STATISTICAL PROPERTIES OF UNIMODAL MAPS

Physical Measures, Periodic Orbits and Pathological Laminations

by ARTUR AVILA⋆ *and* CARLOS GUSTAVO MOREIRA⋆⋆

ABSTRACT

We consider typical analytic unimodal maps which possess a chaotic attractor. Our main result is an explicit combinatorial formula for the exponents of periodic orbits. Since the exponents of periodic orbits form a complete set of smooth invariants, the smooth structure is completely determined by purely topological data ("typical rigidity"), which is quite unexpected in this setting. It implies in particular that the lamination structure of spaces of analytic unimodal maps (obtained by the partition into topological conjugacy classes, see [ALM]) is not transversely absolutely continuous. As an intermediate step in the proof of the formula, we show that the distribution of the critical orbit is described by the physical measure supported in the chaotic attractor.

CONTENTS

1. Introduction

A *unimodal map* is a smooth (at least C^2) map $f : I \to I$, where $I \subset \mathbf{R}$ 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 2. 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 $p_a(x) = a - x^2$, $-1/4 \le a \le 2$. 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-

⋆ Partially supported by the French-Brazilian Agreement in Mathematics.

⋆⋆ Partially supported by Faperj and CNPq, Brazil.

DOI 10.1007/s10240-005-0033-2

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 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: orbits in an open and dense set of full Lebesgue measure converge to one of finitely many hyperbolic attracting periodic cycles, and the complementary set is hyperbolic expanding. Due to the works of Jakobson [J] and Benedicks-Carleson [BC], we know that 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 [L5], [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].

To describe our results, we will need the following from the picture that emerged from those works: a typical non-regular unimodal map f possesses a unique *transitive finite union of intervals*, A_f, and a unique *invariant probability measure absolutely continuous with respect to Lebesgue measure*, μ_f. Moreover A_f is the support of μ_f and periodic points are dense in A_f.

We note that the definition of A_f is given in terms of the topological dynamics of f. We will call A_f the *non-trivial attractor* of f and μ_f the *non-trivial physical measure* of f.

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. It can be visualized in terms of the regularity properties of a certain codimension-one lamination constructed in [ALM]: the resulting rather amusing picture is related to some recently discovered examples of measure-theoretical pathological laminations (Katok's

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

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

"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. *The formula.* — 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 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 Livšic [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 smooth and combinatorial invariants for typical non-regular analytic unimodal maps.

Theorem **1.** — *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_λ 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 2 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 3 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 (the orbit of) a point x is given by a probability measure μ (or equivalently, x is in the basin of μ, or x is typical for μ) if for any continuous function $\phi : I \to \mathbf{R}$

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

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 **2.** — *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.** — *In the setting of Theorem 2, 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 2 to Corollary 1, since $\ln |Df|$ is not continuous.

Previous progress in the direction of Theorem 2 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 1 we are interested in the exponents of (repelling) periodic orbits. More generally, one is led to ask about the geometry of (invariant) hyperbolic subsets $K \subset A_f$ (those are often Cantor sets).

In order to apply Theorem 2 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 (but, for typical maps, is always in L^p, for $1 \leq p < 2$, see §6.1).

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 $\mathscr{B}(K)$ of the gaps of K. The sigma-algebra $\mathscr{B}(K)$ 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 **3.** — *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 $d\mu_{f_\lambda}^K \in L^p$, for every $1 \leq p < \infty$.*

One can see this estimate (together with Theorem 2) as a generalization of Theorem 1, since it allows to compute using μ_f (which, due to Theorem 2 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 1 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 they actually lead us to the opposite end of rigid systems: the "moduli space of smooth structures" (in a fixed topological class) is infinite dimensional[3]. The usual examples of geometrically rigid systems, Diophantine irrational rotations and Feigenbaum attractors[4] 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[5].

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[6]. 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, and possibly empty). However, the set of typical non-regular unimodal maps (satisfying the conclusion of Theorem 1) 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.)

Remark. — Let us point out that one does not need the full power of Theorem 1 to prove that the lamination of [ALM] is not absolutely continuous. Indeed, Theorem 2 already implies "typical rigidity" (though in a less explicit way), see §8.1.6.

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 quasisymmet-

[3] In the case of maps f admiting a (topological) attractor A_f which is a cycle of intervals and which contains a dense (and hence infinite) set of periodic orbits.

[4] Since this paper is concerned only with the typical unimodal maps, it does not touch the very interesting issue of rigidity of attractors of Feigenbaum maps (unimodal maps which are infinitely renormalizable of bounded type), since those maps are relatively rare in parameter space. See [L4] for a thorough account and further references.

[5] 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 countable union of codimension-one analytic varieties.

[6] 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.

ric 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 properties 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 1 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 the 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. This formula is quite simple, but is given 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 1, 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. 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 1 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 second 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, especially the quadratic family, was heavily tied to the possibility of the intertwined use of real and complex techniques. Many of the most beautiful aspects of the theory of unimodal maps (particularly with respect to connections to different fields) show up only when one complexifies the dynamics.

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 work of Lyubich.

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.

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 1, assuming the validity of Theorems 2 and 3.

In §4, we discuss the combinatorics of the principal nest and introduce our basic tool to make parameter estimates: the Phase-Parameter relation. We then present some of the estimates obtained in [AM1].

In §5, we prove Theorem 2. The proof is technical but has a clear strategy, which we describe in §5.1. In §6, we reduce Theorem 3 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.

In §8, we describe in more detail some of the consequences of Theorems 1, 2 and 3 discussed in the introduction (rigidity, singularity of the lamination, and a formula for the Lyapunov exponent of the physical measure).

Acknowledgement

The authors would like to thank Misha Lyubich and Jean-Christophe Yoccoz for their continuous encouragement and the referee for his detailed comments.

2. Preliminaries

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

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

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

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

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 $h : \mathbf{R} \to \mathbf{R}$ such that there exists a constant k such that for any $x \in \mathbf{R}$, $a > 0$,

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

Equivalently, h is quasisymmetric if it has a real-symmetric quasiconformal extension to the whole $\mathbf{C}$ (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[7]. 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.

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

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

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

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

where $h : \mathbf{R} \to \mathbf{R}$ 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

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

We will sometimes use the notation $p(X|T) = p_1(X|T) = |X \cap T|/|T|$.

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

A map $f \in \mathbf{U}^2$ 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 $\{f_\lambda \in \mathbf{U}^2\}_{\lambda \in \Lambda}$ (where $\Lambda \subset \mathbf{R}$ 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].)

Recall that, by [K2], regular maps are dense in both smooth and analytic topologies. Let us say that an analytic family of unimodal maps is *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 (this is clear from Theorem A of [ALM] and also by using Kozlovski's trick of [K2]).

2.4. *Renormalization.* — Let $f \in \mathbf{U}^2$. 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*[8] 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 $\mathscr{F} \subset \mathbf{U}^2$ 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 **4** *(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 1, 2, or 3 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.

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 **1** *(Koebe Principle, see [MS], page 258). — Let $f : T \to \mathbf{R}$ 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 $\lim_{\epsilon\to\infty} \delta(\epsilon) > \frac{9}{100}$.*

The Koebe Principle gives control on the inverse branches of maps with negative Schwarzian derivative (since such inverse branches have positive Schwarzian derivative).

---

[8] 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.

Due to the recent results of Kozlovski, we know that the condition of negative Schwarzian is not needed for application of the Koebe Principle (for unimodal maps in $\mathbf{U}^3$ 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.** — *Let $f \in \mathcal{F}$ be a quasiquadratic map, and let* T *be a nice interval. There exist constants* $C > 0$, $\lambda > 1$ *such that if* $f^k(x) \in I \setminus T$, $k = 0, \ldots, m-1$ *then* $|Df^m(x)| > C\lambda^m$.

Corollary **2.** — *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

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

in the weak* topology, where δ_y denotes the Dirac mass on y. 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 $A = T_0 \cup \ldots \cup T_{k-1}$ 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. — 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 **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 **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. — 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 3 a reasonable estimate for $d\mu$. Notice also that our proof of Theorem 3 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.[9]

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 $\Sigma^{\mathbf{N}}$ 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 $r \in \mathbf{N} \cup \{\infty\}$, 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

[9] Since the critical value (which is associated to a square-root singularity for the density of the physical measure) belongs to an invariant hyperbolic Cantor set.

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 $\Theta:\mathrm{I}\to\Sigma$ be defined by $\Theta|[-1,0)=0$, $\Theta|(0,1]=1$, and $\Theta(0)=c$.

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

The (discontinuous) map $\theta:\mathrm{I}\to\Sigma^{\mathbf{N}}$ 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 $\mathrm{I}_\alpha\subset\mathrm{I}$ 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 1 assuming Theorems 2 and 3.* — We will actually prove the following stronger:

Theorem **7.** — *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 $1\le p<\infty$, $d\mu_f^{\mathrm{K}}\in\mathrm{L}^p$.*

Then $\theta(0)$ *is geometrically normal and for any* $z \in \mathrm{A}$ *periodic (of period m),*

$$|\mathrm{D}f^m(z)| = \rho(\overline{\theta(z)}, \theta(0))^{-1}. \tag{3.5}$$

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

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

Let $z \in \mathrm{A}$ be a periodic orbit, and let $\alpha = \overline{\theta(z)}$. By item 1, we conclude that z is repelling, and since f is quasiquadratic, $\cap \mathrm{I}_{\alpha^k} = \{z\}$, and the length m of α is the period of z. Let $q, q' \in \mathrm{I}_\alpha$ be periodic orbits in opposite sides of z, and let $q_k = (f^{km}|\mathrm{I}_{\alpha^{k+1}})^{-1}(q)$ and $q'_k = (f^{km}|\mathrm{I}_{\alpha^{k+1}})^{-1}(q')$. Let K be the hyperbolic set consisting of z, the forward orbit of q and q' and all q_k and q'_k. Let $\mathrm{T}_k = [q'_k, q_k]$. It is easy to see that there exists $j > 0$ such that for all $k > j$,

$$\mathrm{T}_{k+j} \subset \mathrm{I}_{\alpha^k} \subset \mathrm{T}_{k-j}. \tag{3.6}$$

In particular,

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

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

By Theorem 6, there exists a constant $\mathrm{C} > 0$ such that $d\mu|\mathrm{A} \ge \mathrm{C}$. On the other hand, $\mathrm{T}_k \subset \mathrm{A}$ for k big enough, so $\mu(\mathrm{T}_k) \ge \mathrm{C}|[q'_k, q_k]|$. It is clear that

$$\lim_{n\to\infty} |\mathrm{T}_n|^{1/n} = |\mathrm{D}f^m(z)|^{-1}, \tag{3.9}$$

so $\rho^-(\alpha, \theta(0)) \ge |\mathrm{D}f^m(z)|^{-1}$.

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

$$\left(\int_{\mathrm{T}_k} (d\mu^{\mathrm{K}})^p\right)^{1/p} \le \mathrm{C}_p. \tag{3.10}$$

In particular, by the Hölder inequality,

$$\mu(\mathrm{T}_k) = \int_{\mathrm{T}_k} d\mu^{\mathrm{K}} \le \left(\int_{\mathrm{T}_k} (d\mu^{\mathrm{K}})^p\right)^{\frac{1}{p}} \left(\int_{\mathrm{T}_k} 1^{\frac{p}{p-1}}\right)^{1-\frac{1}{p}} \le \mathrm{C}_p|\mathrm{T}_k|^{1-\frac{1}{p}}. \tag{3.11}$$

Taking $k \to \infty$ we get

$$\rho^+(\alpha, \theta(0)) \le |\mathrm{D}f^m(z)|^{-1+\frac{1}{p}}. \tag{3.12}$$

Since $1 \le p < \infty$ 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 z in A. In the first case, obviously $\rho(\alpha, \theta(0)) = 0$. In the second case, by the previous discussion, $\rho(\alpha, \theta(0)) = |Df^m(z)|^{-1} > 0$, where m is the length of α. In particular, $\theta(0)$ is geometrically normal. □

Remark. — 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 $f \in \mathscr{F}$ 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 **1.** — *Let us say that a family f_λ of quasiquadratic maps satisfies the Topological Phase-Parameter relation at a parameter λ_0 if $f = f_{\lambda_0} \in \mathscr{F}$, 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 $(R_i | I_i^0)^{-1}(C_i^{\underline{d}}) = I_{i+2}$.

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

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

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

Definition **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 **8** *(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 **3** *(Lemma 3.1 of [AM1]). — Let* $X \subset \mathbf{R}$ *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* $\mathbf{R}$ *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.** — *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 **5.** — *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 3, 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

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

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

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

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

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

Lemma **6.** — *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 **7.** — *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* $\underline{d} \in \Omega$, $\operatorname{dist}(R_n^{\underline{d}}) \leq n^{\frac{1}{2}+\delta}$.

The estimate above for distortion of branches $R_n^{\underline{d}}$ is pessimistic in a sense. 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 $\operatorname{dist}(f|I_n^j) - 1$ 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 $\operatorname{dist}(f|I_n^j) = 1 + O(c_{n-1})$. 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

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

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

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

Lemma **8.** — *For almost every non-regular map, if* $\epsilon_0 = \epsilon(\gamma) < 1/100$, *then, for n large enough*

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

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

$$\textbf{(4.9)} \qquad p_\gamma\left(r_n(x) > c_{n-1}^{-1-2\epsilon_0} \,\middle|\, I_n\right) \leq e^{-c_{n-1}^{-\epsilon_0/5}}.$$

5. The critical orbit is typical

5.1. *Outline.* — Let us summarize the main steps in the proof of Theorem 2.
(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).

This argument is detailed in §5.4.1.

(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 $\lim_{n\to\infty} p(|\psi_n^\xi(x) - q| > \epsilon | I_n) = 0$.

This step is carried out in §5.4.2.

(3) We use an explicit Large Deviation Estimate (this key estimate is Proposition 2) 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.

This step is carried out in Proposition 3.

(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 sense) 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.

The arguments related to this step are developed in §5.2.

(6) As a consequence of (4) and (5), 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 $|\psi_{n_s}^\xi - \mu(\xi)| < \epsilon$. As a consequence, those "very good" return branches of level n satisfy $|\psi_n^\xi - \mu(\xi)| < 2\epsilon$. As a bonus 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, see Proposition 1.

(7) Using the Phase-Parameter relation we make the critical point fall in "very good" branches (Lemmas 21 and 22). Thus the distribution of the critical orbit on ξ is 2ϵ close to $\mu(\xi)$. Making ϵ go to 0 we obtain Theorem 2.

5.2. *Inductive estimates.* — In this section we will show that a small (in the quasi-symmetric sense) set of branches of level n_0 has a small effect on most (in the quasi-symmetric 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.

Remark. — The estimates we will obtain in this section hold for all parameters satisfying the estimates of §4.4 (thus there is no parameter exclusion going on in this section).

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 $\mathrm{X} \subset \mathrm{I}$ is a Borelian set, we let

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

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 $(\gamma_0, \mathrm{C}\frac{n+1}{n})$-homeomorphism is a $\hat{\gamma}$-qs homeomorphism. Let $\mathrm{C}_n = \mathrm{C} \cdot \frac{n+1}{n}$, $\tilde{\mathrm{C}}_n = \mathrm{C}\frac{2n+3}{2n+1}$.

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 $\tilde{\mathrm{C}}_n$. We will use (γ, C_n) capacities to estimate the size of sets of return branches of level n and $(\gamma, \tilde{\mathrm{C}}_n)$ 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 **9** *(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 $\mathrm{X} \subset \mathrm{I}_n$ then*

$$p_{\gamma,\tilde{\mathrm{C}}_n}\left(\left(\mathrm{R}_n^d\right)^{-1}(\mathrm{X})\middle|\mathrm{I}_n^d\right) \leq 2^n p_{\gamma,\mathrm{C}_n}(\mathrm{X}|\mathrm{I}_n). \tag{5.2}$$

And if $X \subset I_n$ *and*

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

then

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

Induction applied to (5.2) gives:

Lemma **10** *(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* $Q_n \subset \mathbf{Z}$ *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*

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

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

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*

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

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.

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

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

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

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

Lemma **11** *(Analogous to Lemma 7.1 of [AM1]). — With total probability we have*

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

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

Let $T \subset \mathbf{Z}$ be given. Let us define $VG(T, n_0, n) \subset \mathbf{Z}$ and $LE(T, n_0, n) \subset \Omega$ inductively as follows. Let $VG(T, n_0, n_0) = \mathbf{Z} \setminus T$. 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 $\underline{d} = (j_1, \ldots, j_m)$ and

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

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

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

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

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

The class $VG(T, n_0, n)$ is designed so that those 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 $VG(T, n_0, n)$ corresponds to most branches of level n (by induction). Those two estimates are given below:

Lemma **12** *(see also Lemma 7.2 of [AM1]). — With total probability, for all n_0 sufficiently big, if* T *satisfies (5.14) for some $1 \le \gamma \le \gamma_0$ then for all $n \ge n_0$, we have*

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

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

Furthermore,

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

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

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

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

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

Lemma 11 we get

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

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. □

Lemma **13** *(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 **14.** *— 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

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

Indeed (using the notation of Lemma 13),

$$\sum_{\substack{i<k,\\ f^i(x)\in I_{n_0}^T}} r_{n_0}(f^i(x)) \leq k - \beta_k + \alpha_n\beta_k + c_{n-1}^{-14} \tag{5.25}$$

$$\leq \left(1 - \frac{\beta_k}{k} + \alpha_n + c_{n-1}^{-14}c_n^{2/n}\right) k \leq \alpha_{n+1} k.$$

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