parameter relation, and we will discuss in the next section the way it is applied for measure-theoretical problems.

Definition 4.1. Let us say that a family f_λ of quasiquadratic maps satisfies the Topological Phase-Parameter relation at a parameter λ_0 if $f = f_{\lambda_0} \in \mathcal{F}$, and there exists $i_0 > 0$ and a sequence of nested intervals J_i, $i \geq i_0$ such that:

1. J_i is the maximal interval containing λ_0 such that for all $\lambda \in J_i$ there exists a homeomorphism $H_i[\lambda] : I \to I$ such that $f_\lambda \circ H_i[\lambda] | (I \setminus I_{i+1}) = H_i[\lambda] \circ f$.
2. There exists a homeomorphism $\Xi_i : I_i \to J_i$ such that $\Xi_i(C_i^{\underline{d}})$ (respectively, $\Xi_i(I_i^{\underline{d}})$) is the set of λ such that the first return of 0 to $H_i[\lambda](I_i)$ under iteration by f_λ belongs to $H_i[\lambda](C_i^{\underline{d}})$ (respectively, $H_i[\lambda](I_i^{\underline{d}})$).

Let K_i be the closure of the union of all $\partial C_i^{\underline{d}}$ and $\partial I_i^{\underline{d}}$. Notice that H_i and Ξ_i are only uniquely defined in K_i. Condition (2) of the Topological Phase-Parameter relation can be equivalently formulated as the existence of a homeomorphism $\Xi_i : I_i \to J_i$ such that the first return of the critical point (under iteration by f_λ) to $H_i[\lambda](I_i)$ belongs to $H_i[\lambda](K_i)$ if and only if $\lambda \in \Xi_i(K_i)$.

Let us assume we have a non-trivial family of unimodal maps satisfying the Phase-Parameter relation at a parameter $f = f_{\lambda_0}$. It will be important to estimate the metric properties of $H_i|K_i$ and $\Xi_i|K_i$.

Let $\tilde{I}_{i+2} = (R_i|I_i^0)^{-1}(I_i^{\underline{d}})$, where $\underline{d}$ is such that $\overline{(R_i|I_i^0)^{-1}(C_i^{\underline{d}})} = I_{i+2}$.

Let τ_i be such that $R_i(0) \in I_i^{\tau_i}$. Let $\tilde{K}_i = (\cup_j \partial I_i^j \cup \partial I_i) \setminus \tilde{I}_{i+1}$.

Let $J_i^j = \Xi_i(I_i^j)$.

Let us say that $f \in \mathcal{F}$ is simple if only finitely many R_n have central returns.

Definition 4.2. Let f_λ be a family of unimodal maps. We say that f_λ satisfies the Phase-Parameter relation at λ_0 if $f = f_{\lambda_0}$ is simple, f_λ satisfies the Topological Phase-Parameter relation at λ_0 and for every $\gamma > 1$, there exists $i_0 > 0$ such that for $i > i_0$ we have:

PhPa1: $\Xi_i|(K_i \cap I_i^{\tau_i})$ is γ-qs,
PhPa2: $\Xi_i|\tilde{K}_i$ is γ-qs,
PhPh1: $H_i[\lambda]|K_i$ is γ-qs if $\lambda \in J_i^{\tau_i}$,
PhPh2: the map $H_i[\lambda]|\tilde{K}_i$ is γ-qs if $\lambda \in J_i$.

Theorem 4.1 (Theorem A of [AM3]). *Let f_λ be a non-trivial analytic family of quasiquadratic maps. Then f_λ satisfies the Phase-Parameter relation at almost every non-regular parameter.*

(Theorem A of [AM3] actually covers the non-quasiquadratic case as well.)

4.3. **Using the Phase-Parameter relation.** Let us now explain how the Phase-Parameter relation can be used to prove that some property is typical among non-regular analytic unimodal maps.

Notice that, due to the previous results, it is enough to prove that the property is satisfied by almost every parameter in a non-trivial analytic family of quasiquadratic maps. *From now on we shall always work inside such a fixed family.* We can further restrict our scrutiny to the subset of parameters which are simple and satisfy the Phase-Parameter relation. It is also clearly enough to restrict ourselves to the analysis of unimodal maps which are exactly k-times renormalizable for some fixed (but arbitrary) k. *We shall use "with total probability" to denote some property that is valid for a full measure set of parameters under those restrictions.*

We will now illustrate the basic principle with an example worked out in [AM1].

For a simple map $f = f_{\lambda_0}$ which is quasiquadratic, simple and satisfies the Phase-Parameter relation, let us associate a sequence of "statistical parameters" in the following way. Let s_n be the number of times the critical point 0 returns to I_n before the first return to I_{n+1}. Let $c_n = |I_{n+1}|/|I_n|$. 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 about c_n^{-1}.

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} (to be computed precisely using a statistical argument, in this case, fixing some small $\epsilon > 0$, we can take the neighborhood to be $c_n^{-1+2\epsilon} < k < c_n^{-1-\epsilon}$ for n big). By most, we mean that, say, the complement has at most probability α_n which is some summable sequence. In this case, it is not hard to see that we can take $\alpha_n = c_n^\epsilon$, which indeed decays exponentially, and so is summable, for all simple maps f by [L1].

If the phase-parameter relation were 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 Borel-Cantelli, for almost every parameter this only happens a finite number of times. More precisely, we would use the following version of Borel-Cantelli:

Lemma 4.2 (Lemma 3.1 of [AM1])**.** *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 $|Y| = |X|$.*

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 to not fall into. In order to do so, in the statistical analysis of the sets A_k we control the quasisymmetric capacity (instead of Lebesgue measure) of the complement of the set of points whose entrance times belong to the good neighborhood. This makes the analysis sometimes much more difficult: capacities are not probabilities (since they are not additive), so 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 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. It will indeed be very important for us that the Phase-Parameter relation we use provides constants near 1, since this will allow us to partially get rid of those error terms. This is also the reason that the estimates in [AM2] (which employed weaker Phase-Parameter estimates) are worse than [AM1].

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 $C_n^{\underline{d}}$, and 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:

Lemma 4.3. *For almost every parameter, for every $\epsilon > 0$, there exists $\gamma > 1$ such that $p_\gamma(X_n|I_n^\tau)$ 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.*

And as a consequence of PhPa1 we get:

Lemma 4.4. *With total probability, for all $\epsilon > 0$, for all n sufficiently big,*

$$c_n^{-1+\epsilon} < s_n < c_n^{-1-\epsilon}. \tag{4.1}$$

In the language of Lemma 4.2, X would be the set of simple quasiquadratic parameters satisfying the Phase-Parameter relation and which are exactly k-times renormalizable, $D_n(\lambda)$, $\lambda \in X$ would be $J_n^{\tau_n}(\lambda)$, and $Q_n \subset X$ would be the set of parameters such that either $s_n < c_n^{-1+\epsilon}$ or $s_n > c_n^{-1-\epsilon}$.

4.4. **Some results on the statistics of the principal nest.** Let us collect here some results of [AM1] on the dynamics of typical non-regular analytic unimodal maps (the results were initially proved in the quadratic setting, but hold in general due to [AM3]).

Let $r_n(j)$ be such that $R_n|I_n^j = f^{r_n(j)}$. For $x \in I_n^j$, we let $r_n(x) = r_n(j)$. Let $l_n(\underline{d})$ be such that $L_n|C_n^{\underline{d}} = f^{l_n(\underline{d})}$, and for $x \in C_n^{\underline{d}}$, let $l_n(x) = l_n(\underline{d})$. Let $v_n = r_n(0)$. Recall that we have defined $s_n = |\underline{d}|$ where $R_n(0) \in C_n^{\underline{d}}$, so that $R_{n+1}(0) = R_n^{s_n+1}(0)$. Let $c_n = |I_{n+1}|/|I_n|$.

We define the following convenient notation

$$I_n^X = \bigcup_{j \in X} I_n^j, \qquad I(X,n) = \frac{|I_n^X|}{|I_n|} = \sum_{j \in X} \frac{|I_n^j|}{|I_n|}, \qquad X \subset \mathbb{Z}, \tag{4.2}$$

$$I_n^X = \bigcup_{\underline{d} \in X} I_n^{\underline{d}}, \qquad I(X,n) = \sum_{\underline{d} \in X} \frac{|I_n^{\underline{d}}|}{|I_n|}, \qquad X \subset \Omega, \tag{4.3}$$

$$C_n^X = \bigcup_{\underline{d} \in X} C_n^{\underline{d}}, \qquad C(X,n) = \frac{|C_n^X|}{|I_n|} = \sum_{\underline{d} \in X} \frac{|C_n^{\underline{d}}|}{|I_n|}, \qquad X \subset \Omega. \tag{4.4}$$

(Thus I_n^X and $I(X,n)$ are defined both for $X \subset \mathbb{Z}$ and $X \subset \Omega$.)

The following summarizes Lemma 4.3, Corollaries 6.8 and 6.10, and Remark 6.3 of [AM1].

Lemma 4.5. *Almost every non-regular map satisfies*

$$\lim \frac{\ln v_{n+1}}{\ln c_n^{-1}} = \lim \frac{\ln s_n}{\ln c_n^{-1}} = \lim \frac{\ln \ln c_{n+1}^{-1}}{\ln c_n^{-1}} = \lim \frac{\ln r_n(\tau_n)}{\ln c_{n-1}^{-1}} = 1, \tag{4.5}$$

In particular, c_n decays very fast (this type of decay is called torrential).

4.4.1. *Distortion estimates.* Let us now discuss some estimates on the position of the critical value of the return maps R_n, which are relevant for distortion estimates. The following summarizes Lemmas 4.8 and 4.10 (and their proof) of [AM1].

Lemma 4.6. *For almost every non-regular map, for every $\delta > 0$, for any n big enough, the following holds:*

1. $|R_n(0)| > n^{-1-\delta}|I_n|$, *and in particular,* $R_n(0) \notin \tilde{I}_{n+1}$,
2. *The distance between $R_n(0)$ to ∂I_n is at least $n^{-1-\delta}|I_n|$,*
3. *For any $\underline{d} \in \Omega$, if $R_n(0) \notin C_n^{\underline{d}}$, then the distance between $R_n(0)$ and $C_n^{\underline{d}}$ is at least $n^{-1-\delta}|C_n^{\underline{d}}|$,*
4. *For any $\underline{d} \in \Omega$, $\mathrm{dist}(R_n^{\underline{d}}) \leq n^{\frac{1}{2}+\delta}$.*

The estimate above for distortion of branches $R_n^{\underline{d}}$ is relatively pessimistic. For most branches, we have much better bounds. Indeed, if $R_{n-1}(I_n^j) \subset C_{n-1}^{\underline{d}}$ and $R_{n-1}(0) \notin I_{n-1}^{\underline{d}}$, then $\mathrm{dist}(f|I_n^j) - 1$ is at most of order of the quotient of $|I_n^j|$ by the distance from I_n^j to 0 (this can be bounded from above by $O(|C_{n-1}^{\underline{d}}|/|I_{n-1}^{\underline{d}}|)$ because $R_{n-1}(0) \notin I_{n-1}^{\underline{d}}$), so $\mathrm{dist}(f|I_n^j) = 1 + O(c_{n-1})$. Since $R_n|I_n^j$ is the composition of $f|I_n^j$ and a diffeomorphism onto I_n (which extends to I_{n-1}) with distortion bounded by $1 + O(c_{n-1})$ (by the Koebe principle), we see that for all those branches the distortion of R_n is at most $1 + O(c_{n-1})$.

Notice that for any j, both components of $I_n \setminus I_n^j$ have size at least $|I_n^j| 2^n c_{n-1}^{-1/2}$. Indeed, let $R_{n-1}(I_n^j) \subset C_{n-1}^{\underline{d}}$. Each connected component of $I_{n-1} \setminus C_{n-1}^{\underline{d}}$ must have size at least of order $2^{4n} c_{n-1}^{-1} |C_{n-1}^{\underline{d}}|$ (which implies the desired estimate), unless $|\underline{d}| = 0$ (that is $C_{n-1}^{\underline{d}} = I_n$). In this last case, the first item of the previous lemma implies that each connected component of $I_n \setminus I_n^j$ has size at least of order $2^{-n} c_{n-1}^{-1} |I_n^j| \geq 2^n c_{n-1}^{-1/2} |I_n^j|$.

In particular, if $\operatorname{dist}(R_n|I_n^j) = 1 + O(c_{n-1})$ and the last entry of $\underline{d}$ is j, we can also find better bounds for the distortion of $R_n^{\underline{d}}$. Indeed, $R_n^{\underline{d}}$ is the composition of a map onto I_n^j which extends to I_n, and has distortion bounded by $1 + o(c_{n-1}^{1/2})$ and $R_n|I_n^j$, so we have $\operatorname{dist}(R_n^{\underline{d}}) = 1 + o(c_{n-1}^{1/2})$.

4.4.2. *Estimates on the capacity of some relevant sets.* In the course of proving the above estimates, one obtains several estimates for the quasisymmetric capacities of certain sets, which will be important here. In order to be definite, let $\epsilon = \epsilon(\gamma)$ be the smallest number such that, for $\kappa = 1 + \frac{\epsilon}{5}$ and for any γ-qs map h we have

$$\frac{1}{\kappa}\left(\frac{|J|}{|I|}\right)^{\kappa} \leq \frac{|h(J)|}{|h(I)|} \leq \left(\frac{\kappa|J|}{|I|}\right)^{1/\kappa}, \tag{4.6}$$

so that $\epsilon(\gamma) \to 0$ as $\gamma \to 1$.

The following summarizes Corollaries 6.5 and 6.7 of [AM1].

Lemma 4.7. *For almost every non-regular map, if $\epsilon_0 = \epsilon(\gamma) < 1/100$, then, for n large enough*

$$p_\gamma(r_n(x) > k c_n^{-4}|I_n) \leq e^{-k}, \quad k \geq 1, \tag{4.7}$$

$$p_\gamma(r_n(x) < c_{n-1}^{-1+2\epsilon_0}|I_n) \leq c_{n-1}^{\epsilon_0/10}, \tag{4.8}$$

$$p_\gamma(r_n(x) > c_{n-1}^{-1+2\epsilon_0}|I_n) \leq e^{-c_{n-1}^{-\epsilon_0/5}}. \tag{4.9}$$

5. The critical orbit is typical

5.1. **Outline.** Let us summarize the main steps in the proof of Theorem B.

(1) We must show that (with total probability) the proportion of time the critical orbit spends in any given interval $T \subset I$ is given by $\mu(T)$. It is of course enough to consider a countable class of intervals which generates all Borelians, and then prove the distribution result (with total probability) for each interval in the class. Our choice of intervals will be domains ξ of the first landing map from I to I_{n_0} (for arbitrary n_0).

(2) We must be able to estimate $\mu(\xi)$ in terms of return branches. Let $\psi_n^\xi(x)$ be the frequency of visits to ξ of the iterates of a point $x \in I_n$ before x returns to I_n ($\psi_n^\xi(x)$ only depends on the branch I_n^j containing x). We show that ψ_n^ξ is concentrated around $\mu(\xi)$ and indeed we show that $\mu(\xi)$ is the unique number q such that, for every $\epsilon > 0$, we have $\lim_{n\to\infty} p(|\psi_n^\xi(x) - q| > \epsilon|I_n) = 0$.

(3) We use an explicit Large Deviation Estimate to obtain a quantitative estimate on the rate of decay of $p(|\psi_n^\xi(x) - \mu(\xi)| > \epsilon|I_n)$ (in n) using only the fact that it decays to 0. We obtain a torrential estimate ($p(|\psi_n^\xi(x) - \mu(\xi)| > \epsilon|I_n) < c_{n-1}^{1/20}$).

(4) We would like to show that returns $R_n(0)$ of the critical point belong to branches of R_n with "close to correct" distribution on ξ, that is $|\psi_n^\xi(R_n(0)) - \mu(\xi)| < \epsilon$. The previous estimate indicate that this should be the case, but the Phase-Parameter relation is just quasisymmetric. We show that the torrential rate of decay still holds if instead of probabilities $p(|\psi_n^\xi(x) - \mu(\xi)| > \epsilon|I_n)$ we consider qs-capacities $p_{\gamma(n)}(|\psi_n^\xi(x) - \mu(\xi)| > \epsilon|I_n)$, provided we choose $\gamma(n)$ very close to 1. This argument does not give any reasonable bound on the rate of decay of $\gamma(n)$ to 1, it could be very fast.

(5) We want to show that we may actually take $\gamma(n)$ as a constant γ bigger than 1. For this we argue that a torrentially small set of branches (in the $\gamma(n_s)$-qs) of a fixed level n_s has torrentially small effect (in the γ-qs sense for some fixed $1 < \gamma < \gamma(n_s)$) with respect to total (and partial) time of branches in the subsequent levels. This argument follows the proof of the Collet-Eckmann condition in [AM1], where we used those ideas to control the propagation of weakly hyperbolic branches. A little bit of change is needed in order to avoid a loss of the quasisymmetric constant of level n_s, on

which we do not have control. For this reason, we will work with modified quasisymmetric capacities in some arguments.
(6) As a consequence, we see that except for a set with torrentially small γ-qs capacity, return branches of level n are "very good" in the sense that they spend most of their time following branches of level n_s which satisfy $|\psi^\xi_{n_s} - \mu(\xi)| < \epsilon$. As a consequence, those "very good" return branches of level n satisfy $|\psi^\xi_n - \mu(\xi)| < 2\epsilon$. As a bonus from the previous item we get for free the estimates for intermediate moments (not just full returns), which are needed also in the proof of the Collet-Eckmann condition.
(7) Using the Phase-Parameter relation we make the critical point falls in "very good" branches. Thus the distribution of the critical orbit on ξ is 2ϵ close to $\mu(\xi)$. Making ϵ goes to 0 we obtain Theorem B.

5.2. **Inductive estimates.** In this section we will show that a small (in the quasisymmetric sense) set of branches of level n_0 has a small effect on most (in the quasisymmetric sense) branches of level $n \geq n_0$. This kind of argument was already needed in the analysis of [AM1], so we will keep a similar notation to that work, and will refer to it for some computations.

5.2.1. *Modified capacities.* For our application, we will need a modification of the γ-qs capacities used by [AM1]. This is not the same modification used by [AM3].

We say that h is a (γ, C)-homeomorphism if $h = h_2 \circ h_1$ where h_2 is γ-qs and h_1 is C^1 with distortion bounded by C.

If $X \subset I$ is a Borelian set, we let

$$p_{\gamma,C}(X|I) = \sup \frac{|h(X \cap I)|}{|h(I)|} \tag{5.1}$$

where h ranges over all (γ, C)-homeomorphisms.

Through the end of this section we will fix ϵ_0 very small (say, 1/1000), but we won't need to make $\epsilon_0 \to 0$ later on. Choose $\hat{\gamma}$ very close to 1 so that $\epsilon(\hat{\gamma}) \leq \epsilon_0$, in the notation of §4.4.2

Let us fix C and γ_0 close to 1 so that for n big, any $(C\frac{n+1}{n}, \gamma_0)$ homeomorphism is a $\hat{\gamma}$-qs homeomorphism. Let $C_n = C \cdot \frac{n+1}{n}$, $\tilde{C}_n = C\frac{2n+3}{2n+1}$.

In what follows, we will work with some fixed $1 \leq \gamma \leq \gamma_0$, *but the estimates will be uniform for* γ *in this range*, and with the sequences C_n and $\tilde{C}_n$.

We will use (γ, C_n) capacities to estimate the size of sets of return branches of level n and $(\gamma, \tilde{C}_n)$ for sets of landing branches of level n.

The introduction of those constants is motivated by the following result which can be proved using the methods of [AM1].

Lemma 5.1 (Analogous to Remarks 5.1 and 5.2 of [AM1])**.** *With total probability, there exists n_0 such that for $n > n_0$ and for all $1 \leq \gamma \leq \gamma_0$, the following holds. If $X \subset I_n$ then*

$$p_{\tilde{C}_n,\gamma}((R^{\underline{d}}_n)^{-1}(X)|I^{\underline{d}}_n) \leq 2^n p_{C_n,\gamma}(X|I_n). \tag{5.2}$$

And if $X \subset I_n$ and

$$p_{\tilde{C}_n,\gamma}(X|I_n) \leq \delta \leq 2^{-n^2}. \tag{5.3}$$

then

$$p_{C_{n+1},\gamma}((R_n|I^0_n)^{-1}(X)|I_{n+1}) \leq \delta^{1/5}. \tag{5.4}$$

Induction applied to (5.2) gives:

Lemma 5.2 (Analogous to Lemma 5.4 of [AM1]). *With total probability, there exists n_0 such that for $n > n_0$ and all $1 \leq \gamma \leq \gamma_0$ the following holds. Let $Q_n \subset \mathbb{Z}$ and let $Q_n(m,r)$ be the set of all $\underline{d}$ with length m and at least r entries on Q_n. Let*

$$q_n = p_{\gamma,C_n}(I_n^{Q_n}|I_n), \tag{5.5}$$

$$q_n(m,r) = p_{\gamma,\tilde{C}_n}(I_n^{Q_n(m,r)}|I_n). \tag{5.6}$$

Then

$$q_n(m,r) \leq \binom{m}{r}(2^n q_n)^r. \tag{5.7}$$

More generally, for any fixed $\underline{d}$, defining

$$q_n^{\underline{d}}(m,r) = p_{\gamma,\tilde{C}_n}((R_n^{\underline{d}})^{-1}(I_n^{Q_n(m,r)})|I_n^{\underline{d}}), \tag{5.8}$$

we have

$$q_n^{\underline{d}}(m,r) \leq \binom{m}{r}(2^n q_n)^r. \tag{5.9}$$

This estimate will be mainly used to estimate $q_n(m,r)$ for m large and $\frac{r}{m}$ larger than $(6 \cdot 2^n)q_n$. Notice that if $q^{-1} \geq 6 \cdot 2^n$ and $q \geq q_n$ then by Stirling formula,

$$q_n(m,(6 \cdot 2^n)qm) \leq 2^{-(6 \cdot 2^n)qm}, \tag{5.10}$$

and

$$\sum_{k \geq q^{-2}} q_n(k,(6 \cdot 2^n)qk) \leq 2^{-n}q^{-1}2^{-(6\cdot 2^n)q^{-1}}. \tag{5.11}$$

5.2.2. *Estimates on time.* Following [AM1], we define the set of standard landings at time n, $LS(n) \subset \Omega$ as the set of all $\underline{d} = (j_1,...,j_m)$ satisfying the following.

$$c_n^{-1/2} < m < c_n^{-1-2\epsilon_0}, \tag{LS1}$$

$$r_n(j_i) < c_{n-1}^{-14}, \quad \text{for all } i, \tag{LS2}$$

$$\#\{1 \leq i \leq k,\, r_n(j_i) < c_{n-1}^{-1+2\epsilon_0}\} < (6 \cdot 2^n)c_{n-1}^{\epsilon_0/10}k, \quad \text{for } c_{n-1}^{-2} \leq k \leq m, \tag{LS3}$$

$$\#\{1 \leq i \leq k,\, r_n(j_i) > c_{n-1}^{-1-2\epsilon_0}\} < (6 \cdot 2^n)e^{-c_{n-1}^{-\epsilon_0/5}}k, \quad \text{for } c_n^{-1/n} \leq k \leq m. \tag{LS4}$$

The following estimate was obtained in [AM1], and is a consequence of the estimates of §4.4.2:

Lemma 5.3 (Analogous to Lemma 7.1 of [AM1]). *Let $LS(n)$ denote the set of standard landings. Then*

$$p_{\hat{\gamma}}(C_n^{\Omega \setminus LS(n)}|I_n) < c_n^{1/3}. \tag{5.12}$$

$$p_{\hat{\gamma}}(C_n^{\Omega \setminus LS(n)}|I_n^{\tau_n}) < c_n^{1/3}. \tag{5.13}$$

Let $T \subset \mathbb{Z}$ be given. Let us define $VG(T,n_0,n) \subset \mathbb{Z}$ and $LE(T,n_0,n) \subset \Omega$ inductively as follows. Let $VG(T_{n_0},n) = \mathbb{Z} \setminus T$. Assuming $VG(T,n_0,n)$ defined, let $LE(T,n_0,n)$ be the set of all $\underline{d} \in LS(n)$ such that $\underline{d} = (j_1,...,j_m)$ and

$$\#\{j_i \notin VG(T,n_0,n), 1 \leq i \leq k\} < (6 \cdot 2^n)c_n^{1/20}k, \quad \text{for } c_{n-1}^{-2} \leq k \leq m. \tag{LE}$$

Let $VG(T,n_0,n+1)$ be the set of all j such that $R_n(I_{n+1}^j) \subset LE(T,n_0,n)$.

In what follows, we will put the condition that T is a small set of branches of some (deep) level n_0 in the sense that

$$p_{\gamma,C_{n_0}}(I_{n_0}^T|I_{n_0}) < c_{n_0-1}^{1/20} \tag{5.14}$$

for some n_0 and some $1 \leq \gamma \leq \gamma_0$.

The definition of the class $VG(T, n_0, n)$ is designed so that such branches do not pass very often by T before returning. The precise constants in the definition were chosen so that they allow to show that $VG(T, n_0, n)$ correspond to most branches of level n (by induction). Those two estimates are given below:

Lemma 5.4 (see also Lemma 7.2 of [AM1])**.** *With total probability, for all n_0 sufficiently big, if T satisfies (5.14) for some $1 \leq \gamma \leq \gamma_0$ then for all $n \geq n_0$, we have*

$$p_{\gamma,\tilde{C}_n}(C_n^{\Omega \setminus LE(T,n_0,n)}|I_n) < c_n^{2/7} \tag{5.15}$$

$$p_{\gamma,C_n}(I_n^{\mathbb{Z} \setminus VG(T,n_0,n)}|I_n) < c_{n-1}^{1/20}. \tag{5.16}$$

Furthermore,

$$p_{\gamma,\tilde{C}_n}(C_n^{\Omega \setminus LE(T,n_0,n)}|I_n^{\tau_n}) < c_n^{2/7}. \tag{5.17}$$

Proof. If (5.15) is valid for n then by (5.4) we get

$$p_{\gamma,C_{n+1}}(I_{n+1}^{\mathbb{Z} \setminus VG(T,n_0,n+1)}|I_{n+1}) < c_n^{2/35} < c_n^{1/20} \tag{5.18}$$

which gives (5.16) for $n+1$.

Let us assume the validity of (5.16) for n. Then the $(\tilde{C}_n, \gamma)$-capacity of the set of standard landings which fail to satisfy LE is much less than c_n, by (5.10). Using Lemma 5.3 we get

$$p_{\gamma,\tilde{C}_n}(C_n^{\Omega \setminus LE(T,n_0,n)}|I_n) < c_n^{1/3} + c_n \leq c_n^{2/7}. \tag{5.19}$$

This implies that (5.15) is valid for n. A similar computation gives (5.17) for n.

Since (5.16) is valid for n_0 by hypothesis, we get (5.15), (5.16) and (5.17) for all n by induction. $\square$

Lemma 5.5 (Analogous to Lemma 7.6 of [AM1])**.** *With total probability, for all n_0 big enough and for all $n \geq n_0$, the following holds.*

Let $j \in VG(T, n_0, n+1)$, and let $\underline{d}$ be such that $R_n(I_{n+1}^j) \subset C_n^{\underline{d}}$ and $\underline{d} = (j_1, ..., j_m)$. Let $c_n^{-2/n} < k \leq r_{n+1}(j)$. Let m_k be biggest possible with

$$v_n + \sum_{j=1}^{m_k} r_n(j_i) \leq k \tag{5.20}$$

$$\beta_k = \sum_{\substack{1 \leq i \leq m_k, \\ j_i \in VG(T,n_0,n)}} r_n(j_i). \tag{5.21}$$

Then $1 - \frac{\beta_k}{k} < c_{n-1}^{1/100}$.

Lemma 5.6. *With total probability, for all n_0 big enough and for all $n \geq n_0$, the following holds. Let $j \in VG(T, n_0, n+1)$ and $x \in I_{n+1}^j$, and let $c_n^{-2/n} \leq k \leq r_{n+1}(x)$. Then*

$$\sum_{\substack{i<k, \\ f^i(x) \in I_{n_0}^T}} r_{n_0}(f^i(x)) < c_{n_0-1}^{1/200} k, \tag{5.22}$$

Proof. Let $\alpha_n = \sum_{k=n_0}^{n-1} c_{k-1}^{1/110} < c_{n_0-1}^{1/200}$. We show by induction that if

$$\sum_{\substack{i<r_n(x), \\ f^i(x) \in I_{n_0}^T}} r_{n_0}(f^i(x)) \leq \alpha_n r_n(x), \quad \text{for all } x \in I_n^{VG(T,n_0,n)}, \tag{5.23}$$

then

$$\sum_{\substack{i<k,\\ f^i(x)\in I^T_{n_0}}} r_{n_0}(f^i(x)) < \alpha_{n+1}k, \quad \text{for all } x \in I^{VG(T,n_0,n+1)}_{n+1}, c_n^{-2/n} \le k \le r_{n+1}(x). \tag{5.24}$$

Indeed (using the notation of Lemma 5.5),

$$\sum_{\substack{i<k,\\ f^i(x)\in I^T_{n_0}}} r_{n_0}(f^i(x)) \le k-\beta_k+\alpha_n\beta_k+c_{n-1}^{-14} \le \left(1-\frac{\beta_k}{k}+\alpha_n+c_{n-1}^{-14}c_n^{2/n}\right)k \le \alpha_{n+1}k. \tag{5.25}$$

This gives our result by induction, since for $n=n_0$, the left side of (5.23) is 0. □

5.2.3. *Control of intermediate times.* Let us define $LC(T,n_0,n)\subset\Omega$, $n_0,n\in\mathbb{N}$, $n\ge n_0$ as the set of all $\underline{d}=(j_1,...,j_m)$ in $LE(T,n_0,n)$ satisfying

$$j_i \in VG(T,n_0,n), \quad 1\le i\le c_{n-1}^{-1/30}, \tag{LC1}$$

$$\#\{1\le i\le k,\, r_n(j_i)<c_{n-1}^{-1+2\epsilon_0}\} < (6\cdot 2^n)c_{n-1}^{\epsilon_0/10}k, \quad \text{for } c_{n-1}^{-\epsilon_0/5}\le k\le m, \tag{LC2}$$

$$\#\{1\le i\le k,\, j_i\notin VG(n_0,n)\} < (6\cdot 2^n)c_{n-1}^{1/60}k, \quad \text{for } c_{n-1}^{-1/30}\le k\le m, \tag{LC3}$$

$$\#\{1\le i\le k,\, r_n(j_i)>c_{n-1}^{-1-2\epsilon_0}\} < (6\cdot 2^n)c_{n-1}^{100}k, \quad \text{for } c_{n-1}^{-200}\le k\le m, \tag{LC4}$$

$$r_n(j_i)<c_{n-1}^{-1-2\epsilon_0}, \quad 1\le i\le e^{c_{n-1}^{-\epsilon_0/5}}/2. \tag{LC5}$$

Lemma 5.7 (Analogous to Lemma 7.7 of [AM1]). *For all n_0 sufficiently big and all $n\ge n_0$, if T satisfy (5.14), then*

$$p_{\gamma,\tilde{C}_n}(C_n^{\Omega\setminus LC(T,n_0,n)}|I_n) < c_{n-1}^{1/100} \tag{5.26}$$

and if $\tau_n\in VG(T,n_0,n)$,

$$p_{\gamma,\tilde{C}_n}(C_n^{\Omega\setminus LC(T,n_0,n)}|I_n^{\tau_n}) < c_{n-1}^{1/100}. \tag{5.27}$$

Lemma 5.8. *For all n_0 is sufficiently big, for all $n\ge n_0+1$, for all T, if $\underline{d}\in LC(T,n_0,n)$, then for all $c_{n-1}^{-4/(n-1)}<k\le l_n(\underline{d})$, and for all $x\in C_n^{\underline{d}}$,*

$$\sum_{\substack{i\le k,\\ f^i(x)\in I^T_{n_0}}} r_{n_0}(f^i(x)) < 2c_{n_0-1}^{1/200}k \tag{5.28}$$

Proof. The proof follows closely the argument of Lemma 7.11 of [AM1], but since the claim is formally different, we will repeat some steps here, referring to the computations in [AM1].

Let $\underline{d}=(j_1,...,j_m)$. Assume that $k\le r_n(j_1)$. Since $j_1\in VG(T,n_0,n)$, we get the result as a consequence of Lemma 5.6. This will still work if we take $k\le r_n(j_1)+...+r_n(j_t)$, where j_i is very good for $1\le i\le t$.

Let m_k be the last return completed before k, that is $\sum_{i=1}^{m_k} r_n(j_i)\le k$. We must analyze the case where j_i is not very good for some $i\le m_k+1$. In this case, we must have, by LC1, $m_k\ge c_{n-1}^{-1/30}$. Let

$$\beta_k = \sum_{\substack{i\le m_k,\\ j_i\in VG(T,n_0,n)}} r_n(j_i). \tag{5.29}$$

After some computations, we get (see [AM1])

$$\sum_{\substack{i \leq m_k, \\ j_i \notin VG(T,n_0,n)}} r_n(j_i) \leq 4c_{n-1}^{1/35} k, \tag{5.30}$$

and

$$r_n(j_{m_k+1}) \leq c_{n-1}^{1/80} k, \tag{5.31}$$

(if $m_k = |\underline{d}|$, that is, $k = l_n(\underline{d})$, we will make the convention that $r_n(j_{m_k+1}) = 0$). We obtain

$$\sum_{\substack{i \leq k, \\ f^i(x) \in I_{n_0}^T}} r_{n_0}(f^i(x)) \leq c_{n_0-1}^{1/200} \beta_k + r_n(j_{m_k+1}) + \sum_{\substack{i \leq m_k, \\ j_i \notin VG(T,n_0,n)}} r_n(j_i) \leq 2c_{n_0-1}^{1/200} k. \tag{5.32}$$

□

5.3. **Large deviation estimate.**

5.3.1. *More precise estimates on return times.* We will need several times the following elementary result.

Lemma 5.9. *Let $m > 100$. If $q \in [0,1]$ and $\epsilon \geq m^{-1/4}$ then*

$$\sum_{|\frac{k}{m} - q| > \epsilon} \binom{m}{k} q^k (1-q)^{m-k} \leq e^{-m^{1/4}}. \tag{5.33}$$

Proof. Let $y_k = \binom{m}{k} q^k (1-q)^k$, and recall that $y_k \leq 1$ for all k (since $\sum y_k = 1$). It is enough to show that $\sum_{k/m > q+\epsilon} y_k \leq e^{-m^{1/4}}/2$, since the estimate corresponding to $\frac{k}{m} < q - \epsilon$ reduces to this one after interchanging q and $1-q$. Let $x_k = \frac{y_{k+1}}{y_k} = \frac{1-k/m}{(k+1)/m} \cdot \frac{q}{1-q}$. If $\frac{k}{m} \geq q + \frac{\epsilon}{2}$ then $x_k < \frac{1-q-\epsilon/2}{q+\epsilon/2} \cdot \frac{q}{1-q} < 1 - \frac{\epsilon}{2}$. Notice that if k is minimal with $\frac{k}{m} > q - \epsilon$ then there are about $\frac{\epsilon m}{2}$ integers $j < k$ such that $q + \frac{\epsilon}{2} \leq \frac{j}{m}$. We conclude that

$$\sum_{\frac{k}{m} > q+\epsilon} y_k \leq \sum_{i \geq \frac{\epsilon m}{2}} \left(1 - \frac{\epsilon}{2}\right)^i \leq \frac{2}{\epsilon} e^{-m\epsilon^2/4} \leq \frac{e^{m^{1/4}}}{2}, \tag{5.34}$$

and the result follows. □

Notation warning: In what follows, we will work with a fixed typical unimodal map f. We will use $\delta_1^{(n)}, ..., \delta_{14}^{(n)}$ to denote several small constants (going to 0 with n). We shall always choose $\delta_{i+1}^{(n)}$ after fixing $\delta_i^{(n)}$, and satisfying (among other requirements) $\delta_{i+1}^{(n)} \geq 10\delta_i^{(n)}$. We shall also take $\delta_1^{(n)} > n^{-1}$.

5.3.1.1. Let $\underline{d}_{n-1}$ be such that $R_{n-1}(0) \in C_{n-1}^{\underline{d}_{n-1}}$, and let B_n be the set of all j such that $R_{n-1}(I_n^j) = C_{n-1}^{\underline{d}}$, where $|\underline{d}| < |\underline{d}_{n-1}|$ and $\underline{d} = (j_1, ..., j_k)$ is obtained by considering the first k entries of $\underline{d}_{n-1}$. Let $A_n = \mathbb{Z} \setminus (\{0\} \cup B_n)$. Recall the estimates of §4.4.1. One readily sees that $I(B_n, n) \leq 2^{-n} c_{n-1}^{1/2}$ and for $j \in B_n$, $r_n(j) \leq v_n$. Notice that if $j \in A_n$, the interval I_n^j is far from the critical point in the sense that $c_{n-1}^{-1} |I_n^j|$ is much smaller than the distance from I_n^j to 0. It follows that, for any $\underline{d}$ whose last entry belongs to A_n, $\mathrm{dist}(R_n^{\underline{d}}) < 1 + \delta_1^{(n)} c_{n-1}^{1/2}$. If the last entry belongs to B_n we will use the general estimate $\mathrm{dist}(R_n^{\underline{d}}) \leq n^{2/3}$.

5.3.1.2. Let $m(\underline{d})$ be the number of entries of $\underline{d}$ which belong to A_n. The following easy estimates follow from the previous discussion by induction:

(5.35)
$$(1-(1+\delta_2^{(n)}c_{n-1}^{1/2})c_n)^m \leq I(\{m(\underline{d})=m\},n) \qquad \leq (1+\delta_2^{(n)}c_{n-1}^{1/2})(1-(1-\delta_2^{(n)}c_{n-1}^{1/2})c_n)^m,$$

(5.36)
$$c_n(1-(1+\delta_2^{(n)}c_{n-1}^{1/2})c_n)^m \leq C(\{m(\underline{d})=m\},n) \qquad \leq (1+\delta_2^{(n)}c_{n-1}^{1/2})c_n(1-(1-\delta_2^{(n)}c_{n-1}^{1/2})c_n)^m.$$

Let $Q(m',m) \subset \Omega$ be the set of $\underline{d}$ of size m' and with at least m entries on B_n, that is, $Q(m',m) = \{\underline{d} \in \Omega,\ |\underline{d}| = m',\ |\underline{d}| - m(\underline{d})| \geq m\}$. Let $q(m',m) = I(Q(m',m),n)$. From the definition and the estimates on distortion we have

$$\begin{aligned} q(m',m) &\leq q(m'-1,m) + \delta_3^{(n)}c_{n-1}^{1/2}(q(m'-1,m-1)-q(m'-1,m)) \\ &= (1-\delta_3^{(n)}c_{n-1}^{1/2})q(m'-1,m) + \delta_3^{(n)}c_{n-1}^{1/2}q(m'-1,m-1), \end{aligned} \tag{5.37}$$

which implies by induction,

$$q(m',m) \leq \sum_{k=m}^{m'} \binom{m'}{k} (\delta_3^{(n)}c_{n-1}^{1/2})^k (1-\delta_3^{(n)}c_{n-1}^{1/2})^{m'-k}. \tag{5.38}$$

Let us compute a few consequences of those estimates. Let H be the set of all $\underline{d} \in \Omega$ such that *at least one* of the following holds:

(H1) $|\underline{d}| \geq c_n^{-1/n}$ and $|\underline{d}| - m(\underline{d}) \geq 2c_{n-1}^{1/2}|\underline{d}|$,

(H2) $\underline{d}$ has some entry j_i satisfying $r_n(j_i) \geq c_{n-1}^{-14}$.

Using the present discussion to estimate (H1) and (4.7) to estimate (H2) we get

$$C(H,n) \leq I(H,n) \leq e^{-c_n^{-1/(8n)}} + e^{-c_{n-1}^{-19/2}} \leq e^{-c_{n-1}^{-9}}. \tag{5.39}$$

Let V be the set of $\underline{d}$ with $m(\underline{d}) \leq c_n^{-1/n}$. The present discussion gives

$$C(V,n) \leq 2c_n^{1-1/n}. \tag{5.40}$$

5.3.1.3. We will also need the following easy estimate:

Lemma 5.10. *Fix $P \subset A_n$, and let $p = I(P,n)$. Let $P(m,r) \subset \Omega$ be the set of all $\underline{d}$ with $m(\underline{d}) = m$ and with exactly r entries in P. Let $\overline{P}(m,r)$ (respectively $\underline{P}(m,r)$) denote the union of all $P(m,r)$ with $r' \geq r$ (respectively $r' \leq r$). Let $p(m,r) = I(P(m,r),n)$, $\overline{p}(m,r) = I(\overline{P}(m,r),n)$ and $\underline{p}(m,r) = I(\underline{P}(m,r),n)$.*

We have, with $\overline{p} = p(1+4\delta_1^{(n)}c_{n-1}^{1/2})$ and $\underline{p} = p(1-4\delta_1^{(n)}c_{n-1}^{1/2})$

$$\overline{p}(m,r) \leq (1-\overline{p})\overline{p}(m-1,r) + \overline{p} \cdot \overline{p}(m-1,r-1) \tag{5.41}$$

$$\underline{p}(m,r) \leq (1-\underline{p})\underline{p}(m-1,r) + \underline{p} \cdot \underline{p}(m-1,r-1) \tag{5.42}$$

$$\overline{p}(m,r) \leq (1+2\delta_2^{(n)}c_{n-1}^{1/2}) \sum_{k=r}^{m} \binom{m}{k} \overline{p}^k (1-\overline{p})^{m-k} \tag{5.43}$$

$$\underline{p}(m,r) \leq (1+2\delta_2^{(n)}c_{n-1}^{1/2}) \sum_{k=0}^{r} \binom{m}{k} \underline{p}^k (1-\underline{p})^{m-k} \tag{5.44}$$

Proof. We notice that $p(1,0) \leq 1-p$, $p \leq p(1,1) \leq p(1+2\delta_1^{(n)}c_{n-1}^{1/2})$. Let us consider a connected component E of $I_n^{\overline{P}(m,r)}$. It is either contained in a connected component of $I_n^{\overline{P}(m-1,r)}$ or it is contained in a component $\hat{E}$ of $I_n^{P(m-1,r-1)}$. In this last case, the iterate of R_n which takes $\hat{E}$ to

I_n (necessarily with distortion bounded by $1+\delta_1^{(n)} c_{n-1}^{1/2}$) must take E to a component of $I_n^{P(1,1)}$. It follows that

$$\overline{p}(m,r) \le \overline{p}(m-1,r) + (1+\delta_1^{(n)} c_{n-1}^{1/2}) p(m-1,r-1) p(1,1). \tag{5.45}$$

Since $p(m-1,r-1) = \overline{p}(m-1,r-1) - \overline{p}(m-1,r)$, we get (5.41), and (5.43) follows by induction.

Let us now consider a connected component E of $I_n^{\underline{P}(m,r)}$. It is either contained in a connected component of $I_n^{\underline{P}(m-1,r-1)}$ or it is contained in a component $\hat{E}$ of $I_n^{P(m-1,r)}$. In this last case, the iterate of R_n which takes $\hat{E}$ to I_n (necessarily with distortion bounded by $1+\delta_1^{(n)} c_{n-1}^{1/2}$) must take E to a component of $I_n \setminus I_n^{P(1,1)}$. It follows that

$$\underline{p}(m,r) \le \underline{p}(m-1,r-1) + p(m-1,r)(1-(1-\delta_1^{(n)} c_{n-1}^{1/2}) p(1,1)). \tag{5.46}$$

Since $p(m-1,r) = \underline{p}(m-1,r) - \underline{p}(m-1,r-1)$, we get (5.42), so (5.44) follows by induction. $\square$

5.3.2. *Return times.* Let us fix $\Theta \subset \mathbb{Z} \setminus \{0\}$, $\theta = I(\Theta, n)$. We would like to estimate

$$\zeta = \sum_{j \in \Theta} r_n(j) I(j,n) \tag{5.47}$$

in terms of θ (specially for the case $\Theta = \mathbb{Z} \setminus \{0\}$). In order to do so, it is convenient to write $\zeta = \zeta^A + \zeta^B$, where

$$\zeta^A = \sum_{j \in \Theta \cap A_n} r_n(j) I(j,n), \quad \zeta^B = \sum_{j \in \Theta \cap B_n} r_n(j) I(j,n). \tag{5.48}$$

Notice that it is easy to estimate (using §5.3.1.1)

$$\zeta^B \le v_n I(\Theta \cap B_n, n) \le c_{n-1}^{-1-\delta_4^{(n)}} \min\{\theta, 2^{-n} c_{n-1}^{1/2}\}. \tag{5.49}$$

To estimate ζ^A, we will consider the level sets $M_s = \{j \in A_n \cap \Theta | \, r_n(j) = s\}$, so that $\zeta^A = \sum s m_s$, where $m_s = I(M_s, n)$. Let $L = \{s | m_s \ge c_n^{1/(8n)}\}$, $S = \{s | m_s < c_n^{1/(8n)}\}$. Define

$$\zeta^L = \sum_{s \in L} s m_s, \quad \zeta^S = \sum_{s \in S} s m_s, \tag{5.50}$$

so that $\zeta^A = \zeta^L + \zeta^S$. Notice that by (4.7),

$$\zeta^S = \sum_{\substack{s \in S, \\ s \le c_n^{-1/(32n)}}} s m_s + \sum_{\substack{s \in S, \\ s > c_n^{-1/(32n)}}} s m_s \le c_n^{1/(16n)} + \sum_{t \ge c_n^{-1/(32n)}} t e^{-c_{n-1}^{-4} t} \le c_n^{1/(32n)}. \tag{5.51}$$

5.3.2.1. Let N be the set of all $\underline{d} \in \Omega$ such that $m(\underline{d}) \ge c_n^{-1/n}$ and *at least one* of the following holds:

(N1) For some $s \in L$, the number u of entries j_i of $\underline{d}$ belonging to M_s satisfies either $\frac{u}{m} > (1+4\delta_1^{(n)} c_{n-1}^{1/2}) m_s + c_n^{1/(8n)}$, or $\frac{u}{m} < (1-4\delta_1^{(n)} c_{n-1}^{1/2}) m_s - c_n^{1/(8n)}$,

(N2) For some $s \in S$, the number u of entries j_i of $\underline{d}$ belonging to M_s satisfies $\frac{u}{m} \ge 2 c_n^{1/(8n)}$.

It follows from Lemmas 5.9 and 5.10 that

$$I(N,n) \le 2e^{-c_n^{-1/(10n)}}. \tag{5.52}$$

Let $D = N \cup H \cup V$ and $\hat{D} = N \cup H$. By (5.52), (5.39) and (5.40) we have

$$C(D,n) \le c_n^{1-2/n} \tag{5.53}$$

$$C(\hat{D},n) \le e^{-c_{n-1}^{-17/2}}. \tag{5.54}$$

If $\underline{d} \notin D$, we have

$$\frac{1}{m(\underline{d})} \sum_{j_i \in \Theta} r_n(j_i) \geq (1 - \delta_5^{(n)} c_{n-1}^{1/2}) \zeta^L, \tag{5.55}$$

$$\frac{1}{m(\underline{d})} \sum_{j_i \in \Theta} r_n(j_i) \leq (1 + \delta_5^{(n)} c_{n-1}^{1/2}) \zeta^L + (|\underline{d}| - m(\underline{d})) v_n + 2 c_{n-1}^{-28} c_n^{1/(8n)} \tag{5.56}$$

$$\leq (1 + \delta_5^{(n)} c_{n-1}^{1/2}) \zeta^L + c_{n-1}^{-\frac{1}{2} - \delta_5^{(n)}}.$$

while, if $\underline{d} \notin \hat{D}$, we have either $\underline{d} \notin V$ (in which case (5.55) and (5.56) hold) or $\underline{d} \in V$ in which case we have $l_n(\underline{d}) \leq c_n^{-2/n}$.

Notice that D, ζ and ζ^L depend on Θ (and on n). If needed we will stress this dependence by writing $D(\Theta)$, $\zeta(\Theta)$ and $\zeta^L(\Theta)$.

5.3.2.2. Let

$$\alpha_n = \zeta(\mathbb{Z} \setminus \{0\}) = \sum_{j \neq 0} r_n(j) I(j, n). \tag{5.57}$$

Notice that due to (4.8),

$$\alpha_n > c_{n-1}^{-1 + \delta_6^{(n)}}. \tag{5.58}$$

Lemma 5.11. *We have*

$$\left| \frac{\alpha_n}{\alpha_{n-1} c_{n-1}^{-1}} - 1 \right| < c_{n-2}^{1/30}, \tag{5.59}$$

and for any set $\Theta \subset \mathbb{Z} \setminus \{0\}$ with $\theta = I(\Theta, n)$, we have

$$\zeta(\Theta) = \sum_{j \in \Theta} r_n(j) I(j, n) \leq (3\theta(1 - \ln \theta) + c_{n-1}) \alpha_n \tag{5.60}$$

Proof. Letting $\Theta = \mathbb{Z} \setminus \{0\}$ and keeping the previous notation, we have clearly

$$\zeta^L \leq \alpha_n = \zeta^L + \zeta^B + \zeta^S \leq \zeta^L + v_n I(B_n, n) + c_n^{1/(32n)} \leq \zeta^L + c_{n-1}^{-\frac{1}{2} - \delta_7^{(n)}}, \tag{5.61}$$

and since $\alpha_n \geq c_{n-1}^{-1 + \delta_6^{(n)}}$ by (5.58), we actually have

$$1 \leq \frac{\alpha_n}{\zeta^L} \leq 1 + c_{n-1}^{\frac{1}{2} - \delta_8^{(n)}}. \tag{5.62}$$

The previous discussion in §5.3.2.1 gives for $\underline{d} \notin D$,

$$(1 - c_{n-1}^{\frac{1}{2} - \delta_9^{(n)}}) \alpha_n \leq (1 - \delta_5^{(n)} c_{n-1}^{1/2}) \zeta^L \leq \frac{1}{m(\underline{d})} \sum r_n(j_i) = \frac{l_n(\underline{d})}{m(\underline{d})} \leq (1 + c_{n-1}^{\frac{1}{2} - \delta_9^{(n)}}) \alpha_n. \tag{5.63}$$

Using the estimate (5.36) on the distribution of $m(\underline{d})$, we get

$$(1 - c_{n-1}^{\frac{1}{2} - \delta_{10}^{(n)}}) c_n^{-1} \alpha_n \leq \sum_{\underline{d} \notin D} l_n(\underline{d}) C(\underline{d}, n) \tag{5.64}$$

which implies that for each $j \in A_n$ we have

$$I(j, n)(1 - c_{n-1}^{\frac{1}{2} - \delta_{11}^{(n)}}) c_n^{-1} \alpha_n \leq \sum_{C_n^{\underline{d}} \subset I_n^j} l_n(\underline{d}) C(\underline{d}, n). \tag{5.65}$$

Let us now consider the set $Z \subset A_n$ of all j such that $R_n(I_n^0)$ contains I_n^j, $r_n(j) < c_{n-1}^{-14}$, and such that $R_n(0)$ is at least $c_{n-1}^{1/4}|I_n|$ away from I_n^j. Let $\hat{Z}$ denote the set of $j \in \mathbb{Z}$ such that $R_n(I_{n+1}^j) \subset I_n^Z$. Then $I(\mathbb{Z} \setminus \hat{Z}, n+1) < c_{n-1}^{1/9}$. Since $I(j,n) \leq \delta_{12}^{(n)} c_{n-1}^{1/2}$ for all j, the distortion of $(R_n|I_n^0)^{-1}$ restricted to any component of I_n^Z is bounded by $1 + \delta_{14}^{(n)} c_{n-1}^{1/4}$. We conclude

$$(1 - c_{n-1}^{1/10})\alpha_n c_n^{-1} \leq \sum_{j \in \hat{Z}} r_{n+1}(j) I(j, n+1) \leq \alpha_{n+1}. \tag{5.66}$$

Let X_t be the set of $\underline{d}$ with $l_n(\underline{d}) \geq t(1 + c_{n-1}^{\frac{1}{2} - 10\delta_{14}^{(n)}}) c_n^{-1} \alpha_n$. Notice that

$$t \geq c_n^{1-2/n} \implies X_t \cap D = X_t \cap \hat{D}. \tag{5.67}$$

On the other hand, by (5.63),

$$\underline{d} \in X_t \setminus D \implies m(\underline{d}) \geq t(1 + c_{n-1}^{\frac{1}{2} - 9\delta_{14}^{(n)}}) c_n^{-1}, \tag{5.68}$$

so, by (5.36), $C(X_t \setminus D, n) \leq (1 - c_{n-1}^{\frac{1}{2} - 8\delta_{14}^{(n)}}) e^{-t}$, which gives by (5.54)

$$C(X_t, n) \leq (1 - c_{n-1}^{\frac{1}{2} - 8\delta_{14}^{(n)}})(e^{-t} + e^{-c_{n-1}^{-25/3}}), \quad t \geq c_n^{1-2/n}. \tag{5.69}$$

If $j \in Z$, we can estimate

$$C(X_t \cap \{\underline{d} \in \Omega,\, C_n^{\underline{d}} \subset I_n^j\}, n) \leq I(j,n)(1 - c_{n-1}^{\frac{1}{2} - 5\delta_{14}^{(n)}})(e^{-t} + e^{-c_{n-1}^{-25/3}}), \quad t \geq c_n^{1-2/n}, j \in Z. \tag{5.70}$$

Let Y_t be the set of $j \neq 0$ with $R_n(I_{n+1}^j) = C_n^{\underline{d}}$, $\underline{d} \in X_t$. The following estimates are immediate from (5.69), (5.70):

$$I(Y_t, n+1) \leq 2^n (e^{-t} + e^{-c_{n-1}^{-8}})^{1/2}, \qquad t \geq c_n^{1-2/n} \tag{5.71}$$

$$I(Y_t \cap \hat{Z}, n+1) \leq e^{-t} + e^{-c_{n-1}^{-8}}, \qquad t \geq c_n^{1-2/n}. \tag{5.72}$$

This last estimate implies in particular

$$I(Y_t, n+1) \leq e^{-t} + e^{-c_{n-1}^{-8}} + c_{n-1}^{1/9} \leq (1 + c_{n-1}^{1/20}) e^{-t}, \quad c_n^{1-2/n} \leq t \leq \ln c_{n-1}^{-1/20}. \tag{5.73}$$

Using additionally that by (4.7), $I(Y_t, n+1) \leq e^{-tc_n^4}$ for $t \geq c_n^{-4}$, and that obviously $I(Y_t, n+1) \leq 1$ for all t, we see that (5.71), (5.73) imply

$$Y_t \leq s(t) = \begin{cases} 1 & \text{for } t < c_n^{1-2/n}, \\ (1 + c_{n-1}^{1/20}) e^{-t} & \text{for } c_n^{1-2/n} \leq t < \ln c_{n-1}^{-1/20}, \\ 2^{n+1} e^{-t/2} & \text{for } \ln c_{n-1}^{-1/20} \leq t < c_{n-1}^{-8}, \\ 2^{n+1} e^{-c_{n-1}^{-8}/2} & \text{for } c_{n-1}^{-8} \leq t < c_n^{-5}, \\ e^{-tc_n^4} & \text{for } t \geq c_n^{-5}. \end{cases} \tag{5.74}$$

which gives

$$\int_0^\infty I(Y_t, n+1) dt \leq 1 + c_{n-1}^{1/20}. \tag{5.75}$$

By definition of X_t and Y_t, we have

$$0 \neq j \in Y_t \iff r_{n+1}(j) \geq v_n + t(1 + c_{n-1}^{\frac{1}{2} - 10\delta_{14}^{(n)}}) c_n^{-1} \alpha_n, \tag{5.76}$$

so that (5.75) implies

$$\alpha_{n+1} \leq v_n(1-c_{n+1}) + (1+c_{n-1}^{\frac{1}{2}-10\delta_{14}^{(n)}})\alpha_n c_n^{-1} \int_0^\infty I(Y_t, n+1) \leq (1+c_{n-1}^{1/30})\alpha_n c_n^{-1}. \tag{5.77}$$

Estimates (5.66) and (5.77) imply (5.59), shifting n to $n+1$.

Moreover, for any set $\Theta \subset \mathbb{Z} \setminus \{0\}$, with $\theta = I(\Theta, n+1)$, (5.74) implies

$$\int_0^\infty I(Y_t \cap \Theta, n+1)dt \leq \int_0^\infty \min\{\theta, s(t)\}dt \leq \frac{5}{2}\theta(1-\ln\theta) + \frac{c_n}{2}, \tag{5.78}$$

which (together with (5.76) and (5.66)) implies (5.60), shifting n to $n+1$. □

We can now conclude:

Lemma 5.12 (Large Deviation Estimate). *Let $\Theta \subset \mathbb{Z} \setminus \{0\}$, $\theta = I(\Theta, n)$. Let F be the set of $\underline{d}$ such that*

$$\frac{1}{l_n(\underline{d})} \sum_{j_i \in \Theta} r_n(j_i) \geq 4(\theta(1-\ln\theta) + c_{n-1}^{1/4}). \tag{5.79}$$

Then $C(F, n) \leq 2c_n^{1-2/n}$.

Proof. By the previous considerations §5.3.2.1, except for $\underline{d}$ in an exceptional set $D(\Theta)$ satisfying $C(D(\Theta), n) \leq c_n^{1-2/n}$, (5.56) holds, that is

$$\frac{1}{m(\underline{d})} \sum_{j_i \in \Theta} r_n(j_i) \leq (1+\delta_5^{(n)} c_{n-1}^{1/2})\zeta^L(\Theta) + c_{n-1}^{-\frac{1}{2}-\delta_5^{(n)}}, \tag{5.80}$$

where, by Lemma 5.11,

$$\zeta^L(\Theta) \leq (3\theta(1-\ln\theta) + c_{n-1})\alpha_n. \tag{5.81}$$

By the proof of Lemma 5.11, except for $\underline{d}$ in an exceptional set $D(\mathbb{Z} \setminus \{0\}) \subset \Omega$ satisfying $C(D(\mathbb{Z} \setminus \{0\}), n) \leq c_n^{1-2/n}$, (5.63) holds, that is

$$\frac{l_n(\underline{d})}{m(\underline{d})} \geq (1 - c_{n-1}^{\frac{1}{2}-\delta_9^{(n)}})\alpha_n. \tag{5.82}$$

Estimates (5.80) and (5.82) imply that for $\underline{d} \notin D(\Theta) \cup D(\mathbb{Z} \setminus \{0\})$,

$$\frac{1}{l_n(\underline{d})} \sum_{j_i \in \Theta} r_n(j_i) \leq \frac{(1+\delta_5^{(n)} c_{n-1}^{1/2})\zeta^L(\Theta) + c_{n-1}^{-\frac{1}{2}-\delta_5^{(n)}}}{(1-c_{n-1}^{\frac{1}{2}-\delta_9^{(n)}})\alpha_n} \leq 4\theta(1-\ln\theta) + c_{n-1}^{1/3}. \tag{5.83}$$

thus $F \subset D(\Theta) \cup D(\mathbb{Z} \setminus \{0\})$. The result follows. □

5.4. **Proof of Theorem B.**

5.4.1. *Series of reductions.* We will argue by contradiction. If Theorem B is false, there exists a positive measure set $\mathcal{J}_1$ of non-regular parameters λ such that the critical point is not in the basin of the physical measure μ_{f_λ}. Since almost all parameters in $\mathcal{J}_1$ are finitely renormalizable, there exists a subset $\mathcal{J}_2 \subset \mathcal{J}_1$ of positive measure of parameters which are exactly k times renormalizable, with some fixed k.

For each parameter λ in $\mathcal{J}_2$, let us consider the sequence of partitions Υ_n of the interval I in connected components of the domain of the first landing map from I to I_n. Those partitions get more refined as n increases, the size of the largest component (of order at most c_{n-1}) decreasing to 0 with n. Thus, there exists some $\eta > 0$ and a positive measure set of parameters $\mathcal{J}_3 \subset \mathcal{J}_2$, such that for all parameters in $\mathcal{J}_3$ there exists at least one component $\xi^\lambda \in \Upsilon_\eta$ (that may be chosen to depend measurably on λ) such that the asymptotic frequency of the critical orbit in ξ^λ either does not exists or is different from $\mu_{f_\lambda}(\xi^\lambda)$. Proceeding further, there exists $\epsilon > 0$ and a positive measure set $\mathcal{J}_4 \subset \mathcal{J}_3$ such that for all parameters in $\mathcal{J}_4$.

$$\limsup \left| \frac{1}{k} \#\{i \le k, f_\lambda^i(0) \in \xi^\lambda\} - \mu_{f_\lambda}(\xi^\lambda) \right| > \epsilon. \tag{5.84}$$

The set $\mathcal{J}_4$ is contained in the union of parameter intervals J_η (η fixed) associated to the principal nest (of k-th renormalization). It follows that at least one such interval J_η intersects $\mathcal{J}_4$ in a positive measure set $\mathcal{J}_5$. For any $\lambda_1, \lambda_2 \in J_\eta$, there is a homeomorphism $h[\lambda_1, \lambda_2] : I \to I$ such that $h[\lambda_1, \lambda_2] \circ f_{\lambda_1} | (I \setminus I_{\eta+1}[\lambda_1]) = f_{\lambda_2} \circ h[\lambda_1, \lambda_2]$. Thus, there exists a positive measure subset $\mathcal{J} \subset \mathcal{J}_5$ such that for $\lambda_1, \lambda_2 \in \mathcal{J}$, $h[\lambda_1, \lambda_2]$ takes ξ^{λ_1} to ξ^{λ_2}. In other words, the combinatorics of ξ^λ does not depend on $\lambda \in \mathcal{J}$.

In order to get a contradiction and prove Theorem B, we will show that for almost every parameter in $\mathcal{J}$,

$$\limsup \left| \frac{1}{k} \#\{i \le k, f_\lambda^i(0) \in \xi^\lambda\} - \mu_{f_\lambda}(\xi^\lambda) \right| < \epsilon. \tag{5.85}$$

To simplify the notation, we will write ξ for ξ^λ. We will also write μ for μ_{f_λ}. For $x \in I$, and a measurable set $\Lambda \subset I$, let

$$\Psi(\Lambda, x, k) = \frac{1}{k} \#\{i \le k, f^i(x) \in \Lambda\}. \tag{5.86}$$

Notice that if $\Lambda = \xi$, and $n > \eta$ then $x \mapsto \Psi(\Lambda, x, k)$ is constant in each interval I_n^j for $k \le r_n(x)$, while for $k \le l_n(x)$, $x \mapsto \Psi(\Lambda, x, k)$ is constant in each $C_n^{\underline{d}}$. Those quantities stay unchanged if we vary the parameter λ inside some J_n, if we keep the combinatorics constant, that is, if we choose a varying point x_λ inside $I_n^j[\lambda]$ or $C_n^{\underline{d}}[\lambda]$, j or $\underline{d}$ fixed.

5.4.2. *Computing μ in the principal nest.* For $x \in I$, let $\varsigma_n(x) = \inf\{k,\ f^k(x) \in I_n\}$, so that $f^{\varsigma_n(x)}(x)$ is the first landing of x in I_n.

For $x \in I_n$, $\varsigma_n(x) = 0$, and in general we have $\varsigma_{n+1}(x) - \varsigma_n(x) = l_n(f^{\varsigma_n}(x))$. Notice that Lemma 5.3 implies that

$$\frac{|\{x \in I_n, c_n^{-1/2} < l_n(x) < c_n^{-2}\}|}{|I_n|} < c_n^{1/3}. \tag{5.87}$$

Since each branch of the first landing map from I to I_n has distortion bounded by $1 + O(c_n)$ (see [ALM], Theorem 2.14), we obtain the estimate

$$\frac{|\{x \in I, c_n^{-1/2} < l_n(x) < c_n^{-2}\}|}{|I|} < 2c_n^{1/3}. \tag{5.88}$$

By Borel-Cantelli, for almost every x, for n sufficiently big,

$$c_n^{-1/2} < \varsigma_{n+1}(x) - \varsigma_n(x) < c_n^{-2}. \tag{5.89}$$

In particular

$$\lim \frac{\varsigma_{n+1}(x)}{\varsigma_n(x)} = \infty, \quad \text{for almost every } x \in I. \tag{5.90}$$

Thus, for all $\Lambda \subset I$ measurable, for almost every $x \in I$,

$$\lim \Psi(\Lambda, f^{\varsigma_n(x)}(x), \varsigma_{n+1}(x) - \varsigma_n(x)) = \mu(\Lambda). \tag{5.91}$$

Given a measurable subset $\Lambda \subset I$, we let $M_l(\Lambda, n, \delta) \subset I_n$ be the set of all x such that

$$|\Psi(\Lambda, x, l_n(x)) - \mu(\Lambda)| > \delta. \tag{5.92}$$

We let $M_r(\Lambda, n, \delta) \subset I_n$ be the set of all x such that

$$|\Psi(\Lambda, x, r_n(x)) - \mu(\Lambda)| > \delta. \tag{5.93}$$

Lemma 5.13. *For any measurable set $\Lambda \subset I$, for any $\delta > 0$,*

$$\lim \frac{|M_l(\Lambda, n, \delta)|}{|I_n|} = 0, \tag{5.94}$$

$$\lim \frac{|M_r(\Lambda, n, \delta)|}{|I_n|} = 0. \tag{5.95}$$

Proof. Let H_n be the set of $x \in I$, such that the first landing of x on I_n belongs to $M_l(\Lambda, n, \delta)$. If (5.94) is not true, using the small distortion of the first landing map we conclude that $\limsup |H_n| > 0$, so there exists a positive measure set of x which belong to infinitely many H_n. But this is incompatible with (5.91).

Let $T_n \subset I_n$ be the union of I_n^j with the following properties:

$$I_n^j \subset R_n(I_{n+1}), \tag{5.96}$$

$$\operatorname{dist}((R_n|I_{n+1})^{-1}|I_n^j) < 2, \tag{5.97}$$

$$\operatorname{dist}(R_n|I_n^j) < 2, \tag{5.98}$$

$$r_n(j) < c_{n-1}^{-14}. \tag{5.99}$$

It follows that

$$1 - \frac{|(R_n|I_{n+1})^{-1}(T_n)|}{|I_{n+1}|} < c_{n-1}^{1/10}. \tag{5.100}$$

Let

$$Y_{n+1} = \{x \in (R_n|I_{n+1})^{-1}(T_n), R_n^2(x) \in M_l(\Lambda, n, \delta/2), l_n(R_n^2(x)) > c_n^{-1/2}\}. \tag{5.101}$$

Then

$$\frac{|Y_{n+1}|}{|(R_n|I_{n+1})^{-1}(T_n)|} \leq 4\left(\frac{|M_l(\Lambda, n, \delta/2)|}{|I_n|} + c_n^{1/3}\right), \tag{5.102}$$

thus

$$1 - \frac{|Y_{n+1}|}{|I_{n+1}|} \leq 4\left(\frac{|M_l(\Lambda, n, \delta/2)|}{|I_n|} + c_n^{1/3} + c_{n-1}^{1/10}\right), \tag{5.103}$$

so that, by (5.94), $\lim \frac{|Y_{n+1}|}{|I_{n+1}|} = 1$. On the other hand, if $x \in Y_{n+1}$,

$$\begin{aligned} |\Psi(\Lambda, x, r_{n+1}(x)) - \mu(\Lambda)| &\leq \frac{1}{r_{n+1}(x)}\left(\frac{\delta}{2} l_n(R_n^2(x)) + r_{n+1}(x) - l_n(R_n^2(x))\right) \\ &\leq \left(\frac{\delta}{2} + \frac{2c_{n-1}^{-14}}{c_n^{-1/2}}\right), \end{aligned} \tag{5.104}$$

so that $Y_{n+1} \subset M_r(\Lambda, n, \delta)$ for n sufficiently big. $\square$

5.4.3. *Distribution of the critical orbit.*

Lemma 5.14. *Let $S \subset I_n$ be a union of I_n^j. For any $C \geq 1$:*

$$\lim_{\gamma \to 1} p_{\gamma,C}(S|I_n) = p_{1,C}(S|I_n). \tag{5.105}$$

Proof. If $X \subset I_n$ is any finite union of intervals, by compactness of quasisymmetric maps we get

$$\lim_{\gamma \to 1} p_{\gamma,C}(X|I_n) = p_{1,C}(X|I_n). \tag{5.106}$$

It is clear that for any $\gamma \geq 1$, $p_{\gamma,C}(S|I_n) \geq p_{1,C}(S|I_n)$. On the other hand, since $I_n \setminus \cup I_n^j$ is a regular Cantor set,

$$\lim_{k \to \infty} p_{2,C}(x \in I_n^j, |j| > k|I_n) = 0, \tag{5.107}$$

since the qs-capacity of gaps of generation t decays exponentially with t (see Lemma 6.1 of [AM1] for a related estimate).

Given $\delta > 0$, we can fix a subset $S' \subset S$ which is a union of finitely many I_n^j such that

$$p_{2,C}(S \setminus S'|I_n) < \delta. \tag{5.108}$$

Hence

$$\begin{aligned} \limsup_{\gamma \to 1} p_{\gamma,C}(S|I_n) &\leq \limsup_{\gamma \to 1} p_{\gamma,C}(S \setminus S'|I_n) + \limsup_{\gamma \to 1} p_{\gamma,C}(S'|I_n) \leq \delta + p_{1,C}(S'|I_n) \\ &\leq \delta + p_{1,C}(S|I_n). \end{aligned} \tag{5.109}$$

The result follows. □

We now specify this discussion to $\Lambda = \xi$ (the gap fixed at the beginning). We are now in situation to apply the Large Deviations Estimate to obtain:

Lemma 5.15. *For all $\delta > 0$, for all n_s sufficiently big, there exists $\gamma > 1$ such that*

$$p_{\gamma,10}(M_r(\xi, n_s, \delta)|I_{n_s}) < c_{n_s-1}^{1/20}, \tag{5.110}$$

Proof. Let n_0 be very big and $\delta' > 0$ be such that

$$4(\delta'(1 - \ln \delta') + c_{n_0-1}^{1/4}) < \frac{\delta}{2}, \tag{5.111}$$

and $c_{n_0}^{1/3} \ll \delta$.

Let $n > n_0$ be such that

$$\frac{|M_r(\xi, n, \delta/3)|}{|I_n|} < \delta'. \tag{5.112}$$

Notice that $M_r(\xi, n, \delta/3) = I_n^\Theta$ for some set $\Theta \subset \mathbb{Z}$, and $I(\Theta, n) < \delta'$. Let $F \subset \Omega$ be the set of $\underline{d} = (j_1, ..., j_m)$ such that

$$\frac{1}{l_n(\underline{d})} \sum_{j_i \in \Theta} r_n(j_i) > \frac{\delta}{2} > 4(\delta'(1 - \ln \delta') + c_{n-1}^{1/4}). \tag{5.113}$$

Then, by the Large Deviation Estimate (Lemma 5.12) we get $C(F, n) \leq c_n^{1-3/n}$. Let $F' = F \cup (\Omega \setminus LS(n))$. It follows that $C(F', n) \leq c_n^{2/7}$. Let $E' \subset \mathbb{Z}$ be the set of j such that $R_n(I_{n+1}^j) \subset C_n^{F'}$. Then $I(E', n+1) \leq c_n^{2/35}$.

Notice that if $x \in I_{n+1}^{\mathbb{Z} \setminus E'}$ and $\underline{d} = (j_1, ..., j_m)$ is such that $R_n(x) \in C_n^{\underline{d}}$, then

$$\frac{v_n}{r_{n+1}(x)} < c_n^{1/3} \ll \delta. \tag{5.114}$$

(since $\underline{d} \in LS(n)$ and $r_{n+1}(x) > |\underline{d}|$) so we can conclude

$$|\Psi(\xi, x, r_{n+1}(x)) - \mu(\xi)| \leq \frac{1}{r_{n+1}(x)} \left(v_n + \sum_{i \leq m} r_n(j_i) |\Psi(\xi, R_n^i(x), r_n(j_i)) - \mu(\xi)| \right) \tag{5.115}$$
$$\leq \frac{1}{r_{n+1}(x)} \left(v_n + \sum_{\substack{i \leq m, \\ j_i \in \Theta}} r_n(j_i) + \frac{\delta}{3} \sum_{\substack{i \leq m, \\ j_i \notin \Theta}} r_n(j_i) \right)$$
$$\leq \frac{1}{r_{n+1}(x)} \left(v_n + \left(\frac{\delta}{2} + \frac{\delta}{3} \right) l_n(\underline{d}) \right) < \delta.$$

So $I_{n+1}^{E'} \supset M_r(\xi, n+1, \delta)$. But $I(E', n+1) \leq c_n^{2/35}$ implies that

$$p_{1,10}(I_{n+1}^{E'} | I_{n+1}) < c_n^{2/39}, \tag{5.116}$$

so the result now follows by Lemma 5.14 with $n_s = n + 1$. □

Let us select $\delta = \epsilon/3$, and using the previous lemma, we select n_s very large and such that $c_{n_s-1}^{1/400} < \delta$. Let T be such that $I_{n_s}^T = M_r(\xi, n_s, \delta)$. Using Lemma 5.4 we get

$$p_{\gamma, C_n}(I_n^{\mathbb{Z} \setminus VG(T, n_s, n)} | I_n) \leq c_{n-1}^{1/20}. \tag{5.117}$$

Using PhPa2 we get:

Lemma 5.16. *For almost every parameter in $\mathcal{J}$, for all n sufficiently big, we have $\tau_n \in VG(T, n_s, n)$.*

Using Lemma 5.7, we get, for n sufficiently big,

$$p_{\gamma, \tilde{C}_n}(C_n^{\Omega \setminus LC(T, n_s, n)} | I_n^{\tau_n}) < c_{n-1}^{1/100}. \tag{5.118}$$

Using PhPa1 we get:

Lemma 5.17. *For almost every parameter in $\mathcal{J}$, for all n sufficiently big, $R_n(0) \in C_n^{\underline{d}}$ with $\underline{d} \in LC(T, n_s, n)$.*

Let us now consider a parameter which satisfies the conclusion of the two previous lemmas. Let us show that for k big enough,

$$|\Psi(\xi, 0, k) - \mu(\xi)| < 2\delta < \epsilon. \tag{5.119}$$

Indeed, if $v_n + c_{n-1}^{-4/(n-1)} < k \leq v_{n+1}$, by Lemma 5.8

$$|\Psi(\xi, f^{v_n}(0), k - v_n) - \mu(\xi)| < \delta + 2c_{n_s-1}^{1/200}, \tag{5.120}$$

in particular, for n big enough

$$|\Psi(\xi, 0, v_n) - \mu(\xi)| < 3\delta/2. \tag{5.121}$$

Notice that (5.120) and (5.121) imply (5.119) for n big enough and for $v_n + c_{n-1}^{-4/(n-1)} < k \leq v_{n+1}$. For $v_n \leq k \leq v_n + c_{n-1}^{-4/(n-1)}$, (5.119) follows from (5.121) since $v_n > c_{n-1}^{-1/2} \gg \delta^{-1} c_{n-1}^{-4/(n-1)}$ for n big enough.

Thus, for almost every parameter in $\mathcal{J}$, (5.119) holds, which contradicts (5.84) and completes the proof of Theorem B.

5.5. **Proof of Corollary 1.1.** We want to show that

$$\int \ln |Df| d\mu = \lim \frac{1}{k} \ln |Df^k(f(0))| = \lim \frac{1}{n} \sum_{k=1}^{n} \ln |Df(f^k(0))|. \tag{5.122}$$

The fact that 0 belongs to the basin of μ means that for all continuous ϕ,

$$\frac{1}{n} \sum_{k=1}^{n} \phi(f^k(0)) = \int \phi d\mu. \tag{5.123}$$

Since μ has no atoms, this formula still holds if ϕ is a bounded function with at most finitely many discontinuities. Unfortunately, $\ln |Df|$ is not bounded, so we only have, for every $\delta > 0$ small

$$\int_{I \setminus (-\delta,\delta)} \ln |Df| d\mu = \lim_{n \to \infty} \frac{1}{n} \sum_{\substack{1 \leq k \leq n \\ f^k(0) \in I \setminus (-\delta,\delta)}} \ln |Df(f^k(0))|. \tag{5.124}$$

Since

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{n} \sum_{\substack{1 \leq k \leq n \\ f^k(0) \in (-\delta,\delta)}} \ln |Df(f^k(0))| \leq 0, \tag{5.125}$$

we have to prove that for almost every non-regular parameter,

$$\lim_{\delta \to 0} \liminf_{n \to \infty} \frac{1}{n} \sum_{\substack{1 \leq k \leq n \\ f^k(0) \in (-\delta,\delta)}} \ln |Df(f^k(0))| = 0. \tag{5.126}$$

Condition (5.126) is called Weak Regularity by Tsujii. In [AM3], Theorem 10.2 (see also Remark 10.3 in that paper), it was shown that almost every non-regular parameter in non-trivial analytic families of unimodal maps satisfies (5.126), so, together with Theorem B, it implies Corollary 1.1.

6. Regularity of the physical measure

6.1. **Outline.** Theorem C is a statement of regularity of μ. We can think of $d\mu^K$ as a regularization of μ, designed to allow an understanding of the relation between ergodic and geometric properties of hyperbolic Cantor sets. Before tackling the problem of studying the regularization of μ, it is important to understand the limitations on the regularity of μ and identify the source of the difficulties.

According to Theorem 2.6, $d\mu$ is bounded from below on A (by some constant $C > 0$). As a consequence, if T is an interval of radius ϵ centered on 0 then $\mu(f(T)) = \mu(T) \geq 2C\epsilon \geq 2C\sqrt{|f(T)|}$. This shows that $d\mu$ has a "pole" at the critical value and, due to invariance of μ, there are also poles all along the orbit of the critical value.

In particular, for a general measurable set contained in A, the measure-theoretical quantity $\ln(\mu(A))$ only gives information about the geometric quantity $\ln |A|$ up to a factor of 2 (for $|A|$ small enough). This is the main reason why we have to introduce the regularization procedure. We would not be able to prove Theorem A just with general information on μ.

According to [MS], this estimate on the non-regularity of μ is optimal: it implies that $d\mu \notin L^2(I)$, but it is known that for maps satisfying the Collet-Eckmann condition $d\mu \in L^p(I)$, $p < 2$. This is better explained by Benedicks and Carleson [BC], who, for a smaller set of parameters (contained in the set of good Benedicks-Carleson parameters) described $d\mu$ as a sum of a bounded distribution and infinitely many poles (called square-root singularities by them) along the orbit of the critical value. Although this was not proved in general, this is the picture to imagine as a guide.

Since the critical orbit is distributed according to μ, those poles are everywhere (they are dense in the attractor). However, not all is lost:

(1) The strength of the poles decreases exponentially fast along the critical orbit (because of the Collet-Eckmann condition),
(2) The regularized $d\mu^K$ averages over the gap, and dissipate the pole with a strength proportional to the size of the gap.

Thus, a naive argument to prove Theorem C would be to obtain, with total probability, some "quantitative transversality" of the critical orbit with respect to K which would guarantee that strong poles are located in big gaps. For instance, we could expect that the time of the first visit of the critical point to some gap of K is inversely proportional to the size of the gap. Such a situation would imply that strong poles fall in (very) big gaps and should help[7] us to conclude that $d\mu^K \in L^p(I)$ for $1 \leq p < \infty$.

This would be much easier to deal with if the location of the successive poles was independent and uniformly distributed with respect to Lebesgue measure. However, there is quite a bit of interaction between different poles. In particular, new poles tend to show up more frequently near earlier poles than elsewhere (since the critical orbits distributes according to μ which in turn is more concentrated near the poles).

Our strategy will be to hierarchize the gaps according to the principal nest. To estimate the measure of a given gap, we will study their frequency in return branches. To estimate the possible increase in frequency between levels (caused by the distortion originated on the poles), we introduce a transversality condition (which we call "Strong poles fall in big gaps"), which means that $R_n(0)$ falls transversely enough with respect to the Cantor set of points that never land on I_n (the concept of transverse involves the hierarchy). This analysis (which will be carried out in the next section) will allow us to conclude the "Main estimate", which gives bounds on the μ-measure of gaps.

In this section we state the "Strong poles fall in big gaps" condition, prove that it is a total probability one, and conclude Theorem C assuming the Main estimate.

6.2. **The "Strong poles fall in big gaps condition".** We say that f satisfies the "Strong poles fall in big gaps" condition if

SP1: For all $\underline{d} \in \Omega$, $|\underline{d}| \geq 1$, the distance between $R_n(0)$ and $\partial I_n^{\underline{d}}$ is bounded by

$$\frac{|I_n^{\underline{d}}|}{2^n |\underline{d}|^2}, \tag{6.1}$$

SP2: $R_n(0) \in C_n^{\underline{d}}$, where $\underline{d} = (j_1, ..., j_m)$ satisfies

$$r_n(j_i) \leq c_{n-1}^{-11}, \quad 1 \leq i \leq m, \tag{6.2}$$

$$I(j_i, n) \geq e^{-c_{n-1}^{-12}}, \quad 1 \leq i \leq m, \tag{6.3}$$

SP3: $R_n(0) \in C_n^{\underline{d}}$, where $\underline{d} = (j_1, ..., j_m)$, and for each $1 \leq i \leq e^{c_{n-2}^{-4}}$ we have $R_{n-1}(I_n^{j_i}) \subset C_{n-1}^{\underline{d}_i}$ where $\underline{d}_i = (j_1^i, ..., j_{s(i)}^i)$ and

$$r_{n-1}(j_k^i) \leq c_{n-2}^{-11}, \quad 1 \leq k \leq s(i), \tag{6.4}$$

$$I(j_k^i, n-1) \geq e^{-c_{n-2}^{-12}}, \quad 1 \leq k \leq s(i). \tag{6.5}$$

Lemma 6.1. *Almost every non-regular parameter satisfies the "Strong poles fall in big gaps" condition.*

[7] One also needs to guarantee that strong poles fall well inside a gap in order to control the effect on small nearby gaps.

Proof. Let γ be such that $\epsilon(\gamma) < \delta_0$, in the notation of §4.4.2, with $\delta_0 > 0$ very small (say, 1/1000).

Let us first deal with SP1. We will consider two cases $|\underline{d}| = 1$ and $|\underline{d}| > 1$. Let $d(X, Y)$ denote the distance between X and Y.

In the first case, let A_n be the set of k such that

$$d(I_n^k, \partial I_n^j) \leq \frac{|I_n^j|}{2^{3n/4}}, \quad \text{for some } j \neq k. \tag{6.6}$$

Then

$$p_\gamma(I_n^{A_n} | I_n) < 2^{-n/2}. \tag{6.7}$$

Applying PhPa2, we see that with total probability, $\tau_n \notin A_n$ for n large enough. This implies that, with total probability, for n big, if

$$d(R_n(0), \partial I_n^j) \leq \frac{|I_n^j|}{2^n} \tag{6.8}$$

then $j = \tau_n$. Let B_n be the set of $\underline{d}$ such that $C_n^{\underline{d}} \subset I_n^{\tau_n}$ and

$$d(C_n^{\underline{d}}, \partial I_n^{\tau_n}) \leq \frac{|I_n^{\tau_n}|}{2^{3n/4}}. \tag{6.9}$$

Then

$$p_\gamma(C_n^{B_n} | I_n^{\tau_n}) < 2^{-n/2}, \tag{6.10}$$

and by PhPa1 we see that $R_n(0) \notin C_n^{B_n}$ for n large enough. In particular, we conclude the result for $\underline{d} = 1$.

In the second case, let $E(n)$ be the set of $\underline{d}$ such that there exists some $\tilde{\underline{d}}$ with $|\tilde{\underline{d}}| \geq 2$ and

$$d(C_n^{\underline{d}}, \partial I_n^{\tilde{\underline{d}}}) \leq \frac{|I_n^{\tilde{\underline{d}}}|}{|\tilde{\underline{d}}|^{3/2} 2^{3n/4}}. \tag{6.11}$$

Let us show that

$$p_\gamma(E(n) | I_n^{\tau_n}) \leq \frac{1}{2^{n/2}} \sum_{k \geq 2} k^{-4/3}. \tag{6.12}$$

Notice that if $I_n^{\tilde{\underline{d}}} \subset I_n^j$ with $j \neq \tau_n$, then no $C_n^{\underline{d}} \subset I_n^{\tau_n}$ satisfies (6.11), since

$$d(C_n^{\underline{d}}, \partial I_n^{\tilde{\underline{d}}}) \geq d(I_n^{\tilde{\underline{d}}}, \partial I_n^j) \gg |I_n^{\tilde{\underline{d}}}|. \tag{6.13}$$

On the other hand, for each $I_n^{\tilde{\underline{d}}} \subset I_n^{\tau_n}$, the set $E(\tilde{\underline{d}})$ of $\underline{d}$ satisfying (6.11) has the property that, for any h γ-qs,

$$\frac{|h(C_n^{E(\tilde{\underline{d}})})|}{|h(I_n^{\tilde{\underline{d}}})|} < \frac{1}{2^{n/2} |\tilde{\underline{d}}|^{4/3}}, \tag{6.14}$$

and since all $I_n^{\tilde{\underline{d}}}$ with $|\tilde{\underline{d}}| = k$ are disjoint, letting $E(k, n) = \cup_{|\tilde{\underline{d}}| = k} E(\tilde{\underline{d}})$, we get

$$p_\gamma(C_n^{E(k,n)} | I_n^{\tau_n}) \leq \frac{1}{2^{n/2} k^{4/3}}. \tag{6.15}$$

This implies (6.12).

Applying PhPa1, we see that with total probability, for n big enough, $R_n(0) \notin C_n^{E(n)}$, which gives SP1 for $|\underline{d}| > 1$.

Let us consider SP2. Notice that $r_n(j) < c_{n-1}^{-11}$ implies $I(j,n) > e^{c_{n-1}^{-12}}$ for n big, since the derivative of f is bounded. Let $F(n)$ be the set of j satisfying $r_n(j) > c_{n-1}^{-11}$, and $F'(n)$ be the set of $\underline{d}$ with at least one entry in $F(n)$. We get

$$p_\gamma(I_n^{F(n)}|I_n) \leq e^{-c_{n-1}^{-7}}, \tag{6.16}$$

which implies using PhPa2 that $\tau_n \notin F(n)$ with total probability, and thus

$$p_\gamma(C_n^{F'(n)}|I_n) \leq e^{-c_{n-1}^{-7+\delta}}, \tag{6.17}$$

$$p_\gamma(C_n^{F'(n)}|I_n^{\tau_n}) \leq e^{-c_{n-1}^{-7+\delta}}, \tag{6.18}$$

where δ goes to 0 when n grows. Using PhPa1 we get $R_n(0) \notin C_n^{F'(n)}$ with total probability, which implies SP2.

Let us consider SP3. Keeping the notation of the previous discussion, let $G(n+1)$ be the set of j such that $R_n(I_{n+1}^j) \subset C_n^{\underline{d}}$ with $\underline{d} \in F'(n)$. Let $G'(n+1)$ be the set of $\underline{d}$ with at least one entry in $G(n+1)$ among its first $e^{-c_{n-1}^{-4}}$ entries. It follows that

$$p_\gamma(I_{n+1}^{G(n+1)}|I_{n+1}) \leq e^{-c_{n-1}^{-7+\delta}}. \tag{6.19}$$

which by PhPa2 implies that $\tau_{n+1} \notin G(n+1)$ with total probability and thus

$$p_\gamma(C_{n+1}^{G'(n+1)}|I_{n+1}) \leq e^{-c_{n-1}^{-7+\delta}}, \tag{6.20}$$

$$p_\gamma(C_{n+1}^{G'(n+1)}|I_{n+1}^{\tau_{n+1}}) \leq e^{-c_{n-1}^{-7+\delta}}. \tag{6.21}$$

This implies, using PhPa1, that $R_{n+1}(0) \notin C_{n+1}^{G'(n+1)}$ with total probability, which implies SP3. $\square$

6.3. Main estimate.

Theorem 6.2 (Main estimate). *Let f be a unimodal map with the following properties:*

1. *f is Collet-Eckmann and has an absolutely continuous invariant measure μ;*
2. *The several asymptotic limits and estimates described in §4.4 hold;*
3. *f satisfies the "Strong poles fall in big gaps" condition.*

Then, there exists $n_0 > 0$ such that for every $\delta > 0$, and all $n \geq n_0$, there exists C_n such that for any I_n^j, $\mu(I_n^j) < C_n|I_n^j|^{1-\delta}$.

It turns out that Theorem C implies that we can take $n_0 = 1$ in the Main estimate.

Remark 6.1. We think that it is possible to refine the conditions of the Main estimate (keeping total probability) in order to obtain better estimates for $\mu(I_n^j)$ (of the type $-C|I_n^j|\ln|I_n^j|$ or even better). It is an interesting problem whether a bound of the type $C|I_n^j|$ is valid with total probability. Such a bound is equivalent to obtaining $d\mu_f^K \in L^\infty$ in Theorem C.

6.4. **Proof of Theorem C assuming the Main estimate.** By Lemma 6.1 and the results of [AM1], we get that, with total probability, f satisfies the hypothesis of the Main estimate. Let us now fix such an f.

If K is a hyperbolic set for f, then it avoids a neighborhood of the critical point. On the other hand, if $K \subset K'$ and $d\mu^{K'} \in L^p$, then necessarily $d\mu^K \in L^p$. So we just have to consider the case of K_n, the maximal invariant of $I \setminus I_n$ for n big. The gaps of K_n are connected components of the domain of the first landing map from I to I_n.

We will use the following:

Lemma 6.3. *For all $n > 0$, there exists a finite partition of $I \setminus I_n$ on intervals M_i, such that for each M_i, $f|M_i$ is a diffeomorphism onto the union of some M_i and, possibly, I_n. Moreover, there exists constants $C > 0$ and $t < 1$ such that for any x such that $x, ..., f^k(x) \in I \setminus I_n$, we can associate an interval $M^k(x)$ such that*

1. *$f^k : M^k(x) \to I$ is a diffeomorphism over some M_i;*
2. *Two intervals $M^k(x)$ and $M^k(y)$ are either disjoint or coincide.*
3. *$|M^k(x)| < Ct^k$;*
4. *$\sum_{k\geq 1} |M^k(x)| < C$;*
5. *The distortion of $f^k|M^k(x)$ is bounded by C;*
6. *For each k, $|\cup M^k(x)| < Ct^k$.*

Proof. Let Q be the finite set consisting of all points in the forward orbit of ∂I_n. Let M_i be the connected components of $I \setminus (Q \cup I_n)$. It is clear that the image of M_i consists of a union of M_j, possibly together with I_n. The M_i form a Markov partition of $I \setminus I_n$, and so the first and second item follow. The third item follows from hyperbolicity of $f|I \setminus I_n$ (see Lemma 2.3), and the fourth follows from the third. The fifth follows from the fourth by a classical argument (it is enough to use that $\ln |Df|$ is Hölder in $I \setminus I_n$). Notice that for each i, there exists $j_i > 0$ such that $f^{j_i}(M_i)$ contains I_n. This and the fifth item show that $|\cup M^{k+j}(x)| \leq t|\cup M^k(x)|$ for some $t < 1$ and for $j = \max j_i$, and this gives the sixth item. $\square$

Corollary 6.4. *For all $n > 0$, there exists $\theta_n > 0$ such that*

$$\sum_{\Lambda \text{ gap of } K_n} |\Lambda|^{1-\theta_n} < \infty. \tag{6.22}$$

Proof. Let us say that a gap Λ of K_n is of generation k if $f^k(\Lambda) = I_n$. Let $k(\Lambda)$ be the generation of Λ. Notice that each $M^k(x)$ contains at most one gap of generation $k+1$ (and no gaps of generation $\leq k$). On the other hand, each gap Λ of generation $k+1$ is contained on some $M^k(x)$, which we denote $M(\Lambda)$. Notice that since the derivative of f is bounded by some constant κ,

$$|\Lambda| \geq |I_n|\kappa^{-k(\Lambda)}. \tag{6.23}$$

We can estimate

$$\begin{aligned}\sum_{\Lambda \text{ gap of } K_n} |\Lambda|^{1-\theta_n} &\leq \sum_{k\geq 0} \sum_{k(\Lambda)=k} |M(\Lambda)|^{1-\theta_n} \leq \sum_{k\geq 0} \sum_{k(\Lambda)=k} |\Lambda|^{-\theta_n} |M(\Lambda)| \\ &\leq |I_n|^{-\theta_n} \left(\sum_{k\geq 0} \kappa^{\theta_n k} \sum_{k(\Lambda)=k} |M(\Lambda)| \right) \leq C|I_n|^{-\theta_n} \sum_{k\geq 0} \kappa^{\theta_n k} t^k,\end{aligned} \tag{6.24}$$

where $C > 0$, $t < 1$ comes from item (5) of Lemma 6.3. The result follows with $\theta_n > 0$ such that $\kappa^{\theta_n} t < 1$. $\square$

6.4.1. Let K_n^r be the Cantor set $I_n \setminus \cup I_n^j$, and let $d\mu^{K_n^r}$ be the function which takes, in each gap of K_n^r the average value of $d\mu$ in that gap, and let $d\mu^{K_n^r} = 0$ outside I_n.

Notice that by Corollary 6.4,

$$\sum_j |I_n^j|^{1-\theta_{n+1}} \leq \sum_{\underline{d}} |C_n^{\underline{d}}|^{1-\theta_{n+1}} \leq C < \infty. \tag{6.25}$$

Fix $1 \leq p < \infty$. Using the Main Estimate, let C' be such that

$$\mu(I_n^j) < C' |I_n^j|^{1-\frac{\theta_{n+1}}{p}}. \tag{6.26}$$

We estimate

$$\int_{I_n} (d\mu^{K^r_n})^p = \sum |I^j_n| \left(\frac{\mu(I^j_n)}{|I^j_n|}\right)^p \le C'^p \sum |I^j_n|^{1-\theta_{n+1}} < CC'^p. \tag{6.27}$$

In particular, $d\mu^{K^r_n} \in L^p$.

6.4.2. Given $\Lambda \subset I \setminus I_n$ measurable, let $\hat{\Lambda} \subset I$ be the set of x such that $\min\{k \ge 1, f^k(x) \in \Lambda\} \le \min\{k \ge 1, f^k(x) \in I_n\}$ (that is, the orbit of $f(x)$ intersects Λ before intersecting I_n). Let $\Lambda_l = \hat{\Lambda} \cap (I \setminus I_n)$ and $\Lambda_r = \hat{\Lambda} \cap I_n$.

Notice that $f^{-1}(\Lambda_l \cup \Lambda) = \Lambda_r \cup \Lambda_l = \hat{\Lambda}$, thus

$$\mu(\Lambda) = \mu(\Lambda_r) \quad \text{provided } \Lambda_l \cap \Lambda = \emptyset. \tag{6.28}$$

Let $\Lambda(k) \subset \Lambda_l$ be the set of points x with $k = \min\{i > 0, f^i(x) \in \Lambda\}$. Then, by Lemma 6.3, $\Lambda(k)$ is covered by disjoint intervals $M^k(y_i)$. By items 5 and 6 of Lemma 6.3, there exists $C > 0$, $t < 1$ such that

$$|\Lambda(k)| < Ct^k \max_i \frac{|\Lambda(k) \cap M^k(y_i)|}{|M^k(y_i)|} < C't^k|\Lambda|, \tag{6.29}$$

since the density of $\Lambda(k)$ inside some $M^k(y_i)$ is comparable with the density of Λ in $f^k(M^k(y_i))$ (by bounded distortion) which is at most $\max_j |\Lambda||M_j|^{-1}$. Thus, there exists a constant $C > 0$ such that

$$|\Lambda_l| < C|\Lambda| \tag{6.30}$$

independently of Λ.

6.4.3. Let now Λ be a gap of K_n. Assume first that Λ is a gap of K_n which does not intersect $\{f^i(0), 0 \le i < v_n\}$. In particular, $\Lambda \ne I_n$ and Λ_r does not contain the critical point. Since Λ is a connected component of the domain of the first landing map from I to I_n, we have that $\Lambda_l \cap \Lambda = \emptyset$, and thus (6.28) holds. By (6.30),

$$|\Lambda_r| \le 2|\Lambda_l| \sup_{x \in I_n \setminus I^0_n} |Df(x)|^{-1} \le C|\Lambda| \tag{6.31}$$

for some constant $C > 0$ independent of Λ. Using the Hölder inequality we get

$$\mu(\Lambda) = \mu(\Lambda_r) \le \left(\int_{\Lambda_r} (d\mu^{K^r_n})^p\right)^{\frac{1}{p}} |\Lambda_r|^{1-\frac{1}{p}} < C_p|\Lambda_r|^{1-\frac{1}{p}} \le C'_p|\Lambda|^{1-\frac{1}{p}}, \tag{6.32}$$

where C'_p depends on p but not on Λ.

The set of gaps Λ of K_n which intersect $\{f^i(0), 0 \le i < v_n\}$ are in finite number, so there exists $C > 0$ such that for any such Λ,

$$\mu(\Lambda) \le C|\Lambda|. \tag{6.33}$$

Putting together (6.32) and (6.33), and varying p, we see that for any $\delta > 0$ there exists a constant $C(\delta)$ such that for any Λ gap of K_n we have

$$\mu(\Lambda) \le C(\delta)|\Lambda|^{1-\delta}. \tag{6.34}$$

By Corollary 6.4, for $\delta < \theta_n/p$ we have

$$\int (d\mu^{K_n})^p = \sum_{\Lambda \text{ gap of } K_n} |\Lambda| \left(\frac{\mu(\Lambda)}{|\Lambda|}\right)^p \le C(\delta)^p \sum |\Lambda|^{1-\theta_n} < CC(\delta)^p. \tag{6.35}$$

7. Proof of the Main Estimate

7.1. **Outline.** Our problem is to analyze the asymptotics of the physical measure of I_n^j as the Lebesgue measure of I_n^j decreases, n fixed. Let us fix some advanced level η. Fix a small interval I_η^ι. To the end of this section, $a = |I_\eta^\iota|$. The critical step l is defined as the unique number with

$$c_{l+1} \leq a < c_l. \tag{7.1}$$

Since our estimate is only relevant if I_η^ι is small, we can assume that l is very big.

The proof will be based on the analysis, for each level $n \geq \eta$, of the frequency of visits to I_η^ι before a return to I_n. Those estimates can be passed from level to level if one can control the distortion introduced by the critical orbit. The argument will take distinct steps.

In the early stages (before $l-1$), very few branches (measure of order a) pass at all in I_η^ι before returning. The critical orbit falls in big holes away from the hierarchic structure of those branches and does not distort much the measure.

In the later stages (after $l+2$) most branches have total time much bigger then a^{-1}, and they spend a proportion of time of order a in I_η^ι. The exceptional branches have measure much smaller than a, and we use the inductive estimate of §5.2 to show that they do not contribute much for the next levels.

In the intermediate stage, there is a delicate transition between those two situations. To complicate further, at this moment the position of the critical point could introduce distortion of strength comparable with a. We will need to use the hierarchical structure of the set of branches passing through I_η^ι combined with our conditions on the critical orbit to control the distortion of pullbacks.

In what follows, δ will denote several constants which go to 0 uniformly as the critical step l goes to infinity.

7.2. **Preliminaries.** Let us define

$$X_n(j) = \#\{k < r_n(j), f^k(I_n^j) \subset I_\eta^\iota\}, \quad n \geq \eta, \tag{7.2}$$

$$\mathbb{X}_n(\underline{d}) = \#\{k < l_n(\underline{d}), f^k(C_n^{\underline{d}}) \subset I_\eta^\iota\}, \quad n \geq \eta, \tag{7.3}$$

so that

$$\mathbb{X}_n(\underline{d}) = \sum_{i=1}^m X_n(j_i), \quad \underline{d} = (j_1, ..., j_m). \tag{7.4}$$

$$X_{n+1}(j) = X_n(0) + \mathbb{X}_n(\underline{d}), \quad R_n(I_{n+1}^j) \subset C_n^{\underline{d}}. \tag{7.5}$$

Define

$$x_n(r) = I(\{X_n(j) \geq r\}, n), \quad x_n = x_n(1), \tag{7.6}$$

$$\mathbf{x}_n(r) = C(\{\mathbb{X}_n(\underline{d}) \geq r\}, n), \quad \mathbf{x}_n = \mathbf{x}_n(1). \tag{7.7}$$

One immediately gets (see §7.5.1.1 for a derivation)

$$\mathbf{x}_n \leq n c_n^{-1} x_n. \tag{7.8}$$

Before getting into the more complicate intermediate steps, let us deal with the initial steps and discuss our strategy for the later steps.

7.2.1. *Initial steps.* Notice that if $x_n < c_n$ then $X_n(0) = 0$, and if additionally $\mathbf{x}_n < c_{n+1}^5$ then we conclude that $x_{n+1} < c_{n+1}$ and $X_{n+1}(0) = 0$ as well. In this case we can estimate

$$x_{n+1} \leq \operatorname{dist}(R_n|(I_{n+1} \setminus I_{n+2}))2^n \mathbf{x}_n \leq 2^{2n} c_{n+1}^{-1} x_n. \tag{7.9}$$

Since $X_\eta(0) = 0$ and $x_\eta(0) = a$, we conclude by induction that for $\eta \leq n \leq l-1$ we have

$$X_n(0) = 0, \quad x_n \leq a c_n^{-1} c_{n-1}^{-5/2} \ll a^{1-\delta} < c_n, \quad \mathbf{x}_n \leq a c_n^{-9/4} \tag{7.10}$$

(using that $a c_n^{-9/4} < c_{n+1}^5$ for $n \leq l-2$). Thus we have, just before the critical time:

$$x_{l-1}, \mathbf{x}_{l-1} \leq a c_{l-1}^{-5/2} \ll a^{1-\delta}, \quad X_{l-1}(0) = 0. \tag{7.11}$$

7.2.2. *Later steps.* Let us fix, for the end of this section, some very small ϵ. Our aim is to estimate

$$\mu(I_\eta^\iota) \leq a^{1-10\epsilon^{1/2}}. \tag{7.12}$$

To attack this problem, we will need to compute $\mu(I_\eta^\iota)$ somehow. Using the idea of §5.4.2, one sees that we only have to show that

$$\lim_{n\to\infty} I(\{j, X_n(j) \leq a^{1-10\epsilon^{1/2}} r_n(j)\}, n) = 1. \tag{7.13}$$

This will be done in the following way. We will show that there exists n_0 such that, defining

$$T = \{j, X_{n_0}(j) > a^{1-8\epsilon^{1/2}} r_{n_0}(j)\}, \tag{7.14}$$

we have

$$I(T, n_0) < c_{n_0-1}^{1/20}. \tag{7.15}$$

Our actual choice of n_0 will be $n_0 = l+3$ if $a < c_{l+1}^{\epsilon^{1/2}}$ and $n_0 = l+2$ otherwise. Notice that in both cases $c_{n_0-1}^{1/200} < a$.

By the inductive estimate, see §5.2.2 (we only need the case $\gamma = 1$ corresponding to Lebesgue measure), we get, for $n \geq n_0$,

$$I(\mathbb{Z} \setminus VG(T, n_0, n), n) \ll c_{n-1}^{1/20}, \tag{7.16}$$

and for $j \in VG(T, n_0, n)$

$$X_n(j) \leq (a^{1-8\epsilon^{1/2}} + c_{n_0-1}^{1/200}) r_n(j) < a^{1-10\epsilon^{1/2}} r_n(j). \tag{7.17}$$

Estimates (7.16) and (7.17) imply (7.13).

7.3. **Transition from $l-1$ to l.** The analysis of the transition from the $l-1$ level to the l level is more complicated. We will need to consider a special sequence E_t of nested intervals in level I_{l-1} around the critical value $R_{l-1}(0)$, where we can analyze (using SP2) the density of the set of points visiting I_η^ι. We then pullback this information to a sequence F_t of nested intervals in level l around the critical point, with a control of the distortion by SP1.

7.3.1. Let $\underline{d} = (j_1, ..., j_s)$ be such that $R_{l-1}(0) \in C_{l-1}^{\underline{d}}$ and for each $0 \le t \le s$, let $\underline{d}(t) = (j_1, ..., j_t)$. By condition SP2 we can estimate

$$I(j_i, l-1) > e^{-c_{l-2}^{-12}} \gg a^{\delta}, \quad 1 \le i \le s. \tag{7.18}$$

Define

$$E_t = I_{l-1}^{\underline{d}(t)}, \quad 0 \le i \le s. \tag{7.19}$$

From (7.18) we get

$$\frac{|E_{t+1}|}{|E_t|} > 2^{-l} e^{-c_{l-2}^{-12}} \gg a^{\delta}. \tag{7.20}$$

Denote

$$F_t = (R_{l-1}|I_l)^{-1}(E_t). \tag{7.21}$$

Notice that (7.18), (7.11) imply that $X_{l-1}(j_i) = 0$. In particular,

$$C_{l-1}^{\{\mathbb{X}_{l-1}>0\}} \cap E_t = (R_{l-1}^t|E_t)^{-1} C_{l-1}^{\{\mathbb{X}_{l-1}>0\}}. \tag{7.22}$$

This forces the density estimate

$$\frac{|C_{l-1}^{\{\mathbb{X}_{l-1}>0\}} \cap E_t|}{|E_t|} < a c_{l-1}^{-3} \le a^{1-\delta}. \tag{7.23}$$

Moreover, since $\mathbb{X}_{l-1}(\underline{d}) = \sum_{i=1}^{s} X_{l-1}(j_i) = 0$ and $X_{l-1}(0) = 0$ (by 7.11), we have that

$$X_l(0) = 0. \tag{7.24}$$

7.3.2. Let us define a_t and b_t by $E_t = (R_{l-1}(0) - a_t, R_{l-1}(0) + b_t)$. Let $\mathbf{m}_t = \min\{a_t, b_t\}$ and $\mathbf{M}_t = \max\{a_t, b_t\}$, so that $\frac{|E_t|}{2} \le \mathbf{M}_t < |E_t|$. By SP1, we have $\mathbf{m}_t > \frac{|E_t|}{2^{l-1}(t+1)^2}$, so we have

$$\frac{\mathbf{M}_t}{\mathbf{m}_{t+1}} \le \frac{|E_t|}{\mathbf{m}_{t+1}} \le 2^{l-1}(t+2)^2 \frac{|E_t|}{|E_{t+1}|} \le e^{c_{l-2}^{-14}} \le a^{-\delta}. \tag{7.25}$$

This allows us to estimate

$$\operatorname{dist}(R_{l-1}|(F_t \setminus F_{t+1})) \le e^{c_{l-2}^{-15}} \le a^{-\delta}, \quad 0 \le t < s. \tag{7.26}$$

If $0 \le t \le s-1$, let E_t^1, E_t^2 be the connected components of $E_t \setminus E_{t+1}$, and if $t = s$, let E_t^1 and E_t^2 be the connected components of $I_{l-1}^{\underline{d}(s)} \setminus C_{l-1}^{\underline{d}(s)}$, (recall that $E_s = I_{l-1}^{\underline{d}(s)}$). Using SP1 and (7.23) we see that

$$\frac{|C_{l-1}^{\{\mathbb{X}_{l-1}>0\}} \cap E_t^i|}{|E_t^i|} < a c_{l-1}^{-3} 2^l (t+1)^2 \le a^{1-\delta}, \quad 0 \le t \le s, \quad i = 1, 2. \tag{7.27}$$

Using (7.26), (7.27) and SP1, we obtain,

$$\frac{|(F_t \setminus F_{t+1}) \cap I_l^{\{X_l>0\}}|}{|F_t|} \le a e^{c_{l-2}^{-16}} \ll a^{1-\delta}, \quad 0 \le t < s. \tag{7.28}$$

Notice that $X_l(0) = 0$, so $F_s \cap I_l^{\{X_l>0\}} = (F_s \setminus I_{l+1}) \cap I_l^{\{X_l>0\}}$. Notice that R_{l-1} takes each component of $F_s \setminus I_{l+1}$ to either E_s^1 or E_s^2, and we have the obvious estimate $\operatorname{dist}(R_{l-1}|(F_s \setminus I_{l+1})) < 2^l c_{l-1}^{-1/2}$. By (7.27)

$$\frac{|F_s \cap I_l^{\{X_l>0\}}|}{|F_s|} \le a 2^{2l} c_{l-1}^{-7/2} (s+1)^2 \le a e^{-c_{l-2}^{-16}} \ll a^{1-\delta}, \tag{7.29}$$

as well.

We have

$$I_l = F_s \cup \bigcup_{t=0}^{s-1} (F_t \setminus F_{t+1}), \tag{7.30}$$

so, by (7.28) and (7.29),

$$x_l \leq a e^{c_{l-2}^{-16}} \ll a^{1-\delta}, \tag{7.31}$$

and as a consequence,

$$\mathbf{x}_l \leq a 2^l c_l^{-1} e^{c_{l-2}^{-16}} \ll a^{1-\delta} c_l^{-1}. \tag{7.32}$$

7.4. **The critical step.** We will consider two cases: $a < e^{-c_{l-1}^{-20}}$ (Case 1) and $a \geq e^{-c_{l-1}^{-20}}$ (Case 2).

7.4.1. *Case 1.* The first case can be dealt by an argument which is analogous to the analysis in §7.3. We consider a sequence of nested intervals E_t in level l around $R_l(0)$, and also their pullback F_t in level $l+1$ (the definitions are the same of §7.3 up to a shift in the indexes). Using SP2 we analyze the density of $C_l^{\{X_l>0\}}$ in E_t, which we bound by $a^{1-\delta}$ (this only works in Case 1), and since $a^{1-\delta} \ll c_l$ (this only works in Case 1 also) we conclude that $X_{l+1}(0) = 0$. We use SP1 to control the pullback to F_t. The reader can check the estimate

$$x_{l+1} < a^{1-\delta}, \quad \text{in Case 1.} \tag{7.33}$$

7.4.2. *Case 2.* Let $r = e^{c_{l-2}^{-4}}$ and $q = (\ln a)^2$, so that $q < r$ (since we are in Case 2).

Let $\underline{d} = (j_1, ..., j_s)$ be such that $R_l(0) \in C_l^{\underline{d}}$. For $1 \leq u \leq q$, we let $\underline{d}_u$ be defined by

$$I_l^{j_u} = (R_{l-1}|I_l)^{-1}(C_{l-1}^{\underline{d}_u}). \tag{7.34}$$

and we let $\underline{d}_u = (j_1^u, ..., j_{s(u)}^u)$.

Notice that by SP2,

$$s(u) \leq r_l(j_u) \leq c_{l-1}^{-14}. \tag{7.35}$$

For $1 \leq u \leq q$ and $0 \leq v \leq s(u)$, let us define a sequence of nested intervals $S_{u,v}$ containing $C_{l-1}^{\underline{d}_u}$ by $S_{u,v} = I_{l-1}^{(j_1^u,...,j_v^u)}$. Let us define nested intervals $T_{u,v}$ containing $I_l^{j_u}$ by taking $T_{u,v}$ as the connected component of $(R_{l-1}|I_l)^{-1}(S_{u,v})$ containing $I_l^{j_u}$. Let

$$W_{u,v} = (R_l^{u-1}|I_l^{(j_1,...,j_{u-1})})^{-1}(T_{u,v}), \tag{7.36}$$

which is some interval containing $R_l(0)$. Notice that if $(u_1, v_1) \leq (u_2, v_2)$ in the lexicographic order we have $W_{u_1,v_1} \supset W_{u_2,v_2}$.

Note also that $|W_{q,s(q)}| < e^{-q}|I_l| < a^{10}$ so that

$$\frac{|(R_l|I_{l+1})^{-1}(W_{q,s(q)})|}{|I_{l+1}|} \ll a. \tag{7.37}$$

By our choice of q, we can apply SP3 and conclude

$$I(j_v^u, l-1) > e^{-c_{l-2}^{-12}} > a^{1/2} > x_{l-1}, \quad 1 \leq u \leq q, \quad v \leq s(u). \tag{7.38}$$

This gives

$$X_{l-1}(j_v^u) = 0, \quad 1 \leq u \leq q, \quad v \leq s(u). \tag{7.39}$$

If $X_l(j_u) > 0$ then there exists some $v \leq s(u)$ such that $X_{l-1}(j_v^u) > 0$ (since $X_{l-1}(0) = 0$), so (7.39) implies

$$X_l(j_u) = 0, \quad 1 \leq u \leq q. \tag{7.40}$$

Notice that (7.39), (7.40) imply

$$S_{u,v} \cap C_{l-1}^{\{\mathbb{X}_{l-1}>0\}} = (R_{l-1}^{v}|S_{u,v})^{-1}(C_{l-1}^{\{\mathbb{X}_{l-1}>0\}}), \tag{7.41}$$

$$C_l^{\{\mathbb{X}_l>0\}} \cap W_{u,v} = (R_l^{u-1}|I_l^{(j_1,\dots,j_{u-1})})^{-1}(T_{u,v} \cap C_l^{\{\mathbb{X}_l>0\}}). \tag{7.42}$$

In particular, (7.41) gives

$$\frac{|S_{u,v} \cap C_{l-1}^{\{\mathbb{X}_{l-1}>0\}}|}{|S_{u,v}|} < 2^l \mathbf{x}_{l-1} \leq a^{1-\delta}. \tag{7.43}$$

7.4.2.1. Let us now show that

$$\frac{|T_{u,v} \cap I_l^{\{X_l>0\}}|}{|T_{u,v}|} \leq a^{1-\delta}. \tag{7.44}$$

There are two cases: $T_{u,v} \supset I_{l+1}$ and otherwise.

In the first case, $T_{u,v} \supset I_{l+1}$, we have $S_{u,v} = E_v$ and $T_{u,v} = F_v$ in the notation of §7.3, and (7.44) follows from (7.28) and (7.29).

In the second case, $T_{u,v} \not\supset I_{l+1}$, using SP1 and that $v \leq s(u) \leq c_{l-1}^{-14}$ (see (7.35)), we get

$$\operatorname{dist}(R_{l-1}|T_{u,v}) < v^2 2^{l-1} \ll c_{l-1}^{-50} \leq a^{-\delta}. \tag{7.45}$$

Since $X_{l-1}(0) = 0$, we have

$$T_{u,v} \cap I_l^{\{X_l>0\}} = (R_{l-1}|T_{u,v})^{-1}(C_{l-1}^{\{\mathbb{X}_{l-1}>0\}} \cap S_{u,v}), \tag{7.46}$$

and (7.44) follows from (7.43) and (7.45).

7.4.2.2. Since $X_l(0) = 0$ (see (7.24)),

$$T_{u,v} \cap C_l^{\{\mathbb{X}_l>0\}} = T_{u,v} \cap \left(I_l^{\{X_l>0\}} \cup \bigcup_{\substack{I_l^j \subset T_{u,v},\\ j\neq 0}} (R_l|I_l^j)^{-1} C_l^{\{\mathbb{X}_l>0\}} \right). \tag{7.47}$$

And by (7.31), (7.32) and (7.44) we get

$$\frac{|T_{u,v} \cap C_l^{\{\mathbb{X}_l>0\}}|}{|T_{u,v}|} \leq a^{1-\delta} c_l^{-1}. \tag{7.48}$$

By (7.48) and (7.42) we get

$$\frac{|C_l^{\{\mathbb{X}_l>0\}} \cap W_{u,v}|}{|W_{u,v}|} < a^{1-\delta} c_l^{-1}, \quad 1 \leq u \leq q, \quad v \leq s(u). \tag{7.49}$$

7.4.2.3. Let $Z_0 \supset Z_1 \supset ... \supset Z_t$ be an enumeration of the $W_{u,v}$. Let us show that

$$\frac{|Z_k|}{|Z_{k+1}|} \ll a^{-\delta}, \quad 0 \leq k \leq t-1. \tag{7.50}$$

Notice that if $v < s(u)$, and $T_{u,v+1}$ does not contain 0,

$$\frac{|T_{u,v}|}{|T_{u,v+1}|} \leq 10 \frac{|S_{u,v}|}{|S_{u,v+1}|} \tag{7.51}$$

(using that $R_{l-1}|I_l$ is almost purely quadratic[8]). So in this case,

$$\frac{|W_{u,v}|}{|W_{u,v+1}|} \leq 2^l \frac{|T_{u,v}|}{|T_{u,v+1}|} \leq 2^{2l} \frac{|S_{u,v}|}{|S_{u,v+1}|} \leq 2^{4l} I(j^u_{v+1}, l-1)^{-1} \leq e^{c_{l-2}^{-14}} \ll a^{-\delta} \tag{7.52}$$

using (7.38) to estimate $I(j^u_{v+1}, l-1)$.

Let us now assume that $v < s_u$ but that $T_{u,v+1}$ contains 0 (in this case $T_{u,v+1} = F_{v+1}$, $T_{u,v} = F_v$, $S_{u,v+1} = E_{v+1}$, $S_{u,v} = E_v$ in the notation of §7.3). Notice that $R_{l-1}(0) \in S_{u,v+1}$, so we can apply SP1 to see that

$$|R_{l-1}(T_{u,v+1})| \geq \frac{|S_{u,v+1}|}{2^l (v+1)^2}. \tag{7.53}$$

Thus,

$$\frac{|W_{u,v}|}{|W_{u,v+1}|} \leq 2^l \frac{|T_{u,v}|}{|T_{u,v+1}|} \leq 2^l \cdot 10 \left(\frac{R_{l-1}(T_{u,v})}{R_{l-1}(T_{u,v+1})} \right)^{1/2} \leq 2^l \cdot 10 \left(\frac{2^l (v+1)^2 |S_{u,v}|}{|S_{u,v+1}|} \right)^{1/2}. \tag{7.54}$$

By (7.35), $v+1 \leq c_{l-1}^{-14}$, while SP2 implies that $|S_{u,v+1}| \geq 2^{-l} e^{-c_{l-2}^{-12}} |S_{u,v}|$ (this is the same estimate as (7.20)), so we get

$$\frac{|W_{u,v}|}{|W_{u,v+1}|} \leq 10 \cdot 2^{2l} c_{l-1}^{-14} e^{c_{l-2}^{-12}/2} \leq e^{c_{l-2}^{-14}} \ll a^{-\delta}. \tag{7.55}$$

Consider now the case of $v = s(u)$. Notice that $W_{u+1,0} = (R_l^{u-1} | I_l^{(j_1,\ldots,j_{u-1})})^{-1} (I_l^{j_u})$, so

$$\frac{|W_{u,s(u)}|}{|W_{u+1,0}|} \leq 2^l \frac{|T_{u,s(u)}|}{|I_l^{j_u}|}. \tag{7.56}$$

Moreover, $I_l^{j_u}$ is a connected component of $(R_{l-1}|I_l)^{-1}(C_{l-1}^{\underline{d}_u})$ and $T_{u,s(u)}$ is a connected component of $(R_{l-1}|I_l)^{-1}(I_{l-1}^{\underline{d}_u})$. Since $C_{l-1}^{\underline{d}_u}$ does not contain $R_{l-1}(0)$ (since $j_u \neq 0$), we conclude

$$\frac{|W_{u,s(u)}|}{|W_{u+1,0}|} \leq 2^l \frac{|T_{u,s(u)}|}{|I_l^{j_u}|} \leq 2^{2l} \frac{|I_{l-1}^{\underline{d}_u}|}{|C_{l-1}^{\underline{d}_u}|} \leq c_{l-1}^{-2} \ll a^{-\delta}. \tag{7.57}$$

[8]This follows from the following estimate: if $0 \leq a < b < c < d$ then $2\frac{d^2-a^2}{c^2-b^2} \geq 2\frac{d+a}{c+b}\frac{d-a}{c-b} \geq \frac{d-a}{c-b}$. Thus the quadratic part of the pullback can not decrease the relative size of $S_{u,v+1} \subset S_{u,v}$ by a factor worse than 2 if $0 \notin S_{u,v}$, or 4 if $0 \in S_{u,v} \setminus S_{u,v+1}$.

7.4.2.4. As in §7.3, define $a_i, b_i > 0$ by $Z_i = (R_l(0) - a_i, R_l(0) + b_i)$, and let $\mathbf{m}_i = \min\{a_i, b_i\}$, $\mathbf{M}_i = \max\{a_i, b_i\}$. Notice that for $i < t$ we have $m_i > 2M_{i+1}$ (both components of $Z_i \setminus Z_{i+1}$ are much bigger than Z_{i+1}). It follows that for $i < t-1$, we have (using (7.50))

$$\frac{\mathbf{M}_i}{\mathbf{m}_{i+1}} \leq \frac{\mathbf{M}_i}{\mathbf{M}_{i+1}} \frac{\mathbf{M}_{i+1}}{\mathbf{M}_{i+2}} \leq a^{-\delta}. \tag{7.58}$$

Let $V_i = (R_l|I_{l+1})^{-1}(Z_i)$. Repeating the argument used to obtain (7.28) and (7.29) we get

$$\frac{|(V_i \setminus V_{i+1}) \cap I_{l+1}^{\{X_{l+1}>0\}}|}{|V_i|} \leq a^{1-\delta} c_l^{-1}. \tag{7.59}$$

On the other hand,

$$|Z_{t-2}| \ll e^{-q}|I_l| \ll a^{10}|I_l|, \tag{7.60}$$

so $|V_{t-2}| \ll a|I_{l+1}|$. Repeating the argument of (7.31) and (7.32) we get

$$x_{l+1} \leq a^{1-\delta} c_l^{-1}, \quad \text{in Case 2.} \tag{7.61}$$

Remark 7.1. The above estimate from above could be bigger than one if a is near c_l. This means that X_{l+1} could be supported on most branches of R_{l+1}.

7.5. **Dealing with the later steps.** We are now in position to work out the later steps, aiming at the estimates outlined in §7.2.2. Before doing so, let us present a couple of tools that will be repeatedly used.

7.5.1. *Useful estimates.* We will need several times the following easy estimates.

7.5.1.1. Let $T \subset \mathbb{Z}$ and let $q = I(T, n)$. Let $\hat{T}$ be the set of $\underline{d}$ with at least one entry in T. Then

$$\begin{aligned} C(\hat{T}, n) &\leq \sum_{k \geq 1} I(\{\underline{d} = (j_1, ..., j_k),\, j_k \in T\}, n) \leq n^{1/2+\delta} \sum_{k \geq 1} qI(\{|\underline{d}| = k-1\}, n) \\ &\leq n \sum_{k \geq 1} qc_n^{-1} C(\{|\underline{d}| = k-1\}, n) = nc_n^{-1} q. \end{aligned} \tag{7.62}$$

7.5.1.2. Let $T \subset \mathbb{Z}$, let $q = I(T, n)$, and assume $nq < 1/2$. Assume that $k > k_0 > c_{n-1}^{-2}$ satisfy $kk_0^{-1} c_n > n^2 q$. Let $\hat{T}$ be the set of $\underline{d}$ with at least k entries in T. Then

$$C(\hat{T}, n) < e^{-k_0/8}. \tag{7.63}$$

To see this, let $\hat{T}(m)$ be the set of $\underline{d} \in \hat{T}$ of length m. Then

$$C(\hat{T}(m), n) \leq \sum_{j=k}^{m} \binom{m}{j} (nq)^j (1-nq)^{m-j}. \tag{7.64}$$

Let $m_0 = c_n^{-1} k_0$. For $m \leq m_0$,

$$C(\hat{T}(m), n) \leq \sum_{j=k}^{m_0} \binom{m_0}{j} (nq)^j (1-nq)^{m_0-j}. \tag{7.65}$$

Notice that for $j \geq k$,

$$\binom{m_0}{j} (nq)^j (1-nq)^{m_0-j} \leq \binom{m_0}{k} (nq)^k (1-nq)^{m_0-k} \tag{7.66}$$

so for $k \le m \le m_0$ we have

$$C(\hat{T}(m),n) \le m_0 \binom{m_0}{k} (nq)^k (1-nq)^{m_0-k} \le m_0 \left(\frac{em_0}{k}\right)^k (nq)^k (1-nq)^{m_0-k} \le m_0 e^{-2k} \le e^{-k}, \tag{7.67}$$

since $(m_0 k^{-1})nq \ll 1$.

Thus

$$C(\hat{T},n) \le C(\{|\underline{d}| > m_0\},n) + \sum_{m \le m_0} C(\hat{T}(m),n) \le e^{-k_0/4} + m_0 e^{-k} \le e^{-k_0/8}. \tag{7.68}$$

7.5.1.3. Let $T \subset \mathbb{Z}$ and let $q = I(T,n)$. Let $q_0 > n^2 q$ be such that $q_0 > c_n^{1/5}$. Let $\hat{T}$ be the set of $\underline{d}$ with $|\underline{d}| > c_n^{-1/2}$ and with at least $q_0|\underline{d}|$ entries in T. Then

$$C(\hat{T},n) \le e^{-c_n^{-1/4}}. \tag{7.69}$$

Indeed, by similar considerations as in §7.5.1.2,

$$\begin{aligned} C(\hat{T},n) &\le \sum_{m > c_n^{-1/2}} \sum_{j=q_0 m}^{m} \binom{m}{j} (nq)^j (1-nq)^{m-j} \\ &\le \sum_{m > c_n^{-1/2}} \sum_{j=q_0 m}^{m} \binom{m}{q_0 m} (nq)^{q_0 m} (1-nq)^{m-q_0 m} \\ &\le \sum_{m > c_n^{-1/2}} m \left(\frac{e}{q_0}\right)^{q_0 m} (nq)^{q_0 m} (1-nq)^{m-q_0 m} \\ &\le \sum_{m > c_n^{-1/2}} m e^{-q_0 m} \le e^{-c_n^{-1/4}}. \end{aligned} \tag{7.70}$$

7.5.2. *Standard landings.* We will need only a couple of properties of standard landings (see §5.2.2) which we will put together here.

$$r_n(j_i) < c_{n-1}^{-14}, \qquad \underline{d} = (j_1, ..., j_m) \in LS(n), \tag{7.71}$$

$$|\underline{d}| > c_n^{-1/2}, \qquad \underline{d} \in LS(n), \tag{7.72}$$

$$l_n(\underline{d}) > c_{n-1}^{-1+\epsilon/2} |\underline{d}|, \qquad \underline{d} \in LS(n), \tag{7.73}$$

$$C(\Omega \setminus LS(n), n) < c_n^{1/3}. \tag{7.74}$$

7.6. **The $l+1$ level.** Recall that ϵ was fixed in advance, so that $\delta \ll \epsilon$ if l is big.

Assume that $\mathbb{X}_l(\underline{d}) \ge c_l^{-\epsilon}$, with $\underline{d} = (j_1, ..., j_m)$. Then $\underline{d} \in A \cup B$, where A is the set of $\underline{d}$ with $X_l(j_i) > c_l^{-\epsilon/2}$ for some j_i and B is the set of $\underline{d}$ with $\#\{i | X_l(j_i) > 0\} \ge c_l^{-\epsilon/2}$.

Since $I(\{r_l(j) > c_l^{-\epsilon/2}\}, l) < e^{-c_l^{-\epsilon/3}}$, the estimate of §7.5.1.1 gives

$$C(A,l) \le e^{-c_l^{-\epsilon/4}}. \tag{7.75}$$

Let $k = c_l^{-\epsilon/4}$, $k_0 = c_l^{-\epsilon/5}$. Since $k k_0^{-1} c_l > l^2 c_l^{1-\delta} \ge l^2 a^{1-\delta} \ge l^2 x_l$, the estimate of §7.5.1.2 gives

$$C(B,l) \le e^{-c_l^{-\epsilon/8}}. \tag{7.76}$$

We now conclude easily

$$\mathbf{x}_l(c_l^{-\epsilon}) \le C(A \cup B, l) < e^{-c_l^{-\epsilon/4}} + e^{-c_l^{-\epsilon/8}} \le e^{-c_l^{-\epsilon/11}} \le c_l^{l^3}. \tag{7.77}$$

After pullback by $R_l|I_{l+1}$ we get

$$x_{l+1}(c_l^{-2\epsilon}) \leq x_{l+1}(c_l^{-\epsilon}+v_l) < e^{-c_l^{-\epsilon/12}} \leq c_l^{l^3}. \tag{7.78}$$

7.7. **Levels $l+2$ and $l+3$.** We will need to consider two cases according to the size of a:

$$\text{Case A} \qquad c_{l+1} \leq a < c_{l+1}^{\epsilon^{1/2}}, \tag{7.79}$$

$$\text{Case B} \qquad c_{l+1}^{\epsilon^{1/2}} \leq a < c_l. \tag{7.80}$$

7.7.1. *Case A.* Notice that $x_{l+1} < a^{1-\delta}$ in this case.

Let us say that $\underline{d}$ is a BAD landing (of level $l+1$) if

$$\mathbb{X}_{l+1}(\underline{d}) > c_{l+1}^{-1+\epsilon} a^{1-5\epsilon^{1/2}}, \quad \text{BAD landing in Case A.} \tag{7.81}$$

7.7.1.1. Let us see that

$$C(\{\underline{d} \text{ is BAD}\}, l+1) \leq e^{-c_{l+1}^{-\epsilon/8}}, \quad \text{in Case A.} \tag{7.82}$$

Indeed, if $\underline{d}$ is BAD and $\underline{d}=(j_1,...,j_m)$ then $\underline{d} \in A \cup B$ where A is the set of $\underline{d}$ with some j_i with $r_{l+1}(j_i) > c_{l+1}^{-\epsilon/2}$ and B is the set of $\underline{d}$ with at least k entries in $\{X_{l+1}(j) > 0\}$ where $k = c_{l+1}^{-1+3\epsilon/2} a^{1-5\epsilon^{1/2}}$.

As before, the estimate of §7.5.1.1 gives

$$C(A, l+1) \leq e^{-c_{l+1}^{-\epsilon/4}}. \tag{7.83}$$

Let $k_0 = c_{l+1}^{-\epsilon}$. Since

$$k k_0^{-1} c_{l+1} > c_{l+1}^{3\epsilon} a^{1-5\epsilon^{1/2}} > a^{1-2\epsilon^{1/2}} > (l+1)^2 x_{l+1} \tag{7.84}$$

the estimate of §7.5.1.2 gives

$$C(B, l+1) \leq e^{-c_{l+1}^{-\epsilon/4}}, \tag{7.85}$$

and (7.82) follows.

7.7.1.2. Define the set of BAD returns (of level $l+2$) as the set of j such that $R_{l+1}(I_{l+2}^j) \subset C_{l+1}^{\underline{d}}$ where $\underline{d}$ is a BAD landing, so that $I(\{j \text{ is BAD}\}, l+2) < e^{-c_{l+1}^{-\epsilon/11}}$. Notice that a non BAD return j satisfies

$$X_{l+2}(j) \leq c_{l+1}^{-1+\epsilon} a^{1-5\epsilon^{1/2}} + v_{l+1}, \quad j \text{ is not a BAD return.} \tag{7.86}$$

Let us define a set VB of landings (of level $l+2$) as the set of $\underline{d}=(j_1,...,j_m)$ which are either non-standard or $\#\{i, j_i \text{ is a BAD return}\} > c_{l+1}^l |\underline{d}|$. Notice that $c_{l+1}^l > (l+2)^2 I(\{j \text{ is BAD}\}, l+2)$, so using the estimate of §7.5.1.3 (with $q_0 = c_{l+1}^l > c_{l+2}^{1/5}$), we conclude that

$$C(VB, l+2) \leq c_{l+2}^{2/7}. \tag{7.87}$$

If $\underline{d} \notin VB$, all returns have time at most $r_n(j_i) < c_{l+1}^{-14}$, so

$$\begin{aligned} \mathbb{X}_{l+2}(\underline{d}) &\leq (c_{l+1}^{-14} c_{l+1}^l + c_{l+1}^{-1+\epsilon} a^{1-5\epsilon^{1/2}} + v_{l+1})|\underline{d}| \\ &\leq (c_{l+1}^{l-14} + c_{l+1}^{-1+\epsilon} a^{1-5\epsilon^{1/2}} + c_l^{-1-\delta}) c_{l+1}^{1-\epsilon/2} l_{l+2}(\underline{d}) \leq a^{1-5\epsilon^{1/2}} l_{l+2}(\underline{d}) \end{aligned} \tag{7.88}$$

since $l_{l+2}(\underline{d}) > c_{l+1}^{-1+\epsilon/2} |\underline{d}|$ for a standard landing.

7.7.1.3. Define a VB return as the set of returns of level $l+3$ that fall in VB landings. Then $I(VB, l+3) < c_{l+2}^{1/20}$. Each non VB return satisfies (notice that $r_{l+3}(j) \geq c_{l+2}^{-1/2}$ when j is not a VB return)

$$X_{l+3}(j) \leq a^{1-5\epsilon^{1/2}} r_{l+3}(j) + v_{l+2} \leq a^{1-7\epsilon^{1/2}} r_{l+3}(j). \tag{7.89}$$

By the estimate of §7.2.2, we conclude that $\mu(I_\eta^\iota) \leq a^{1-10\epsilon^{1/2}}$ in Case A.

7.7.2. *Case B.* Let GOOD denote the set of $\underline{d} \in LS(l+1)$ such that $\underline{d} = (j_1, ..., j_m)$, and

$$\#\{i,\, X_{l+1}(j_i) \geq 1\} < a^{1-\epsilon} c_l^{-1} |\underline{d}|, \qquad \text{GOOD landing, Case B,} \tag{7.90}$$

$$\#\{i,\, X_{l+1}(j_i) \geq c_l^{-3\epsilon^{1/2}}\} < c_l^{l^3} |\underline{d}|, \qquad \text{GOOD landing, Case B.} \tag{7.91}$$

By (7.33), (7.61), and (7.78) we have that $a^{1-\epsilon} c_l^{-1} \geq (l+1)^2 x_{l+1}$ and $c_l^{l^3} \geq (l+1)^2 x_{l+1}(c_l^{-3\epsilon^{1/2}})$.

7.7.2.1. Let D_1 (respectively D_2) be the set of $\underline{d}$ such that $|\underline{d}| > c_{l+1}^{-1/2}$ and which do not satisfy (7.90) (respectively (7.91)). The argument of §7.5.1.3 with $q_0 = a^{1-\epsilon} c_l^{-1}$ (respectively, $q_0 = c_l^{l^3}$) implies that (notice that in both cases $q_0 > c_{l+1}^{1/5}$)

$$C(D_1, l+1), C(D_2, l+1) < e^{-c_{l+1}^{-1/4}}. \tag{7.92}$$

We conclude

$$C(\{\underline{d} \text{ is not GOOD}\}, l+1) \leq c_{l+1}^{2/7}. \tag{7.93}$$

7.7.2.2. Let us say that a return j of level $l+2$ is BAD if $R_{l+1}(I_{l+2}^j) \subset C_{l+1}^{\underline{d}}$ where $\underline{d}$ is not GOOD. Notice that $I(\{j \text{ is BAD}\}, l+2) < c_{l+1}^{1/20}$.

If j is not BAD, with $R_{l+1}(I_{l+2}^j) \subset C_{l+1}^{\underline{d}}$, let us consider two subcases. If $c_{l+1}^{\epsilon^{1/2}} \leq a < c_l^{\sqrt{l}}$, using (7.90) we get

$$X_{l+2}(j) \leq (c_l^{-14} c_l^{-1} a^{1-\epsilon}) |\underline{d}| + v_{l+1} \leq a^{1-2\epsilon} |\underline{d}| \leq a^{1-2\epsilon} r_{l+2}(j). \tag{7.94}$$

If $c_l^{\sqrt{l}} \geq a > c_l$, a similar estimate can be obtained using (7.90) and (7.91)

$$\begin{aligned} X_{l+2}(j) &\leq (c_l^{l^3} c_l^{-14} + c_l^{-3\epsilon^{1/2}} c_l^{-1} a^{1-\epsilon}) |\underline{d}| + v_{l+1} \leq (c_l^{l^3-14} + c_l^{-1-3\epsilon^{1/2}} a^{1-\epsilon} + c_{l+1}^{1/2-\delta}) |\underline{d}| \\ &\leq (c_l^{-1-5\epsilon^{1/2}} a^{1-\epsilon}) |\underline{d}| \leq a^{1-7\epsilon^{1/2}} l_{l+1}(\underline{d}) \leq a^{1-7\epsilon^{1/2}} r_{l+2}(j). \end{aligned} \tag{7.95}$$

Thus we have in both subcases

$$X_{l+2}(j) \leq a^{1-7\epsilon^{1/2}} r_{l+2}(j), \quad j \text{ is not BAD, Case B.} \tag{7.96}$$

By the argument of §7.2.2, we conclude that $\mu(I_\eta^\iota) \leq a^{1-10\epsilon^{1/2}}$ in Case B.
This concludes the proof of the main estimate.

8. Pathological laminations and other consequences

8.1. **Laminations in spaces of analytic unimodal maps.**

8.1.1. Let $\mathbb{F}$ be a Banach space. A codimension-one *holomorphic lamination* $\mathcal{L}$ on an open subset $\mathcal{W} \subset \mathbb{F}$ is a family of disjoint codimension-one Banach submanifolds of $\mathbb{F}$, called the *leaves* of the lamination such that for any point $p \in \mathcal{W}$, there exists a holomorphic local chart $\Phi : \tilde{\mathcal{W}} \to \mathcal{V} \oplus \mathbb{C}$, where $\tilde{\mathcal{W}} \subset \mathcal{W}$ is a neighborhood of p and $\mathcal{V}$ is an open set in some complex Banach space $\mathbb{E}$, such that for any leaf L and any connected component L_0 of $L \cap \mathcal{W}$, the image $\Phi(L_0)$ is a graph of a holomorphic function $\mathcal{V} \to \mathbb{C}$.

The local theory of codimension-one holomorphic laminations coincide with the theory of holomorphic motions (see [ALM], §2.5 and references therein). It follows from the λ-Lemma that holonomy maps of holomorphic laminations have quasiconformal extensions.

8.1.2. For $a > 0$, let $\Omega_a \subset \mathbb{C}$ be the set of z at distance at most a of I. Let $\mathcal{E}_a$ be the space of even holomorphic maps $f : \Omega_a \to \mathbb{C}$, continuous up to $\overline{\Omega}_a$. We endow $\mathcal{E}_a$ with the sup norm. Let $\mathcal{A}_a = \{f \in \mathcal{E}_a, f(-1) = f(1) = 1\}$ and let $\mathcal{A}_a^{\mathbb{R}} = \{f \in \mathcal{A}_a,\, f(z) = \overline{f(\overline{z})}\}$. Let $\mathcal{U}_a$ be the space of analytic quasiquadratic maps which belong to $\mathcal{A}_a^{\mathbb{R}}$.

One of the main results of [ALM] is that the partition of $\mathcal{U}_a$ on topological conjugacy classes has the structure of a codimension-one analytic lamination "almost everywhere".

Theorem 8.1 (Theorem A of [ALM]). *Let $f \in \mathbb{U}_a$ be a Kupka-Smale quasiquadratic map. There exists a neighborhood $\mathcal{V} \subset \mathcal{A}_a$ of f endowed with a codimension-one holomorphic lamination $\mathcal{L}$ (also called hybrid lamination) with the following properties:*

(1) the lamination is real-symmetric;

(2) if $g \in \mathcal{V} \cap \mathcal{A}_a^{\mathbb{R}}$ is non-regular, then the intersection of the leaf through g with $\mathcal{A}_a^{\mathbb{R}}$ coincides with the intersection of the topological conjugacy class of g with $\mathcal{V}$;

(3) Each $g \in \mathcal{V} \cap \mathcal{A}_a^{\mathbb{R}}$ belongs to some leaf of $\mathcal{L}$.

(See also [AM3] for the non-quasiquadratic case.)

Notice that the set of non-Kupka-Smale maps is contained on a countable union of codimension-one analytic submanifolds.

The lamination $\mathcal{L}$ has automatically quasisymmetric holonomy. Quasisymmetric maps are not always absolutely continuous (even though quasiconformal maps are). It turns out that $\mathcal{L}$ is very far from being absolutely continuous, at least at the set of non-regular leaves (the lamination restricted to regular maps is not uniquely defined, but can be chosen in a quite natural way to be locally analytic).

8.1.3. Let $\hat{\mathcal{L}}$ be the lamination consisting of the non-isolated non-regular leaves of $\mathcal{L}$. If f_λ is a one-dimensional family transversal to $\hat{\mathcal{L}}$, it intersects $\hat{\mathcal{L}}$ in a positive measure set[9].

Let $X \subset \mathcal{U}_a$ be the set of Collet-Eckmann maps satisfying the conclusion of Theorem A. Then X intersects each leaf of $\hat{\mathcal{L}}$ in a set of maps which are analytically conjugate in the attractor, and this is a set of infinite codimension (possibly empty): just notice that we can vary the exponent of any finite number of periodic orbits independently.

Thus, $\hat{\mathcal{L}}$ exhibits the same pathology described by Milnor in [Mi]: a full measure set intersecting the leaves of a finite codimension lamination in tiny sets. (The example described by Milnor is also an analytic codimension-one lamination, on two dimensions, and the intersections of the leaves with the full measure set are points. In this finite dimensional setting, this translates in the complete failure of Fubini's Theorem).

[9]Since preperiodic combinatorics are dense in $\hat{\mathcal{L}}$, and the generic unfolding of preperiodic combinatorics generates a positive measure set of non-hyperbolic parameters (see for instance [T] for a proof of a more general statement).

8.1.4. Although in our description we have to make use of transverse measures (since our setting is infinite-dimensional), one can interpret this pathology by taking finite dimensional sections as follows.

Let $\{f_\lambda\}_{\lambda\in\Lambda}$ be a small analytic k-dimensional transverse section to $\hat{\mathcal{L}}$. The lamination $\hat{\mathcal{L}}$ induces a lamination $\hat{\mathcal{L}}_\Lambda$ on Λ. Notice that $\hat{\mathcal{L}}_\Lambda$ has positive k-dimensional Lebesgue measure.

For $\lambda_1, \lambda_2 \in \Lambda$ distinct, let $P(\lambda_1, \lambda_2)$ be the number of periodic orbits of f_{λ_1} which have the same exponent of a periodic orbit of f_{λ_2}. A transversality argument shows that, for most of those sections (actually the complement has infinite-codimension), $P(\lambda_1, \lambda_2) < \infty$ whenever $\lambda_1 \neq \lambda_2$. For any family f_λ with this property, we obtain the same phenomena, but in k dimensions: the set of parameters X_Λ which are Collet-Eckmann and satisfy the conclusion of Theorem A intersects each leaf of $\hat{\mathcal{L}}_\Lambda$ in at most one point.

8.1.5. We point out that the set of *recurrent* parameters which do not satisfy the conclusion of Theorem A (or Theorem B) has Hausdorff dimension one in any one-dimensional transversal f_λ to $\hat{\mathcal{L}}$. Indeed, using the previous argument, we can select another transversal $\tilde{f}_\lambda$ arbitrarily close to f_λ, and such that for any λ_1, λ_2, the number of periodic orbits of f_{λ_1} which have the same exponent as some periodic orbit of $\tilde{f}_{\lambda_2}$ is finite. Let $\tilde{X}$ be the set of parameters satisfying the conclusion of Theorems A and B for $\tilde{f}_\lambda$. Let h be the holonomy map from f_λ to $\tilde{f}_\lambda$. Then the quasisymmetric constant of h (and thus the Hölder constant) is close to 1, provided $\tilde{f}_\lambda$ is close to f_λ (by the λ-Lemma). Since $\tilde{X}$ has positive Lebesgue measure, $h^{-1}(\tilde{X})$ has Hausdorff dimension close to one. But parameters λ in $h^{-1}(\tilde{X})$ do not satisfy the conclusion of Theorem A (or Theorem B, using §8.1.6): otherwise each of the infinitely many periodic orbits in the attractor of f_λ would have the same exponent of the corresponding (by the topological conjugacy) periodic orbit for $\tilde{f}_{h(\lambda)}$.

Remark 8.1. It is easy to see that the conclusions of Theorems A, B, and C fail for all Misiurewicz (non-recurrent Collet-Eckmann) parameters. This set of parameters has Hausdorff dimension one in any one-dimensional transversal to $\hat{\mathcal{L}}$.

8.1.6. By [MM], stochastic unimodal maps satisfying the conclusion of Theorem A are geometrically rigid: two such maps are smoothly (and automatically analytically) conjugate on the attractor. The same conclusion can be obtained for maps satisfying the conclusion of Theorem B. Indeed, the asymptotic distribution of the critical orbit, if it exists, is a topological invariant, and Theorem B implies that the conjugacy must be absolutely continuous on the support of the invariant measure (the attractor). The conjugacy is then easily promoted to being smooth (and automatically analytic) by a well known argument, see Exercise 3.1, Chapter V, page 375 in [MS].

We should point out that this conjugacy is not, in general, analytic on the whole interval I, as can be shown by a simple example.

Example 8.1. Let us consider the families

$$f_a(x) = ax(1-x), \tag{8.1}$$

$$g_a(x) = \frac{2}{\pi}\sin^{-1}\left(\frac{\sqrt{a}}{2}\sin(\pi x)\right). \tag{8.2}$$

Then for f_a and g_a are analytic families of quasiquadratic maps on the interval $[0,1]$ for $2 < a < 4$. Notice that

$$h_a \circ g_a = f_a \circ h_a, \quad h(x) = \frac{1-\cos(\pi x)}{2}, \tag{8.3}$$

so f_a and g_a are analytically conjugate on $(0,1)$, and the holonomy map between both families is trivial. Whenever f_a is Collet-Eckmann and satisfies the conclusion of Theorem A, g_a also does.

However, f_a and g_a are not analytically conjugate on $[0,1]$: indeed, $Df_a(0) = a$ and $Dg_a(0) = \sqrt{a}$, so the exponent of the fixed point 0 is not preserved.

8.2. **Formula for the exponent of μ_f.** In order to compute the Lyapunov exponent of μ_f combinatorially, one just has to find an expression for the Lyapunov exponent of the critical value. There are several ways to proceed, for instance, one can find convenient approximations of the critical orbit by periodic orbits and apply Theorem A. However, there exists a very simple expression using the combinatorics of the principal nest, whose proof only involves Corollary 1.1 and the asymptotic limits of §4.4.

Theorem 8.2. *Let f_t be an analytic family of unimodal maps. For almost every non-regular parameter, the Lyapunov exponent of μ_{f_t} (which is equal to the Lyapunov exponent of the critical value) is given by an explicit combinatorial formula:*

$$\lambda(\mu_{f_t}) = \lim_{n\to\infty} \frac{2\ln v_{n+1}}{v_n}, \tag{8.4}$$

where, as usual, v_n is the return time of the critical point to the n-th level of the principal nest.

Proof. Notice that

$$\lambda(f(0)) = \lim \frac{\ln |Df^{v_n-1}(f(0))|}{v_n - 1}, \tag{8.5}$$

and by Lemma 4.5,

$$\lim \frac{\ln v_{n+1}}{\ln c_n^{-1}} = 1. \tag{8.6}$$

Thus, we only have to show that

$$\lim \frac{\ln |Df^{v_n-1}(f(0))|}{\ln c_n^{-1}} = 2. \tag{8.7}$$

Notice that f^{v_n-1} takes $f(I_{n+1})$ to $R_n(I_{n+1})$ with torrentially small distortion. By Lemma 4.6, we have $n^{-2}|I_n| \leq |R_n(I_{n+1})| \leq |I_n|$. So we conclude

$$\lim \frac{\ln |Df^{v_n-1}(f(0))|}{\ln c_n^{-1}} = \lim \ln\left(\frac{|I_n|}{|I_{n+1}|^2}\right)\frac{1}{\ln c_n^{-1}} = \lim \ln\left(\frac{1}{|I_n| c_n^2}\right)\frac{1}{\ln c_n^{-1}} = 2, \tag{8.8}$$

since $|I_n| > c_{n-1}^{1+\delta}$ for n big. □

References

[ALM] A. Avila, M. Lyubich, W. de Melo. Regular or stochastic dynamics in real analytic families of unimodal maps. Preprint IMS Stony Brook 2001/15. To appear in Invent. Math.

[AM1] A. Avila, C. G. Moreira. Statistical properties of unimodal maps: the quadratic family. Preprint (http://www.arXiv.org). To appear in Annals of Math.

[AM2] A. Avila, C. G. Moreira. Statistical properties of unimodal maps: smooth families with negative Schwarzian derivative. Preprint (http://www.arXiv.org). To appear in Astérisque.

[AM3] A. Avila, C. G. Moreira. Phase-Parameter relation and sharp statistical properties for general families of unimodal maps. Preprint (http://www.arXiv.org).

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

[BL] Blokh, A. M.; Lyubich, M. Yu. Measurable dynamics of S-unimodal maps of the interval. Ann. Sci. École Norm. Sup. (4) 24 (1991), 545–573.

[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.

[K1] O. S. Kozlovski. Getting rid of the negative Schwarzian derivative condition. Ann. Math. 152 (2000), 743–762.

[K2] O.S. Kozlovski. Axiom A maps are dense in the space of unimodal maps in the C^k topology. Ann. Math. 157 (2003), 1–43.

[Li] Livšic, A. N. The homology of dynamical systems. Uspehi Mat. Nauk 27 (1972), no. 3(165), 203–204.
[L1] M. Lyubich. Combinatorics, geometry and attractors of quasi-quadratic maps. Ann. Math, **140** (1994), 347-404. Note on the geometry of generalized parabolic towers. Manuscript (2000) (http://www.arXiv.org).
[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. Astérisque, v. 261 (2000), 173 - 200.
[L4] M. Lyubich. Almost every real quadratic map is either regular or stochastic. Ann. Math. 156 (2002), 1-78.
[M] R. Mañé. Hyperbolicity, sinks and measures for one-dimensional dynamics. Comm. Math. Phys. 100 (1985), 495-524.
[MM] M. Martens, W. de Melo. The multipliers of periodic points in one-dimensional dynamics. Nonlinearity 12 (1999), 217–227.
[MS] W. de Melo, S. van Strien. One-dimensional dynamics. Springer, 1993.
[Mi] J. Milnor. Fubini foiled: Katok's paradoxical example in measure theory. Math. Intelligencer 19 (1997), 30–32.
[MT] J. Milnor, W. Thurston. On iterated maps of the interval. "Dynamical Systems", Proc. U. Md., 1986-87, ed. J. Alexander, Lect. Notes Math., v. 1342 (1988), 465-563.
[NS] T. Nowicki, D. Sands. Non-uniform hyperbolicity and universal bounds for S-unimodal maps. Invent. Math. 132 (1998), 633–680.
[RW] D. Ruelle, A. Wilkinson. Absolutely singular dynamical foliations. Comm. Math. Phys. 219 (2001), 481–487.
[ShSu] M. Shub, D. Sullivan. Expanding endomorphisms of the circle revisited. Erg. Th. Dyn. Sys. 5 (1985), 285–289.
[SW] M. Shub, A. Wilkinson. Pathological foliations and removable zero exponents. Invent. Math. 139 (2000), 495–508.
[T] M. Tsujii. Positive Lyapunov exponents in families of one dimensional dynamical systems. Invent. Math. 111 (1993), 113-137.

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`