



Probabilistic Universality in Two-Dimensional Dynamics

M. Lyubich, M. Martens

SUNY at Stony Brook, Stony Brook, USA. E-mail: mlyubich@math.stonybrook.edu; marco@math.sunysb.edu

Received: 14 September 2017 / Accepted: 15 January 2021

Published online: 1 April 2021 – © The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021

Abstract: In this paper we continue to explore infinitely renormalizable Hénon maps with small Jacobian. It was shown in a previous paper by the authors, joint with A. de Carvalho, (J Stat Phys 121(5/6):611–669, 2005) that contrary to the one-dimensional intuition, the Cantor attractor of such a map is non-rigid and the conjugacy with the one-dimensional Cantor attractor is at most $1/2$ -Hölder. Another formulation of this phenomenon is that the scaling structure of the Hénon Cantor attractor differs from its one-dimensional counterpart. However, in this paper we prove that the unique invariant measure on the attractor assigns a weight to these bad spots which tends to zero on microscopic scales. This phenomenon is called *Probabilistic Universality*. It implies, in particular, that the Hausdorff dimension of the invariant measure on the attractor is universal. In this way, universality and rigidity phenomena of one-dimensional dynamics assume a probabilistic nature in the two-dimensional world.

Contents

1. Introduction	1296
2. Outline	1298
3. Preliminaries	1309
4. Regular Pieces	1317
5. Sticks	1322
6. Scaling	1325
7. Universal Sticks	1328
8. Probabilistic Universality	1339
9. Recovery	1344
10. Probabilistic Rigidity	1346
Appendix: Open Problems	1358
References	1359

1. Introduction

Renormalization ideas have played a central role in Dynamics since the discovery of the Universality and Rigidity phenomena by Feigenbaum [11], and independently by Coulet and Tresser [6], in the mid 1970s. Roughly speaking, it means that different systems in the same “universality class” have the same small scale geometry. In the one-dimensional setting this phenomenon has been viewed from many angles (statistical physics, geometric function theory, Teichmüller theory, hyperbolic geometry, infinite-dimensional complex geometry) and by now has been fully and rigorously justified, see [8, 10, 14, 15, 19, 20, 22] and references therein.

In [6] Coulet and Tresser also conjectured that these phenomena would also be valid in higher dimensional systems, even in infinite dimensional situations. Indeed, computer and physical experiments that followed suggested that universality and rigidity hold in much more general context. The simplest test case for it is the dissipative Hénon family which can be viewed as a small perturbation of the one-dimensional quadratic family. However, it was shown in [7] that Universality and Rigidity break down already in this case. This puts in question the relevance of one-dimensional models for higher dimensional problems.

In this paper we provide a resolution of this unsatisfactory situation: namely, we show that for dissipative Hénon maps, small scale universality is actually valid in *probabilistic* sense, almost everywhere with respect to the invariant measure on the attractor. *Probabilistic universality* and *probabilistic rigidity* phenomena may be valid for higher dimensional (including infinite dimensional) systems which are contracting in all but one direction.

Let us now formulate our results more precisely. We consider a class of dissipative¹ Hénon-like maps on the unit box $B^0 = [0, 1] \times [0, 1]$ of form

$$F(x, y) = (f(x) - \varepsilon(x, y), x), \quad (1.1)$$

where $f(x)$ is a unimodal map with non-degenerate critical point and ε is small. It maps B^0 on a slightly thickened parabola $x = f(y)$. Such a map is called *renormalizable* if there exists a smaller quadrilateral box $B^1 \subset B^0$ around the tip of the parabola which is mapped into itself under F^2 . The *renormalization* for F is the map $RF = \Psi^{-1} \circ F^2 \circ \Psi$, where $\Psi : B^0 \rightarrow B^1$ is an explicit non-linear change of variable (“rescaling”) that brings F^2 to the normal form of type (1.1).

If RF is in turn renormalizable then F is called *twice renormalizable*, etc. In this paper we will be concerned with *infinitely renormalizable* Hénon-like maps. Such a map admits a nest of 2^n -periodic quadrilateral boxes $B^0 \supset B^1 \supset B^2 \supset \dots$ shrinking to the *tip* τ of F . The n th-renormalization level is the orbit $\mathcal{B}^n = \{B_i^n = F^i(B^n), i = 0, 1, \dots, 2^n - 1\}$. The quadrilateral boxes B_i^n are, for given n , pairwise disjoint topological disks. We obtain a hierarchy of such levels shrinking to the *Cantor attractor*

$$\mathcal{O}_F = \bigcap_{n=0}^{\infty} \bigcup_{i=0}^{2^n-1} B_i^n$$

on which F acts as the dyadic adding machine. In particular, the dynamics on $\mathcal{O}_F$ is uniquely ergodic, so we obtain a unique invariant measure μ supported on the attractor

¹ A map is called *dissipative* if the jacobian is nonnegative and smaller than 1 in each point of the domain. The map is *strongly dissipative* if the Jacobian is small enough such that the results from [7] can be applied.

$\mathcal{O}_F$. We define the *average Jacobian* of F as follows:

$$b_F = \exp \int_{\mathcal{O}_F} \ln \text{Jac } F d\mu.$$

Consider a strongly dissipative infinitely renormalizable Hénon-like map. The attractor of such a map is topologically conjugate to the attractor of the one-dimensional period doubling renormalization fixed point f_* . The geometry of a piece $B \in \mathcal{B}^n$ can be very different from the geometry of the corresponding piece I of the one-dimensional renormalization fixed point f_* . The pieces of the one-dimensional system are small intervals. Take a piece $B \in \mathcal{B}^n$ and the two pieces $B_1, B_2 \in \mathcal{B}^{n+1}$ with $B_1, B_2 \subset B$. Let I, I_1, I_2 be the corresponding pieces of f_* . The piece B of F has ϵ -precision if after applying one affine map $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ we have that the (Hausdorff) distance between I and $A(B)$, I_1 and $A(B_1)$, I_2 and $A(B_2)$ is at most $\epsilon \cdot \text{diam}(I)$. The triples $B_1, B_2 \subset B$ and $I_1, I_2 \subset I$ are geometrically almost the same. Indeed, most pieces of a given renormalization level are exponentially close to their one-dimensional counterpart. This is the content of Theorem 1.1 and Theorem 1.2.

Collect the pieces of the n th-level with ϵ -precision in

$$\mathcal{S}_n(\epsilon) = \{B \in \mathcal{B}^n \mid B \text{ has } \epsilon - \text{precision}\}.$$

Definition 1.1. The geometry of the Cantor attractor $\mathcal{O}_F$ of a dissipative infinitely renormalizable Hénon-like map is probabilistically universal if there exists $\theta < 1$ such that

$$\mu(\mathcal{S}_n(\theta^n)) \geq 1 - \theta^n.$$

Theorem 1.1 (Probabilistic universality). *The geometry of the Cantor attractor of a strongly dissipative infinitely renormalizable Hénon-like map is probabilistically universal.*

Definition 1.2. The Cantor attractor $\mathcal{O}_F$ of a dissipative infinitely renormalizable Hénon-like map is probabilistically rigid if the conjugation $h : \mathcal{O}_F \rightarrow \mathcal{O}_{f_*}$ to the attractor $\mathcal{O}_{f_*}$ of the one-dimensional renormalization fixed point f_* has the following property. There exist $\beta > 0$, and a sequence $X_1 \subset X_2 \subset X_3 \subset \dots \subset \mathcal{O}_F$ such that $h : X_N \rightarrow h(X_N) \subset \mathcal{O}_{f_*}$ is $(1 + \beta)$ -differentiable, and $\mu(X_N) \rightarrow 1$.

Theorem 1.2 (Probabilistic Rigidity). *The Cantor attractor of a dissipative infinitely renormalizable Hénon-like map is probabilistically rigid.*

The Cantor attractor $\mathcal{O}_F$ is not part of a smooth curve, see [7]. However, large parts of it, the sets

$$X_N = \bigcap_{n \geq N} \mathcal{S}_n(\theta^n)$$

where $\theta < 1$ is close enough to 1, satisfy

Theorem 1.3. *Each set $X_N \subset \mathcal{O}_F$ is part of a smooth $C^{1+\beta}$ -curve.*

Let μ_* be the invariant measure on $\mathcal{O}_{f_*}$, the attractor of the one-dimensional renormalization fixed point. A consequence of probabilistic rigidity is

Theorem 1.4. *The Hausdorff dimension is universal:*

$$HD_\mu(\mathcal{O}_F) = HD_{\mu_*}(\mathcal{O}_{f_*}).$$

The theory of universality and rigidity became a probabilistic geometric theory for Hénon dynamics.

We prove the above results by introducing the so-called *pushing-up* machinery. This method locates the pieces in the n th-renormalization level that have exponential precision. The difficulty is that the orbit between two such good pieces may pass through poor pieces, so one cannot recover all good pieces by simple iteration of the original map. Instead, the pushing-up machinery relates pieces in the same renormalization level by means of the diffeomorphic rescalings built into the notion of renormalization. The distortion of these rescalings can be controlled if the two pieces under consideration, viewed from an appropriate scale, do not lie “too deep” (in the sense precisely defined below). This machinery might have applications beyond the present situation.

For the reader’s convenience, the pushing-up machinery will be informally outlined in Sect. 2. Also, more special notations are collected in the Nomenclature. For a survey on Hénon renormalization see [17]. For early experiments and results on Hénon renormalization see [3, 5, 12].

2. Outline

This section will give an outline of the proof of probabilistic universality, Theorem 1.1 (and Theorem 2.3 in this outline). From a distant viewpoint the main ideas involved are as follows.

The renormalization process analyzes the geometry at one specific point of the Cantor set, i.e. the tip. It shows universal small scale geometry at this point. Unfortunately one can not reconstruct universal geometry in other areas of Cantor set by iteration, applying the map to small universal neighborhoods of the tip. Some of these iterates will have strong distortion. There will be areas in the Cantor with strongly distorted geometry. This is the non-rigidity explained in [7].

The $(k + 1)^{\text{th}}$ -renormalization $F_{k+1} = R^{k+1}F$ is obtained by a nonlinear rescaling ψ_v^k from the k^{th} -renormalization $F_k = R^k F$. The domain of F_k contains two dynamical pieces defining F_{k+1} . One is the image of the rescaling ψ_v^k and the other is the image of $\psi_c^k = F_k \circ \psi_v^k$, see Fig. 14. The Cantor set is not reconstructed by iteration but it seen as the invariant set of the iterated function system generated by the rescaling maps ψ_c^k and ψ_v^k , $k = 1, 2, \dots$, as illustrated in Fig. 14. In particular, given $n \geq 1$ the iterated function system is used to generate the n -dynamical partition $\mathcal{B}^n$ consisting of pieces

$$B_\omega^n = \text{Im}(\psi_{\omega_1}^1 \circ \psi_{\omega_2}^2 \circ \dots \circ \psi_{\omega_n}^n),$$

with $\omega = (\omega_1, \omega_2, \dots, \omega_n) \in \{c, v\}^n$.

The renormalizations F_k have a universal asymptotic form, see (2.5). Also the rescaling maps ψ_v^k have a universal form. And hence $\psi_c^k = F_k \circ \psi_v^k$ is well controlled. Even long compositions of the form

$$\Psi_k^l = \psi_v^k \circ \psi_v^{k+1} \circ \dots \circ \psi_v^l \quad (2.1)$$

are well controlled, see (2.6). However, compositions of maps in the iterated function system of the form

$$\psi_c^k \circ \psi_v^{k+1} \circ \dots \circ \psi_v^l$$

with

$$l - k > \alpha 2^k - A$$

might distort strongly. This is a “deep jump”. The part of the Cantor set with universal geometric properties corresponds with the part obtained by applying the iterated function systems avoiding these “deep jumps” with $l - k > \alpha 2^k - A$. This mechanism is the *pushing-up* regime.

Unfortunately, applying the iterated function system avoiding all deep jumps would not reach a full measure subset of the Cantor set. The remedy is to apply the pushing-up regime only on the scales

$$\ln n \leq k \leq (1 - q_0) \cdot n,$$

with $q_0 > 0$ small.

In the regime $(1 - q_0) \cdot n < k \leq n$ one applies all maps from the iterated function system. This corresponds to reconstruction of the Cantor set by iteration of the renormalization $F_{(1-q_0)n}$. This map is exponentially close to the one-dimensional renormalization fixed point. The number of iterates one needs is very small. Namely, $2^{q_0 n}$. This few iterates can not lead to large distortion. This is the *one-dimensional regime*.

In the regime $k \leq \ln n$ one applies all maps from the iterated function system. This corresponds to reconstruction of the Cantor set by iteration of the original map F . In this regime the number of iterates is proportional to n . The pieces of the Cantor set are exponentially small and these n iterates can not build up distortion too much. This is the *brute-force regime*.

By organizing the iterated function system in these three regimes one reached a subset of full measure in the Cantor set with universal small scale geometry, i.e. *probabilistic universality*.

The main technical aspects are the control of distortion during the pushing-up regime on one hand. And on the other hand the probabilistic aspect of avoiding deep jumps. There is also classical one-dimensional dynamics to control distortion in the one-dimensional regime. The brute-force regime does not require delicate analysis. It is brute force after all.

2.1. Infinitely renormalizable Hénon-like maps. We will start with outlining the set-up developed in [7, 16]—see Sect. 3 for details.

We consider a class $\mathcal{H} = \mathcal{H}(\bar{\varepsilon})$ of Hénon-like maps of the form

$$F: (x, y) \mapsto (f(x) - \varepsilon(x, y), x),$$

acting on the unit box $B^0 = [0, 1] \times [0, 1]$, where $f(x)$ is a unimodal map subject of certain regularity assumptions, and $\|\varepsilon\| < \bar{\varepsilon}$ is small (for an appropriate norm). If the unimodal map f is renormalizable then the renormalization $F_1 = RF \in \mathcal{H}$ is defined as $(\Psi_0^1)^{-1} \circ (F^2|_{B^1}) \circ \Psi_0^1$, where B^1 is a certain piece around the *tip*, a point which plays the role of the “critical value”, and $\Psi_0^1: \text{Dom}(F_1) \rightarrow B^1$ is an explicit *non-linear* change of variables.

Inductively, we can define n times renormalizable maps for any $n \in \mathbb{N}$, and consequently, *infinitely renormalizable* Hénon-like maps. For such a map the n -fold renormalization $F_n = R^n F \in \mathcal{H}$ is obtained as $(\Psi_0^n)^{-1} \circ (F^{2^n}|_{B^n}) \circ \Psi_0^n$, where B^n is an

appropriate *renormalization piece*, $\Psi_0^n : \text{Dom}(F_n) \rightarrow B^n$ is a non-linear change of variables.

These pieces B^n form a nest around the *tip* τ of F :

$$B^0 \supset B^1 \supset \dots \supset B^n \supset \dots \ni \tau$$

Taking the iterates $F^k B^n$, $k = 0, 1, \dots, 2^n - 1$, we obtain a family $\mathcal{B}^n$ of 2^n pieces $\{B_\omega^n\}$, called the *n*th *renormalization level*, that can be naturally labelled by strings $\omega \in \{c, v\}^n$ in two symbols, c and v , with $B_{v^n}^n \equiv B^n$. See Sect. 3 for details. Then

$$\mathcal{O}_F = \bigcap_n \bigcup_\omega B_\omega^n$$

is an attracting Cantor set on which F acts as the dyadic adding machine. This Cantor set carries a unique invariant measure μ . This allows us to introduce the most important geometric parameter attached to F , its *average Jacobian*

$$b_F = \exp \int_{\mathcal{O}_F} \ln \text{Jac } F \, d\mu.$$

Usually, we will denote the average Jacobian with b .

The size of the pieces decays exponentially:

$$\text{diam } B^n \asymp \sigma^n, \quad \text{diam } B_\omega^n \leq C \sigma^n, \quad (2.2)$$

where $\sigma \in (0, 1)$ is the universal scaling factor (coming from one-dimensional dynamics) while $C = C(\bar{\varepsilon})$ depends on F .

A surprising phenomenon discovered in [7] is that unlike its one-dimensional counterpart, the Cantor set $\mathcal{O}_F$ *does not have universal geometry*: it depends in an essential way on the average Jacobian b . However, the difference appears only in the scales much smaller than b : if the pieces B^n of level n have diameter much larger than b then the geometry of the pieces B_ω^n is controlled by one-dimensional dynamics: the pieces are aligned along the parabola $x = f(y)$ with thickness of order b . According to (2.2), this happens when

$$\alpha \sigma^n \geq b \quad (2.3)$$

with sufficiently small universal² $\alpha > 0$, i.e., when

$$n \leq c |\ln b| - A, \quad \text{where } c = \frac{1}{|\ln \sigma|}, \quad A = \frac{\ln \alpha}{\ln \sigma}. \quad (2.4)$$

We will call these levels *safe*. Compare (4.3). It turns out that the pieces in these levels are safe from being deformed by the dynamics, i.e. they are essentially very thin rectangles. However, some pieces in deeper levels have a shape far from a rectangle, they can be *U-shaped* or even worse.

² The constant $\alpha > 0$ can be chosen for a given compact family of Hénon-like maps. Renormalizations of strongly dissipative Hénon-like maps in the boundary of chaos eventually will be in such a compact set. In this sense, α is universal.

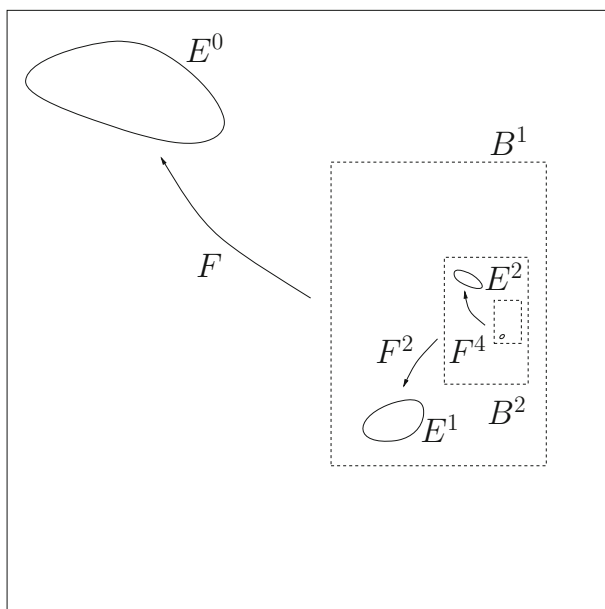


Fig. 1. An infinitely renormalizable Hénon-like map

2.2. Random walk model. The pieces of the n th-renormalization level B^n are obtained by the iteration of the renormalization piece B^n that contains the tip. However, in this way it is difficult to control their geometry. Instead we apply a geometric analysis of the renormalization levels based upon different rescalings involved in the renormalization scheme. Combinatorially we describe it in terms of a natural random walk model that gives a concise way of understanding probabilistic aspects of universality. “Very long jumps” in this model correspond to *unsafe* levels introduced above (2.4) that cause difficulties in the geometric analysis: they have to be avoided. We quantify this situation by a notion of *s-controlled* orbits in this random walk context (in this section) and by an equivalent notion of *not too deep* pieces in the geometric context (see Sect. 4 and in particular (4.3); see also Sect. 2.4 of this outline). This allows us to control geometry of the pieces almost everywhere leading to the Probabilistic Universality.

The *depth* of a piece and its backward closest approaches to the tip contain the essential combinatorial information needed for the analysis.

To any point $x \in \mathcal{O} \equiv \mathcal{O}_F$ we can assign its *depth*

$$\text{depth}(x) \equiv k(x) = \sup\{k : x \in B^k\} \in \mathbb{N} \cup \{\infty\}.$$

Compare to Sect. 4. Here the tip is the only point of infinite depth. If $\text{depth}(x) = k$ then $x \in E^k \equiv F^{2^k}(B^{k+1})$ (see Figs. 1, 2).

We say that a point $x \in \mathcal{O}$ is *combinatorially closer* to τ than $y \in \mathcal{O}$ if $k(x) > k(y)$. We will now encode any point $x \in \mathcal{O}$ by its *closest approaches* to τ in *backward* time. Namely, let us consider the backward orbit $\{F^{-t}x\}_{t=0}^\infty$, and mark the moments t_m ($m = 0, 1, \dots$) of closest approaches, i.e., at the moment t_m the point $x_m := F^{-t_m}x$ is combinatorially closer to τ than all previous points $F^{-t}x$, $t = 0, 1, \dots, t_m - 1$. Since the dynamics of F on $\mathcal{O}$ is the dyadic adding machine, this is an infinite sequence of

moments for any point $x \notin \text{orb}(\tau)$, i.e. any point not in the forward orbit of the tip. If $x = F^l(\tau)$, we terminate the code at the moment t . The encoding of a point $x \in \mathcal{O}$ is given by

$$k_m(x) = k(x_m), \quad m = 0, 1, \dots,$$

be the sequence of the corresponding depths. Obviously, both sequences, $\bar{t} = \{t_m\}$ and $\bar{k} = \{k_m\}$ are *strictly increasing*.

For any depth k , let us consider the *first return map* (see Figs. 1, 2).

$$G_k : B^{k+1} \rightarrow B^k, \quad G_k = F^{2^k},$$

and the *first landing map in backward time*

$$L_k : \bigcup_{m=0}^{2^k-1} F^m(B^k) \rightarrow B^k, \quad L_k(x) = F^{-m}x, \text{ for } x \in F^m(B^k).$$

Then we have by definition:

$$x_m = G_{k_m(x)}(x_{m+1}), \quad x_m = L_{k_m(x)}(x)$$

Let Σ stand for the space of strictly increasing sequences $\bar{k} = \{k_m\}$ of symbols $k_m \in \mathbb{N} \cup \{\infty\}$ that terminate at moment m if and only if $k_m = \infty$. Endow Σ with a weak topology and the measure ν corresponding to the following *random walk* on $\mathbb{N}$: the transition probability of jumping from $k \in \mathbb{N}$ to $l \in \mathbb{N}$ is equal to $\nu_{kl} = 1/2^{l-k}$ if $l > k$, and it vanishes otherwise. The initial distribution on $\mathbb{N}$ is given by $\nu\{k\} = 1/2^{k+1}$. In particular, the measure given to a cylinder/path is

$$\nu([k_0, k_1, k_2, \dots, k_n]) = \nu(k_0) \prod_{i=0}^{n-1} \nu_{k_i k_{i+1}}.$$

We let $j_m := k_{m+1} - k_m$ be the *jumps* in our random walk.

Lemma 2.1. *The coding $x \mapsto \bar{k}(x)$ establishes a homeomorphism between $\mathcal{O}$ and Σ and a measure-theoretic isomorphism between $(\mathcal{O}, \mu)$ and (Σ, ν) .*

Proof. The first entry $k_0 = k_0(x)$ of a sequence $\bar{k}$ determines the piece $B_{i_0}^{k_0+1}$ of level $k_0 + 1$ containing x , the second entry k_1 determines the piece $B_{i_1}^{k_1+1}$ of level $k_1 + 1$, etc, so the sequence $\bar{k}$ determines x . It is also clear that all monotonic sequences $\bar{k}$ are realizable, so the coding is one-to-one. Moreover, the conditional probability of passing from k_m to k_{m+1} is equal to

$$\nu(B_{i_{m+1}}^{k_{m+1}})/\nu(B_{i_m}^{k_m}) = 2^{-(k_{m+1}-k_m)},$$

so ν induces the desired random walk distribution on Σ . $\square$

We can also consider the random walk that *stops on depth* n . This means that we consider the orbit $F^{-t}x$ only until the moment it lands in B^n . The corresponding (finite) coding sequence $\{\tilde{k}_m\}_{m=0}^T$ is defined as follows: $\tilde{k}_m = k_m$ whenever $k_m < n$ ($m = 0, 1, \dots, T-1$), while $\tilde{k}_T = n$. (In what follows we will skip “tilde” in the notation as long as it would not lead to confusion.)

Fix an increasing *control function* $s : \mathbb{N} \rightarrow \mathbb{Z}_+$. We say that a sequence $\bar{k} = \{k_m\}_{m=0}^\infty$ is *s-controlled after a moment* N if $j_m \leq s(k_m)$ for all $k_m \geq N$. We say that a point $x \in \mathcal{O}$ is *s-controlled after moment* N if its code $\bar{k}(x)$ is such. The set of these points is denoted by K_N .

The points in K_N correspond to path of the random walk which do not have too long jumps. These are the paths along which a geometric analysis can be performed. Indeed, these controlled paths constitute most of the paths. The following lemma is at the heart of Probabilistic Universality.

Lemma 2.2. *Under the summability assumption*

$$\sum_{k=0}^{\infty} \frac{1}{2^{s(k)}} < \infty$$

we have

$$v(K_N) \geq 1 - O\left(\sum_{k=N}^{\infty} \frac{1}{2^{s(k)}}\right).$$

Proof. It follows immediately from the definition of the random walk, using the monotonicity of the control function, that

$$v(K_N) \geq \prod_{k=N}^{\infty} \left(1 - \frac{1}{2^{s(k)}}\right),$$

which implies the Lemma. $\square$

2.3. Geometric estimates. Our analysis depends essentially on the geometric control of the renormalizations and changes of variables established in [7].

The renormalizations have the following nearly *universal shape*:

$$R^n F = (f_n(x) - b^{2^n} a(x) y (1 + O(\rho^n)), x), \quad (2.5)$$

where the f_n converge exponentially fast to the universal unimodal map f_* , $a(x)$ is a universal function, and $\rho \in (0, 1)$ is universal.

The changes of variables $\Psi_k^l : \text{Dom}(F^l) \rightarrow \text{Dom}(F^k)$ have the following form:

$$\Psi_k^l = D_k^l \circ (\text{id} + \mathbf{S}_k^l), \quad (2.6)$$

where

$$D_k^l = \begin{pmatrix} 1 & t_k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} (\sigma^2)^{l-k} & 0 \\ 0 & (-\sigma)^{l-k} \end{pmatrix} (1 + O(\rho^k)). \quad (2.7)$$

is a linear map with $t_k \asymp b_F^{2^k}$, while $\text{id} + \mathbf{S}_k^l : (x, y) \mapsto (x + S_k^l(x, y), y)$ is a horizontal non-linear map with

$$|\partial_x S_k^l| = O(1), \quad |\partial_y S_k^l| = O(\bar{\varepsilon}^{2^k}).$$

2.4. Regular pieces. The main goal is to show that many pieces are essentially very thin rectangular strips. For examples the pieces B^n surrounding the tip are as such, see Fig. 1. The first step towards understanding the shape of pieces, Sect. 4, is to show that some are contained in well-controlled rectangles. In this section we outline the results of Sect. 4.

For any $x \in \mathcal{O}$, we let $B^n(x)$ be the piece $B_\omega^n \in \mathcal{B}^n$ containing x (in particular, $B^n(\tau) = B^n$). Let $\mathcal{B}_*^n = \mathcal{B}^n \setminus \{B^n\}$ stand for the family of pieces B_ω^n that do not contain the tip.

Notice that the depth of all points x in any piece $B \in \mathcal{B}_*^n$ is the same, so it can be assigned to the piece itself. In other words,

$$\text{depth}(B) = \sup\{k : B \subset B^k\} \in \{0, 1, \dots, n-1\}.$$

Let $\mathcal{B}^n[l]$, $l < n$, be the family of all pieces of level n whose depth is l . Note that $\mathcal{B}^n[l]$ contains 2^{n-l-1} boxes.

We can view the piece B in the renormalization coordinates on various scales. Namely, to view B from scale $k \leq n$ means that we consider its preimage $\mathbf{B}$ under the (nonlinear) rescaling $\Psi_0^k : \text{Dom}(F_k) \rightarrow B^k$. The main scale from which B will be viewed is its depth k , so from now on $\mathbf{B} := (\Psi_0^k)^{-1}(B)$ will stand for the corresponding piece (see Fig. 3). This seemingly minor ingredient plays a crucial role in the estimates. The reason is as follows. A piece $B = \Psi_0^k(\mathbf{B})$ is a strongly distorted image of $\mathbf{B}$. The cause of this is the form of the linear part of Ψ_0^k which contracts much more strongly horizontally than vertically, see (2.7). In particular, the geometry of the piece B will degenerate when k increases. However, the domain $\mathbf{B}$ might have a bounded geometry. In particular, one should not expect a bounded control of the geometry of B . On the other hand one might see bounded control of the geometry of $\mathbf{B}$, i.e. B seen from the scale k .

A piece B as above is called *regular* if the horizontal and vertical projections of $\mathbf{B}$ are K -comparable, where $K > 0$ is a universal constant, to be specified in the main body of the paper, see Sect. 4. In other words, $\text{mod } \mathbf{B}$ (the ratio of the length of the vertical and horizontal projections of $\mathbf{B}$) is of order 1.

Indeed, there are pieces which are not regular. For example there are pieces B such that the horizontal projection of $\mathbf{B} \subset \text{Dom}(F_k)$ is much smaller than its vertical projection. The renormalization F_k contracts strongly the vertical direction and might result in a strongly distorted piece $F_k(\mathbf{B})$.

A key ingredient to control distortion is the following mechanism, formally described in Proposition 4.1. If a regular piece B , i.e. $\text{mod}(\mathbf{B}) \asymp 1$, is not too deep then $F^{2^k}(B)$ is also regular, i.e. $\text{mod}(F_k(\mathbf{B})) \asymp 1$. The notion of *not too deep* is discussed in the next paragraphs.

We will control depth by the control function

$$s(k) = a2^k - A \quad \text{where } a = \frac{\ln b}{\ln \sigma}, \quad A = \frac{\ln \alpha}{\ln \sigma}, \quad (2.8)$$

with a sufficiently small universal $\alpha > 0$ to be specified in the main body of the paper. With this choice, we have:

$$\alpha \sigma^{l-k} \geq b^{2^k}. \quad (2.9)$$

The outline of the proof of Proposition 4.1 will show why (2.9) is a sufficient criterium to control distortion.

We say that the piece $B \in \mathcal{B}^n[l]$ is *not too deep* in scale B^k if

$$l - k \leq s(k),$$

with $k < l < n$. Compare to (4.3). Pieces which are too deep in scale B^k might be distorted a lot when one applies F_k to push them up to higher levels. This notion of control is at the heart of the analysis. In [1], Benedicks and Carleson have the notion of deep returns. These are excluded to avoid moments of very strong contraction. The deep pieces in this work are excluded to avoid strong distortion. The relation is that both situations are discussed to avoid undesirable behavior, i.e. in the case of [1], the authors were constructing expanding orbits, contraction was the enemy and in the present work the goal is to find as many pieces which have the right universal geometry. In this case, distortion is the enemy.

There are a number of constants which have to be chosen appropriately, for example α and K . In the main body of this paper it will be shown how to choose these constants carefully such that all Lemmas and Propositions hold. From now on we will assume in this outline that the constants are chosen appropriately and will not mention this matter any more.

We will number the Lemmas and Propositions in this outline as the corresponding statements in the main body. However, the version in the outline should be viewed as an informal version of the actual statements.

Renormalizations are rescaled versions of iterates F^{2^k} restricted to the domains B^k containing the tip. If a piece is not too deep, i.e. not too close to the tip, then the branch $F^{2^k} : B^{k+1} \rightarrow B^k$ applied to this piece will keep it regular. Namely,

Proposition 4.1. *For all sufficiently big levels k , the following is true. If a regular piece $B \in \mathcal{B}_*^n$, $n > k$, is not too deep in scale B^k then $G_k(B)$ is regular.*

Outline of the proof. A piece $B \in \mathcal{B}^n[k]$ of depth k can be rescaled to a piece in the domain of the k th renormalization F_k , namely $B = \Psi_0^k(\mathbf{B})$. Recall, the fact that a piece is regular refers to a geometric property of $\mathbf{B}$, i.e. a geometric property of B viewed from the scale k , see Fig. 3.

Let $B \in \mathcal{B}^n[l]$, $n > l > k$. This piece has depth l and hence we should view B from scale l , i.e. consider the piece $\mathbf{B}$ of level $n - l$ for the renormalization F_l ,

$$B = \Psi_0^l(\mathbf{B}) = \Psi_0^k \circ \Psi_k^l(\mathbf{B}),$$

see Fig. 4. As the piece $\tilde{B} = G_k(B)$ has depth k , it should be viewed from the scale k . So, we consider the corresponding piece $\tilde{\mathbf{B}}$ of level $n - k$ for the renormalization F_k ,

$$\tilde{B} = \Psi_0^k(\tilde{\mathbf{B}}),$$

see Fig. 4. Then, because Ψ_0^k is a conjugacy between F^{2^k} and F_k ,

$$\tilde{\mathbf{B}} = F_k \circ \Psi_k^l(\mathbf{B}).$$

Using geometric estimates for factorization (2.6) we show that

$$\text{mod } \Psi_k^l(\mathbf{B}) \asymp \sigma^{l-k} \text{ mod } \mathbf{B},$$

provided $\mathbf{B}$ is regular. So $\Psi_k^l(\mathbf{B})$ is highly stretched in the vertical direction. The nearly universal map F_k , see (2.5), will contract the vertical size by a factor of order b^{2^k} where

$b^{2^k} \ll \sigma^{l-k}$ since the piece is not too deep. This implies that the image under F_k is essentially the image of the horizontal side. We obtain a piece $\tilde{\mathbf{B}}$, which is essentially a curve, that gets roughly aligned with the parabola, which makes its modulus of order 1. $\square$

2.5. Universal sticks. Given a piece $B \in \mathcal{B}^n[l]$ of a map F , let $\mathcal{O}(B) := \mathcal{O}_F \cap B$ be the part of the Cantor set $\mathcal{O}_F$ contained in B . Respectively, $\mathcal{O}(\mathbf{B}) = \mathcal{O}_{F_l} \cap \mathbf{B}$, where $\mathbf{B}$ is the rescaled piece (rescaled by Ψ_0^l) corresponding to B , i.e. $B = \Psi_0^l(\mathbf{B})$.

The main goal is to show that many pieces are essentially very thin rectangular strips, not necessarily horizontal or vertical. As if they are skinny sticks. The formal way to describe the geometry of a piece is illustrated in Fig. 6 and is discussed in the following paragraph.

The *associated rectangle* of a piece is the smallest rectangle with horizontal and vertical sides which contains the piece. A natural quantification to describe the geometry of a piece is to the measure how far a piece is from the diagonal of its associated rectangle. We say that a piece $B \in \mathcal{B}^n[l]$ is a δ -stick if the Cantor set $\mathcal{O}(\mathbf{B})$ is contained in a diagonal strip Π of thickness δ , relatively the horizontal size of $\mathbf{B}$. The minimal thickness is denoted by $\delta_{\mathbf{B}}$. See Fig. 6.

Let us consider the pieces B_1 and B_2 of level $n + 1$ contained in B . The *scaling number* of the piece $\mathbf{B}_1$ is the ratio of the vertical sizes of the associated rectangles of $\mathbf{B}_1$ and $\mathbf{B}$, see Fig. 7. Similarly, one defines the scaling number of $\mathbf{B}_2$.

Let $\sigma_{\mathbf{B}_1}^*$ and $\sigma_{\mathbf{B}_2}^*$ be the scaling numbers of the corresponding two pieces for the degenerate renormalization fixed point F_* . Let $\Delta\sigma_{\mathbf{B}}$ be the maximal difference between the two corresponding scaling numbers, see Sect. 6 for the precise definitions. A piece $B \in \mathcal{B}^n$ is called ϵ -universal if $\delta_{\mathbf{B}} \leq \epsilon$ and $\Delta\sigma_{\mathbf{B}} \leq \epsilon$. Given a piece B let B^* be the corresponding piece of the renormalization fixed point F_* . Observe, if a piece B is ϵ -universal one can rescale $\mathbf{B}$ and $\mathbf{B}^*$ conformally to unit scale such that both are precisely contained in the rectangle $[0, 1] \times [0, \delta]$. Moreover, the corresponding rescaled pieces of $\mathbf{B}_{1,2}$ are ϵ -close in Hausdorff distance to the corresponding rescaled pieces of $\mathbf{B}_{1,2}^*$.

Choose $q_0 > 0$ small. The pieces $B \in \mathcal{B}^n[k]$, with $(1 - q_0) \cdot n \leq k \leq n$, at the scale $n - k$ are said to be *very deep*. Consider a very deep piece $B \in \mathcal{B}^n[k]$ and view it from scale k , say $B = \Psi_0^k(\mathbf{B})$ with $\mathbf{B} \in \mathcal{B}^{n-k}(F_k)$. Then $\mathbf{B}$ is in the orbit of $B_v^{n-k}(F_k)$ under F_k , i.e. $\mathbf{B} = F_k^t(B_v^{n-k}(F_k))$ with $0 < t < 2^{n-k}$. The exponential convergence of renormalization implies that F_k is at a distance $O(\rho^k)$ to the degenerate renormalization fixed point F_* , see (2.5). When $q_0 > 0$ is small enough, these few iterates, $2^{n-k} = 2^{q_0 \cdot n}$, with a map $O(\rho^{(1-q_0) \cdot n})$ close to the renormalization fixed can be well approximated by iterates of the renormalization fixed point. At this scale, one-dimensional dynamics is a good geometric model. We call this the *one-dimensional regime*. In the one-dimensional regime one has the following exponential small universality.

Proposition 7.2. *There exist $\theta < 1$, $0 < q_0 < q_1$ such that every piece in $\mathcal{B}^n[k]$, with $(1 - q_1) \cdot n \leq k \leq (1 - q_0) \cdot n$, is $O(\rho^n)$ -universal.*

We are going to refine Proposition 4.1, in the sense that we are estimating how ϵ -universality is distorted (or even improved!) when we apply maps G_k to regular pieces which are not too deep in scale B^k . Figure 4 illustrates the following proposition.

Proposition 5.1 and 6.1 *If $B \in \mathcal{B}^n[l]$ is regular and not too deep in B^k then*

$$\delta_{\tilde{\mathbf{B}}} \leq \frac{1}{2} \cdot \delta_{\mathbf{B}} + O(\sigma^{n-l}),$$

and

$$\Delta\sigma_{\tilde{\mathbf{B}}} = \Delta\sigma_{\mathbf{B}} + O(\delta_{\mathbf{B}} + \sigma^{n-l}),$$

where $\tilde{B} = G_k(B) \in \mathcal{B}^n[k]$ and $\tilde{\mathbf{B}} = F_k(\Psi_k^l(\mathbf{B}))$.

Outline of the proof. From (2.6) and (2.7) we have the factorization $\tilde{\mathbf{B}} = F_k(\Psi_k^l(\mathbf{B})) = F_k \circ D_k^l \circ (\text{id} + \mathbf{S}_k^l)(\mathbf{B})$. We consecutively estimate, using geometric estimates of Sect. 2.3, the thickness of the pieces $B_{\text{diff}} = (\text{id} + \mathbf{S}_k^l)(\mathbf{B})$, $B_{\text{aff}} = D_k^l(B_{\text{diff}})$ and $\tilde{\mathbf{B}} = F_k(B_{\text{aff}})$, see Fig. 5. The thickness of D_{diff} is comparable with the thickness of $\mathbf{B}$, up to an error of order σ^{n-l} , since the horizontal map $\text{id} + \mathbf{S}_k^l$ has bounded derivatives and $\text{diam } \mathbf{B} \leq \sigma^{n-l}$.

Let us now represent the affine map D_k^l as a composition of the diagonal part Λ and the shear part T , see (2.7). The diagonal map Λ preserves the horizontal thickness, so the thickness is only affected by the shear part T , which has order $t_k \asymp b^{2^k}$. Using this estimate and that B is not too deep in B^k , we show that $\delta(B_{\text{aff}}) = O(\delta_{B_{\text{diff}}})$.

Finally, we show that the map F_k , being strongly vertically contracting, improves thickness again using that B is not too deep in B^k .

The diffeomorphisms Ψ_k^l map horizontal lines to horizontal lines. From this property one obtains that the maps Ψ_k^l do not distort the scaling numbers at all as a consequence of the definition of scaling numbers. The piece $\tilde{\mathbf{B}}$ is the image under F_k of $B_{\text{aff}} = \Psi_k^l(\mathbf{B})$. A geometric observation implies that $\Delta\sigma_{\tilde{\mathbf{B}}} = \Delta\sigma_{\mathbf{B}} + O(\delta(B_{\text{aff}})) = \Delta\sigma_{\mathbf{B}} + O(\delta(\mathbf{B}) + \sigma^{n-l})$.

□

The pieces $B \in \mathcal{B}^n$ can be reconstructed by applying iterates of F to $B^n = B_{v^n}^n$, i.e. each piece is of the form $F^i(B^n)$ with $0 \leq i < 2^n$. However, it is not possible to control the distortion of the maps $F^i|B^n$. In fact the distortion of those iterates is not uniformly bounded. The nonrigidity of the Cantor set $\mathcal{O}_F$ is a consequence of this, see [7].

Recall that we describe the geometry of a piece $B \in \mathcal{B}^n[k]$ in terms of the geometry of the piece viewed from the scale k , denoted by $\mathbf{B}$. These pieces are in the domain of F_k . In particular, each piece $\mathbf{B}$ can be reconstructed by applying F_k to $B_v^{n-k}(F_k)$, i.e. $\mathbf{B} = F_k^t(B_v^{n-k}(F_k))$ with $0 < t < 2^{n-k}$. When the piece is very deep, in the one-dimensional regime, we can control the distortion of the iterates $F_k^t|B_v^{n-k}(F_k)$, see Proposition 7.2.

For pieces which are not very deep we reconstruct the pieces by applying rescaling maps $\Psi_k^{k'}$ and the renormalization F_k to pieces obtained in the one-dimensional regime. In particular, for every $B \in \mathcal{B}^n[k]$ there is a sequence $k < k_1 < k_2 < \dots < k_t = (1 - q_1) \cdot n$ such that

$$\mathbf{B} = F_k \circ \Psi_k^{k_1} \circ F_{k_1} \circ \Psi_{k_1}^{k_2} \circ \dots \circ F_{k_{t-1}} \circ \Psi_{k_{t-1}}^{k_t}(B'),$$

with $B' \in \mathcal{B}^{q_1 \cdot n}(F_{(1-q_1) \cdot n})$ from the one-dimensional regime. Consider the piece

$$\mathbf{B}_{k_i} = F_{k_i} \circ \Psi_{k_i}^{k_{i+1}} \circ \dots \circ F_{k_{t-1}} \circ \Psi_{k_{t-1}}^{k_t}(B').$$

This piece is the piece $\Psi_0^{k_i}(\mathbf{B}_{k_i})$ viewed from scale k_i . In particular, we are reconstructing the piece B by *pushing-up* the piece B' , obtained in the one-dimensional regime, along a sequence of pieces $\mathbf{B}_{k_i}$ which are pushed-up in scale from k_i to k_{i-1} . Compare Fig. 4.

Observe, the maps F_{k_i} and $\Psi_{k_i}^{k_{i+1}}$ are very well controlled, see (2.5) and (2.6). Nevertheless, a map F_{k_i} might distort a piece $\mathbf{B}_{k_{i+1}}$ to produce

$$\mathbf{B}_{k_i} = F_{k_i}(\Psi_{k_i}^{k_{i+1}}(\mathbf{B}_{k_{i+1}})).$$

However, if the piece $\Psi_{k_i}^{k_{i+1}}(\mathbf{B}_{k_{i+1}})$ is not too deep, according to Proposition 5.1 and 6.1, F_{k_i} will distort a regular piece of the form $\Psi_{k_i}^{k_{i+1}}(\mathbf{B}_{k_{i+1}})$ by a small amount. According to Proposition 5.1 and 6.1 the newly created pieces will have an improved thickness and their scaling numbers are essentially not distorted.

One can control the geometry of a piece if it is obtained by pushing up avoiding pieces which are too deep. The starting pieces are from the one-dimensional regime and have good precision. If one avoids pieces $\Psi_{k_i}^{k_{i+1}}(\mathbf{B}_{k_{i+1}})$ which are too deep one can preserve the good precision. This process is called the *pushing-up regime*.

The pieces created by the combined one-dimensional and pushing-up regimes are $O(\rho^n)$ -universal. This can be seen as follows. Proposition 7.2, states that the pieces from the one-dimensional regime are exponentially universal. These pieces are the starting pieces of the pushing-up regime. Propositions 5.1 and 6.1, state that the error in scaling ratios caused by pushing-up is of order of the sum of the thicknesses observed during the pushing-up process. Moreover, the thicknesses are essentially contracted each pushing-up step.

Unfortunately, the pieces generated by the combination of the one-dimensional and pushing-up regimes, do not have a total measure which tends to 1. In particular, Proposition 8.2 states that asymptotically, these pieces will be missing a fraction of the order $O(2^k(b^\gamma)^{2^k})$ of B^k , where $\gamma > 0$. This is an immediate consequence of the fact that during the pushing-up regime we only pushed-up pieces which are not too deep.

The solution to this problem is to stop the pushing-up regime at the level $\kappa(n) \asymp \ln n$. Then $B^{\kappa(n)}$ will be filled except for an exponentially small fraction with $O(\rho^n)$ -universal pieces. After level $\kappa(n)$ we start the *brute-force* regime, push-up all pieces without considering whether they are too deep or not. In other words, just apply the original map F for $2^{\kappa(n)}$ steps. But under these iterates the $O(\rho^n)$ -universal sticks get spoiled at most by factor $O(C^{\kappa(n)}) = O(n^c)$ with some $c > 0$. Hence, they are $O(n^c \rho^n)$ -universal sticks, and we still see $O(\theta^n)$ -universality, for some $\theta < 1$.

Denote the pieces in B^n generated by combining these three regimes by $\mathcal{P}_n$. These pieces are θ^n -universal. The Sects. 6 and 7 are devoted to the precise proof.

2.6. Probabilistic universality. We say that the geometry of $\mathcal{O}$ is *probabilistically universal* if there exists a $\theta \in (0, 1)$ such that the total μ -measure of pieces $B \in \mathcal{B}^n$ which are θ^n -universal sticks is at least $1 - O(\theta^n)$.

Theorem 2.3. *The geometry of $\mathcal{O}$ is probabilistically universal.*

Proof. Let $n \geq 1$. The pieces in $\mathcal{P}_n$ are θ^n -universal. Left is to estimate $\mu(\mathcal{P}_n)$.

The one-dimensional regime deals with the pieces of $\mathcal{B}^n$ in B^k with $(1 - q_1) \cdot n \leq k \leq (1 - q_0) \cdot n$. They occupy a fraction $1 - O(\frac{1}{2^{(q_1 - q_0) \cdot n}})$ of the μ -measure of $B^{(1 - q_1) \cdot n}$. Push them up until B^0 without restriction whether they are too deep or not. They will occupy B^n except for an exponentially small fraction. Let R_n be the corresponding set of paths

of the random walk. These are the paths which hit the interval $[(1 - q_1) \cdot n, (1 - q_0) \cdot n]$ at least once but are not necessarily s -controlled. So

$$v(R_n) = 1 - O\left(\frac{1}{2^{(q_1 - q_0) \cdot n}}\right).$$

Recall, the set $K_{\kappa(n)}$ consists of the paths which are s -controlled after depth $\kappa(n) \asymp \ln n$. Lemma 2.2 gives

$$v(K_{\kappa(n)}) = 1 - O\left(\sum_{k=\kappa(n)}^{\infty} \frac{1}{2^{s(k)}}\right) = 1 - O\left(\frac{1}{2^{a^{2\kappa(n)}}}\right) = 1 - O(\rho^n)$$

for some $\rho \in (0, 1)$.

Observe, the set of paths corresponding to $\mathcal{P}_n$ is $R_n \cap K_{\kappa(n)}$. Hence,

$$\mu(\mathcal{P}_n) = v(R_n \cap K_{\kappa(n)}) = 1 - O(\theta^n),$$

for some $\theta \in (0, 1)$. $\square$

3. Preliminaries

A complete discussion of the following definitions and statements can be found in part I and part II, see [7, 16], of this series on renormalization of Hénon-like maps.

Let $\Omega^h, \Omega^v \subset \mathbb{C}$ be neighborhoods of $[-1, 1] \subset \mathbb{R}$ and $\Omega = \Omega^h \times \Omega^v$. The set $\mathcal{H}_{\Omega}(\bar{\epsilon})$ consists of maps $F : [-1, 1]^2 \rightarrow [-1, 1]^2$ of the following form.

$$F(x, y) = (f(x) - \epsilon(x, y), x),$$

where $f : [-1, 1] \rightarrow [-1, 1]$ is a unimodal map which admits a holomorphic extension to Ω^h and $\epsilon : [-1, 1]^2 \rightarrow \mathbb{R}$ admits a holomorphic extension to Ω and finally, $|\epsilon| \leq \bar{\epsilon}$. The critical point c of f is non degenerate, $D^2 f(c) < 0$. A map in $\mathcal{H}_{\Omega}(\bar{\epsilon})$ is called a *Hénon-like* map. Observe that Hénon-like maps map vertical lines to horizontal lines.

A unimodal map $f : [-1, 1] \rightarrow [-1, 1]$ with critical point $c \in [-1, 1]$ is *renormalizable* if $f^2 : [f^2(c), f^4(c)] \rightarrow [f^2(c), f^4(c)]$ is unimodal and $[f^2(c), f^4(c)] \cap f([f^2(c), f^4(c)]) = \emptyset$. The renormalization of f is the affine rescaling of $f^2|_{[f^2(c), f^4(c)]}$, denoted by Rf . The domain of Rf is again $[-1, 1]$. The renormalization operator R has a unique fixed point $f_* : [-1, 1] \rightarrow [-1, 1]$. The introduction of [8] presents the history of renormalization of unimodal maps and describes the main results.

The *scaling factor* of this fixed point f_* is

$$\sigma = \frac{|[f_*^2(c), f_*^4(c)]|}{|[-1, 1]|}.$$

A Hénon-like map is renormalizable if there exists a domain $D \subset [-1, 1]^2$ such that $F^2 : D \rightarrow D$. The construction of the domain D is inspired by renormalization of unimodal maps. In particular, it should be a topological construction. However, for small $\bar{\epsilon} > 0$ the actual domain $A \subset [-1, 1]^2$, used to renormalize as was done in [7], has an analytical definition. The precise definition can be found in Sect. 3.5 of part I. If the renormalizable Hénon-like maps is given by $F(x, y) = (f(x) - \epsilon(x, y))$ then

the domain $A \subset [-1, 1]^2$, an essentially vertical strip, is bounded by two curves of the form

$$f(x) - \varepsilon(x, y) = \text{Const.}$$

These curves are graphs over the y -axis with a slope of the order $\bar{\varepsilon} > 0$. The domain A satisfies similar combinatorial properties as the domain of renormalization of a unimodal map:

$$F(A) \cap A = \emptyset,$$

and

$$F^2(A) \subset A.$$

Unfortunately, the restriction $F^2|_A$ is not a Hénon-like map as it does not map vertical lines into horizontal lines. This is the reason why the coordinated change needed to define the renormalization of F is not an affine map, but it rather has the following form. Let

$$H(x, y) = (f(x) - \varepsilon(x, y), y)$$

and

$$G = H \circ F^2 \circ H^{-1}.$$

The map H preserves horizontal lines and it is designed in such a way that the map G maps vertical lines into horizontal lines. Moreover, G is well defined on a rectangle $U \times [-1, 1]$ of full height. Here $U \subset [-1, 1]$ is an interval of length $2/|s|$ with $s < -1$. Let us rescale the domain of G by the affine s -dilation Λ , such that the rescaled domain is of the form $[-1, 1] \times V$, where $V \subset \mathbb{R}$ is an interval of length $2|s|$. Define the renormalization of F by

$$RF = \Lambda \circ G \circ \Lambda^{-1}.$$

Notice that RF is well defined on the rectangle $[-1, 1] \times V$. The coordinate change $\psi = H^{-1} \circ \Lambda^{-1}$ maps this rectangle onto the topological rectangle A of full height.

The set of n -times renormalizable maps is denoted by $\mathcal{H}_\Omega^n(\bar{\varepsilon}) \subset \mathcal{H}_\Omega(\bar{\varepsilon})$. If $F \in \mathcal{H}_\Omega^n(\bar{\varepsilon})$ we use the notation

$$F_n = R^n F.$$

The set of infinitely renormalizable maps is denoted by

$$\mathcal{I}_\Omega(\bar{\varepsilon}) = \bigcap_{n \geq 1} \mathcal{H}_\Omega^n(\bar{\varepsilon}).$$

The renormalization operator acting on $\mathcal{H}_\Omega^1(\bar{\varepsilon})$, $\bar{\varepsilon} > 0$ small enough, has a unique fixed point $F_* \in \mathcal{I}_\Omega(\bar{\varepsilon})$. It is the degenerate map

$$F_*(x, y) = (f_*(x), x).$$

This renormalization fixed point is hyperbolic and the stable manifold has codimension one. Moreover,

$$W^s(F_*) = \mathcal{I}_\Omega(\bar{\varepsilon}).$$

If we want to emphasize that some set, such as A , is associated with a certain map F we use notation like $A(F)$.

The coordinate change which conjugates $F_k^2|A(F_k)$ to F_{k+1} is denoted by

$$\psi_v^{k+1} = (\Lambda_k \circ H_k)^{-1} : \text{Dom}(F_{k+1}) \rightarrow A(F_k). \quad (3.1)$$

Here H_k is the non-affine part of the coordinate change used to define $R^{k+1}F$ and Λ_k is the dilation by $s_k < -1$. Now, for $k < n$, let

$$\Psi_k^n = \psi_v^{k+1} \circ \psi_v^{k+2} \circ \cdots \circ \psi_v^n : \text{Dom}(F_n) \rightarrow A_{n-k}(F_k), \quad (3.2)$$

where

$$A_k(F) = \Psi_0^k(\text{Dom}(F_k)) \cap B.$$

The renormalizations $F_k = R^k F$ are obtained by nonlinear rescalings Ψ_0^k of the first return map F^{2^k} to the domain A_k . Notice, that each $A_k \subset [-1, 1]^2$ is an almost vertical strip of exponential small width and of full vertical height and Ψ_0^k conjugates $R^k F$ to $F^{2^k}|A_k$. Furthermore, one has $A_{k+1} \subset A_k$ because the strips A_k are the renormalization domains of consecutive renormalizations.

The change of coordinates conjugating the renormalization RF to F^2 is denoted by $\psi_v^1 := H^{-1} \circ \Lambda^{-1}$. To describe the attractor of an infinitely renormalizable Hénon-like map we also need the map $\psi_c^1 = F \circ \psi_v^1$. The subscripts v and c indicate that these maps are associated to the critical *value* and the *critical* point, respectively.

Similarly, let ψ_v^2 and ψ_c^2 be the corresponding changes of variable for RF , and let

$$\psi_{vv}^2 = \psi_v^1 \circ \psi_v^2, \quad \psi_{cv}^2 = \psi_c^1 \circ \psi_v^2, \quad \psi_{vc}^2 = \psi_v^1 \circ \psi_c^2, \quad \psi_{cc}^2 = \psi_c^1 \circ \psi_c^2.$$

and, proceeding this way, for any $n \geq 0$, construct 2^n maps

$$\psi_w^n = \psi_{w_1}^1 \circ \cdots \circ \psi_{w_n}^n, \quad w = (w_1, \dots, w_n) \in \{v, c\}^n.$$

The notation $\psi_w^n(F)$ will also be used to emphasize the map under consideration, and we will let $W = \{v, c\}$ and $W^n = \{v, c\}^n$ be the n -fold Cartesian product. The following Lemma and its proof can be found in [7, Lemma 5.1].

Lemma 3.1. *Let $F \in \mathcal{I}_\Omega(\bar{\varepsilon})$. There exist $C > 0$ such that for $w \in W^n$, $\|D\psi_w^n\| \leq C\sigma^n$, $n \geq 1$.*

Let $F \in \mathcal{I}_\Omega(\bar{\varepsilon})$ and consider the domains

$$B_\omega^n = \text{Im } \psi_\omega^n.$$

These form a family $\mathcal{B}^n = \{B_\omega^n\}$ of 2^n pieces, called the n th *renormalization level*.

The first return maps to the domains

$$B_{v^n}^n = \text{Im } \Psi_0^n = \text{Im } \psi_{v^n}^n$$

correspond to the different renormalizations. Notice,

$$B_{v^{n+1}}^{n+1} \subset B_{v^n}^n.$$

An infinitely renormalizable Hénon-like map has an invariant Cantor set:

$$\mathcal{O}_F = \bigcap_{n \geq 1} \bigcup_{i=0}^{2^n-1} F^i(B_{v^n}^n) = \bigcap_{n \geq 1} \bigcup_{\omega \in W^n} B_{\omega}^n.$$

The dynamics on this Cantor set is conjugate the dyadic adding machine. Its unique invariant measure is denoted by μ . The *average Jacobian*

$$b_F = \exp \int \ln \text{Jac } F d\mu$$

with respect to μ is an important parameter that influences the geometry of $\mathcal{O}_F$, see [7, Theorem 10.1].

The critical point (and critical value) of a unimodal map plays a crucial role in its dynamics. The counterpart of the critical value for infinitely renormalizable Hénon-like maps is the *tip*

$$\{\tau_F\} = \bigcap_{n \geq 1} B_{v^n}^n.$$

3.1. One-dimensional maps. Recall that $f_*: [-1, 1] \rightarrow [-1, 1]$ stands for the one-dimensional renormalization fixed point normalized so that $f_*(c_*) = 1$ and $f_*^2(c_*) = -1$, where $c_* \in [-1, 1]$ is the critical point of f_* . The renormalization fixed point f_* has a nested sequence of renormalization cycles $\mathcal{C}_n, n \geq 1$. A cycle consists of the following intervals. The critical point of f_* is c_* and the critical value $v_* = f_*(c_*)$

$$I_j^*(n) = [f_*^j(v_*), f_*^{j+2^n}(v_*)] \in \mathcal{C}_n,$$

with $j = 0, 1, 2, \dots, 2^n - 1$. The collection $\mathcal{C}_n$ consists of pairwise disjoint intervals. Notice, for $j = 0, 1, 2, \dots, 2^n - 2$

$$f_*(I_j^*(n)) = I_{j+1}^*(n),$$

and

$$f_*(I_{2^n-1}^*(n)) = I_0^*(n).$$

The interval in $\mathcal{C}_n$ which contains the critical point is denoted by

$$U_n = I_{2^n-1}^*(n) \ni c_*.$$

The *nonlinearity* of a C^2 -diffeomorphism $\phi: I \rightarrow \phi(I) \subset \mathbb{R}, I \subset \mathbb{R}$, is

$$\eta_\phi = D \ln D\phi. \quad (3.3)$$

The *Distortion* of a diffeomorphism $\phi: I \rightarrow J$ between intervals $I, J \subset \mathbb{R}$ is defined as

$$\text{Dist}(\phi) = \max_{x,y} \ln \frac{D\phi(y)}{D\phi(x)}.$$

If η is the nonlinearity of ϕ then

$$\text{Dist}(\phi) \leq \|\eta\|_{C^0} \cdot |I|. \quad (3.4)$$

The distortion of a map does not change if we rescale domain and range.

Given $r > 0$. The r -neighborhood $T \supset I$ of an interval $I \subset \mathbb{R}$ is the interval such that both components of $T \setminus I$ have length $r|I|$.

Lemma 3.2. *There exist $r > 0$ and $D > 1$ such that the r -neighborhoods $T_j(n) \supset I_j^*(n)$ have the following property. For all $n \geq 1$ the following holds. Let $\xi_j \in T_j(n)$ then*

$$\prod_{l=k_1}^{k_2-1} \frac{|Df_*(\xi_j)|}{\frac{|I_{j+1}^*(n)|}{|I_j^*(n)|}} \leq D.$$

with $0 \leq k_1 < k_2 < 2^n$.

Proof. The a priori bounds on the cycles $\mathcal{C}_n$ are described in [9], see also [4]. The a priori bounds state that for some $r > 0$ the gap between $I_j^*(n+1)$ and $I_{j+2^{n+1}}^*(n)$ satisfies

$$|I_j^*(n) \setminus (I_j^*(n+1) \cup I_{j+2^{n+1}}^*(n))| \geq 5r \cdot |I_j^*(n)|.$$

Hence, we have $T_j(n+1) \cap T_{j+2^{n+1}}(n+1) = \emptyset$ and

$$|T_j(n+1)| + |T_{j+2^{n+1}}(n)| \leq (1-r) \cdot |T_j(n)|.$$

Let $\eta_j(n)$ be the nonlinearity, see (3.3), of the rescaling of $f_* : I_j^*(n) \rightarrow I_{j+1}^*(n)$. The rescaling turns domain and range into $[-1, 1]$. Lemma 3.1 in [19] says that

$$\begin{aligned} \|\eta_j(n+1)\|_{C_0} &\leq \frac{|T_j(n+1)|}{|T_j(n)|} \cdot \|\eta_j(n)\|_{C_0}, \\ \|\eta_{j+2^{n+1}}(n+1)\|_{C_0} &\leq \frac{|T_{j+2^{n+1}}(n+1)|}{|T_j(n)|} \cdot \|\eta_j(n)\|_{C_0}. \end{aligned}$$

Hence,

$$\|\eta_j(n+1)\|_{C_0} + \|\eta_{j+2^{n+1}}(n+1)\|_{C_0} \leq (1-r) \cdot \|\eta_j(n)\|_{C_0},$$

for $j = 0, 1, 2, \dots, 2^n - 2$. The a priori bounds also imply a universal bound

$$\|\eta_{2^n-1}(n+1)\|_{C_0} \leq K.$$

Inductively, this gives a universal bound

$$\sum_{j=0}^{2^n-2} \|\eta_j(n)\|_{C_0} \leq K_0.$$

Use (3.4) and observe,

$$\ln \frac{|Df_*(\xi_j)|}{\frac{|I_{j+1}^*(n)|}{|I_j^*(n)|}} = O(\|\eta_j(n)\|_{C_0}).$$

The Lemma follows. $\square$

Proposition 3.3. *There exists $\rho < 1$, independent of n , such that the following holds. Let $0 < q_0 < 1$ and $I \in \mathcal{C}_n$ and $I \subset U_k \setminus U_{(1-q_0) \cdot n}$, with $k < (1 - q_0) \cdot n$. Let $t_I = 2^k$ be the first return time to U_k . Then for every $j \leq t_I$*

$$\text{Dist}(f_*^j | I) = O(\rho^{q_0 \cdot n}).$$

Proof. Let $s_I \geq t_I$ be the first return time of I to $U_{(1-q_0) \cdot n}$. There exists $J_0 \subset U_k$ with $I \subset J_0$ such that $f_*^{s_I} : J_0 \rightarrow U_{(1-q_0) \cdot n}$ diffeomorphically, [18]. Let $J_k = f_*^k(J_0)$ and $I_k = f_*^k(I)$. The *a priori* bounds on the geometry of the cycles $\mathcal{C}_n$ imply

$$\frac{|I_k|}{|J_k|} = O(\rho^{q_0 \cdot n}).$$

This estimate hold because the intervals J_k are in $\mathcal{C}_{(1-q_0) \cdot n}$ and the intervals I_k are in $\mathcal{C}_n$.

The nonlinearity of the rescaled map $f_* : J_k \rightarrow J_{k+1}$ which has the unit interval as its domain and range, is denoted by η_k . As in the proof of Lemma 3.2 we obtain

$$\sum_{k=0}^{s_I-1} \|\eta_k\|_{C^0} \leq K_0.$$

The nonlinearity of the rescaled version of the map $f_* : I_k \rightarrow I_{k+1}$ which has the unit interval as its domain and range, is denoted by η_k^I . Lemma 3.1 in [19] says that the nonlinearity of the restriction $f_* : I_k \rightarrow I_{k+1}$ of $f_* : J_k \rightarrow J_{k+1}$ satisfies

$$\|\eta_k^I\|_{C^0} \leq \frac{|I_k|}{|J_k|} \cdot \|\eta_k\|_{C^0}.$$

Hence,

$$\sum_{k=0}^{s_I-1} \|\eta_k^I\|_{C^0} = O(\rho^{q_0 \cdot n}).$$

The distortion of a map $f_*^t : I_k \rightarrow I_{k+t}$ is bounded as follows.

$$\begin{aligned} \text{Dist}(f_*^s | I_k) &\leq \sum_{j=k}^{k+t-1} \text{Dist}(f_* | I_j) \\ &\leq \sum_{j=0}^{s_I-1} \|\eta_j^I\|_{C^0} = O(\rho^{q_0 \cdot n}). \end{aligned}$$

This finishes the proof of the Lemma. $\square$

3.2. Geometric properties of the Cantor attractor.

Theorem 3.4 (Universality). *For any $F \in \mathcal{I}_\Omega(\bar{\varepsilon})$ with sufficiently small $\bar{\varepsilon}$, we have:*

$$R^n F = (f_n(x) - b^{2^n} a(x) y (1 + O(\rho^n)), x),$$

where $f_n \rightarrow f_*$ exponentially fast, b is the average Jacobian, $\rho \in (0, 1)$, and $a(x)$ is a universal function. Moreover, a is analytic and positive.

Corollary 3.5. *There exists a universal $d_1 > 0$ such that for $k \geq 1$ large enough*

$$\frac{1}{d_1} \leq \left| \frac{\partial F_k}{\partial x}(z) \right| \leq d_1,$$

for every $z \in B_v^1(F_k)$.

Let τ_n be the tip of $F_n = R^n F$ and τ^* be the tip of F_* .

Lemma 3.6. *There exists $\rho < 1$ such the conjugations*

$$h_n : \mathcal{O}_{F_*} \rightarrow \mathcal{O}_{F_n}$$

with $h_n(\tau_*) = \tau_n$ satisfy

$$|h_n(z) - z| = O(\rho^n),$$

for every $z \in \mathcal{O}_{F_*}$.

Proof. Choose $z^* \in \mathcal{O}_{F_*}$ and let $z = h_n(z^*)$. There are unique sequence $w_{n+1}, \dots, w_m, \dots$, and $z_n, z_{n+1}, \dots, z_m, \dots$, and $z_n^*, z_{n+1}^*, \dots, z_m^*, \dots$ with $w_k \in \{c, v\}$, $z_k \in \mathcal{O}_{F_k}$, and $z_k^* \in \mathcal{O}_{F_*}$ such that $z = z_n, z^* = z_n^*$ and for $k \geq n$

$$\begin{aligned} z_k &= \psi_{w_{k+1}}^{k+1}(z_{k+1}) \\ z_k^* &= (\psi_{w_{k+1}}^{k+1})^*(z_{k+1}^*). \end{aligned}$$

This follows from the construction of $\mathcal{O}_F$ in [7].

Theorem 3.4 implies

$$|\psi_w^{k+1} - (\psi_w^{k+1})^*| = O(\rho^k)$$

for some $\rho < 1$. The proof of Lemma 5.6 in [7] gives that $(\psi_w^{k+1})^* = \psi_w^*$ are contractions, $|D\psi_w^*| \leq \sigma < 1$. Then for $k \geq n$

$$\begin{aligned} |z_k - z_k^*| &= |\psi_{w_{k+1}}^{k+1}(z_{k+1}) - (\psi_{w_{k+1}}^{k+1})^*(z_{k+1}^*)| \\ &\leq |\psi_{w_{k+1}}^{k+1}(z_{k+1}) - (\psi_{w_{k+1}}^{k+1})^*(z_{k+1})| + \\ &\quad |(\psi_{w_{k+1}}^{k+1})^*(z_{k+1}) - (\psi_{w_{k+1}}^{k+1})^*(z_{k+1}^*)| \\ &\leq O(\rho^k) + \sigma \cdot |z_{k+1} - z_{k+1}^*| \end{aligned}$$

Then for every $m > n$ we have

$$|z_n - z_n^*| \leq \sum_{k=n+1}^m O(\rho^k) \cdot \sigma^{k-n-1} + \sigma^{m-n} \cdot |z_m - z_m^*|.$$

Observe, $|z_m - z_m^*| \leq 1$ and the Lemma follows by taking $m > n$ sufficiently large. $\square$

3.3. Analytical properties of the coordinate changes. This section collects the main estimates for the coordinate changes used to define the renormalizations. Observe, the coordinate changes are not affine. However, they have a universal asymptotic shape. These estimates will play a role throughout the manuscript.

Fix an infinitely renormalizable Hénon-like map $F \in \mathcal{I}_\Omega(\bar{\varepsilon})$ to which we can apply the results of [7, 16], $\bar{\varepsilon} > 0$ is small enough. For such an F , we have a well defined *tip*:

$$\tau \equiv \tau(F) = \bigcap_{n \geq 0} B_{v^n}^n$$

Consider the tips of the renormalizations, $\tau_k = \tau(R^k F)$. To simplify the notations, we will translate these tips to the origin by letting

$$\Psi_k = \psi_v^0(R^k F)(z + \tau_{k+1}) - \tau_k.$$

Denote the derivative of the maps Ψ_k at 0 by D_k and decompose it into the unipotent and diagonal factors:

$$D_k = \begin{pmatrix} 1 & t_k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_k & 0 \\ 0 & \beta_k \end{pmatrix}. \quad (3.5)$$

Let us factor this derivative out from Ψ_k :

$$\Psi_k = D_k \circ (\text{id} + \mathbf{s}_k),$$

where $\mathbf{s}_k(z) = (s_k(z), 0) = O(|z|^2)$ near 0. Lemma 7.4 in [7] states

Lemma 3.7. *There exists $\rho < 1$ such that for $k \in \mathbb{Z}_+$ the following estimates hold:*

- (1) $\alpha_k = \sigma^2 \cdot (1 + O(\rho^k))$, $\beta_k = -\sigma \cdot (1 + O(\rho^k))$, $t_k = O(\bar{\varepsilon}^{2^k})$;
- (2) $|\partial_x s_k| = O(1)$, $|\partial_y s_k| = O(\bar{\varepsilon}^{2^k})$;
- (3) $|\partial_{xx}^2 s_k| = O(1)$, $|\partial_{xy}^2 s_k| = O(\bar{\varepsilon}^{2^k})$, $|\partial_{yy}^2 s_k| = O(\bar{\varepsilon}^{2^k})$.

Lemma 3.8. *The numbers t_k quantifying the tilt satisfy*

$$t_k \asymp -b_F^{2^k}.$$

We will use the following notation

$$B_{v^{n-k}}^{n-k}(F_k) = \text{Im } \Psi_k^n.$$

Consider the derivatives of the maps Ψ_k^n at the origin:

$$D_k^n = D_k \circ D_{k+1} \circ \cdots \circ D_{n-1}.$$

We can reshuffle this composition and obtain:

$$D_k^n = \begin{pmatrix} 1 & t_k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} (\sigma^2)^{n-k} & 0 \\ 0 & (-\sigma)^{n-k} \end{pmatrix} (1 + O(\rho^k)). \quad (3.6)$$

Factoring the derivatives D_k^n out from Ψ_k^n , we obtain:

$$\Psi_k^n = D_k^n \circ (\text{id} + \mathbf{S}_k^n), \quad (3.7)$$

where $\mathbf{S}_k^n(z) = (S_k^n(z), 0) = O(|z|^2)$ near 0.

The following Lemma is Lemma 7.6 in [7]

Lemma 3.9. *For $k < n$, we have:*

- (1) $|\partial_x S_k^n| = O(1)$, $|\partial_y S_k^n| = O(\bar{\varepsilon}^{2^k})$;
 (2) $|\partial_{xx}^2 S_k^n| = O(1)$, $|\partial_{yy}^2 S_k^n| = O(\bar{\varepsilon}^{2^k})$, $|\partial_{xy}^2 S_k^n| = O(\bar{\varepsilon}^{2^k} \sigma^{n-k})$.

Lemma 3.10. *There exists a universal $d_0 > 0$ such that for $k \geq 1$ large enough*

$$\frac{1}{d_0} \leq \left| \frac{\partial(\text{id} + S_k^n)}{\partial x} \right| \leq d_0$$

Proof. According to proposition 7.8 in [7], the diffeomorphisms $\text{id} + S_k^n$ stay within a compact family of diffeomorphisms. This gives the upperbound on the derivative. The partial derivative $\frac{\partial(\text{id} + S_k^n)}{\partial x}$ can not be zero in a point because otherwise the derivative of the diffeomorphism would become singular in point. This gives the positive lower bound on the partial derivative. $\square$

Lemma 3.11. *There exists $\rho < 1$ such that*

$$|\Psi_k^n - (\Psi_k^n)^*|_{C^0} = O(\rho^k).$$

Proof. The proof is a small modification of the proof of Lemma 3.6. Use the same notation: $w_l = v$ for all $l \geq k$. We have to incorporate the translation which center the maps around the tips. The estimates in the proof of Lemma 3.6 become

$$|\Psi_k^n(z) - (\Psi_k^n)^*(z)| \leq O(\rho^k) + \sum_{l=k+1}^n O(\rho^l) \cdot \sigma^{l-k-1} + \sigma^{n-k} \cdot |z_n - z_n^*|,$$

where $z_n = z + \tau_n$ and $z_n^* = z + \tau^*$. From Lemma 3.6 we get that $|z_n - z_n^*| = O(\rho^n)$ and the Lemma follows. $\square$

3.4. General notions. We will use the following general notions and notations throughout the text.

Let $X \subset \mathbb{R}^2$ and $Q = [a, a+h] \times [b, b+v]$ be the smallest rectangle containing X . Then $h \geq 0$ is the horizontal size of X and $v \geq 0$ the vertical size.

$Q_1 \asymp Q_2$ means that $C^{-1} \leq Q_1/Q_2 \leq C$, where $C > 0$ is an absolute constant or depending on, say F . Similarly, we will use $Q_1 \gtrsim Q_2$.

4. Regular Pieces

By saying that something depends on the geometry of F , we mean that it depends on the C^2 -norm of F . Below, all the constants depend only on the geometry of F unless otherwise is explicitly said.

The piece $B^k \equiv B_{v^k}^k$ of level $k \in \mathbb{N}$, a neighborhood of the tip, contains two pieces of level $k+1$, B^{k+1} , which is a smaller neighborhood of the tip, and the lateral one

$$E^k = B_{v^k}^{k+1}.$$

They are illustrated in Fig. 1, and more schematically, in Fig. 2.

For $n \geq k \geq 0$, let

$$\mathcal{B}^n[k] \equiv \mathcal{B}^n(F)[k] = \{B \in \mathcal{B}^n \mid B \subset E^k\}.$$

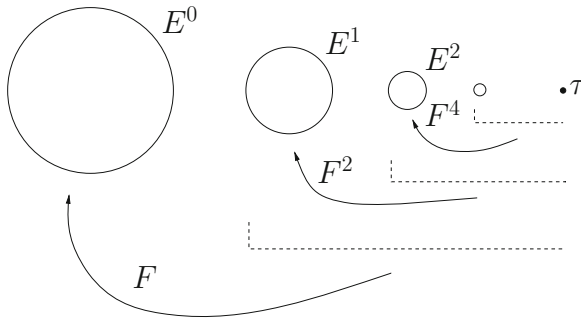


Fig. 2. An infinitely renormalizable Hénon-like map

We call k the *depth* of any piece $B \in \mathcal{B}^n[k]$. A piece B_ω^n belongs to $\mathcal{B}^n[k]$ if and only if

$$\omega = v^k c \omega_{k+2} \dots \omega_n.$$

Observe

$$\mu \left(\bigcup_{B \in \mathcal{B}^n[k]} B \right) = \mu(E^k) = \frac{1}{2^{k+1}},$$

where μ is the invariant measure on $\mathcal{O}_F$. Let

$$G_k = F^{2^k} : \bigcup_{l > k} E^l \rightarrow E^k, \quad k \geq 0.$$

Given a piece $B \in \mathcal{B}^n[k]$, there is a unique sequence

$$k = k_0 < k_1 < \dots < k_t = n, \quad k_i = k_i(B)$$

such that

$$B = G_{k_0} \circ G_{k_1} \circ \dots \circ G_{k_{t-1}} \circ G_{k_t}(B^n).$$

To see this, consider the backward orbit $\{F^{-m}B\}$ that brings B to the tip piece B^n . Let $F^{-m_i}(B)$ be the moments of its closest combinatorial approaches to the tip, in the sense of the nest $B^0 \supset B^1 \supset \dots$. Then k_i is the depth of $F^{-m_i}(B)$. Thus, $F^{-m_i}(B) \in E^{k_i}$, while $F^{-m}(B) \cap B^{k_i} = \emptyset$ for all $m < m_i$, compare with Sect. 2.2. The pieces

$$B^{(i)} := F^{-m_i}(B) = G_{k_i} \circ \dots \circ G_{k_{t-1}} \circ G_{k_t}(B^n) \in \mathcal{B}^n[k_i],$$

with $i = 1, 2, \dots, t$, are called *predecessors* of B . Let us view a piece $B = B_{v^k c \omega_{k+2} \dots \omega_n}^n \in \mathcal{B}^n[k]$ from *scale* k , i.e., let us consider the following piece $\mathbf{B}$ of depth 0 for the renormalization $F_k \equiv R^k F$:

$$\mathbf{B} = B_{c \omega_{k+2} \dots \omega_n}^{n-k}(F_k) \in \mathcal{B}^{n-k}(F_k)[0], \quad \text{so } B = \Psi_0^k(\mathbf{B}), \quad (4.1)$$

see Fig. 3.

Below, a “rectangle” means a quadrilateral with horizontal and vertical sides. Given a piece $B = B_\omega^n \in \mathcal{B}^n$, let us consider the smallest rectangle $Q = Q_\omega^n$ containing $B \cap \mathcal{O}_F$. We say that $Q = Q(B)$ is *associated with* B .

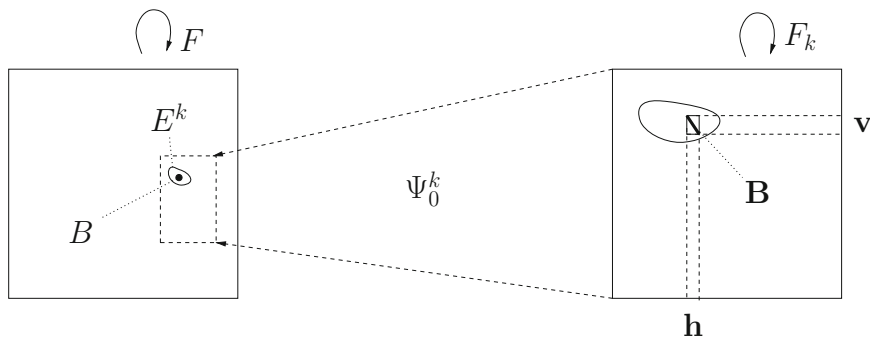


Fig. 3. A regular piece

Remark 4.1. We are primarily interested in the geometry of the Cantor attractor $\mathcal{O}_F$. For this reason we consider rectangles Q superscribed around $\mathcal{O}_F \cap B$ rather than the ones superscribed around the actual pieces B . However, our results apply to the latter rectangles as well.

Given $B \in \mathcal{B}^n[k]$, let us consider the rectangle Q associated to $B \in \mathcal{B}^{n-k}(F_k)$. Let h and v be its horizontal and vertical sizes of Q respectively. We also call them the *sizes of B viewed from scale k* .

The pieces of the n th-renormalization level $\mathcal{B}^n$ are defined as the iterations of the renormalization box B^n . In first instance these pieces are studied by using the iterates F^{2^k} . These high iterates correspond to the respective renormalizations. These renormalizations are very well understood and have an universal asymptotic shape. Hence, by considering pieces as they appear in the domain of the renormalizations one can study very precisely their shape. This is the reason why the pieces are viewed from their proper scale.

We say that the piece $B \in \mathcal{B}^n[k]$ is *regular* if these sizes are comparable, or, in other words, if $\text{mod } B := h/v$ is of order 1:

$$\frac{1}{C_0} \leq \text{mod } B \leq C_0, \quad (4.2)$$

with $C_0 = 3d_1$, where $d_1 > 0$ is the bound on $\partial F_k / \partial x$ from Corollary 3.5 (see Fig. 3³).

Notice that in the degenerate case, $F(x, y) = (f(x), x)$, where f is an infinitely renormalizable unimodal map, every piece is regular since the slope of f in E^0 is squeezed in between d_1 and $1/d_1$.

Next, we will specify an exponential control function $s(k) = s_\alpha(k) = a2^k - A$, see (2.8). Namely, we let

$$a = \frac{\ln b}{\ln \sigma}, \quad A = \frac{\ln \alpha}{\ln \sigma},$$

where $\alpha > 0$ is a small parameter. The actual choice of $\alpha = \alpha(\bar{\epsilon}) > 0$ depending only on the geometry of F will be made in the course of the paper (see Propositions 4.1, 5.1, etc.).

³ The rectangles in Figs. 3, 4 and 5 with solid lines are rectangles. The rectangles with dashed lines are essentially rectangles.

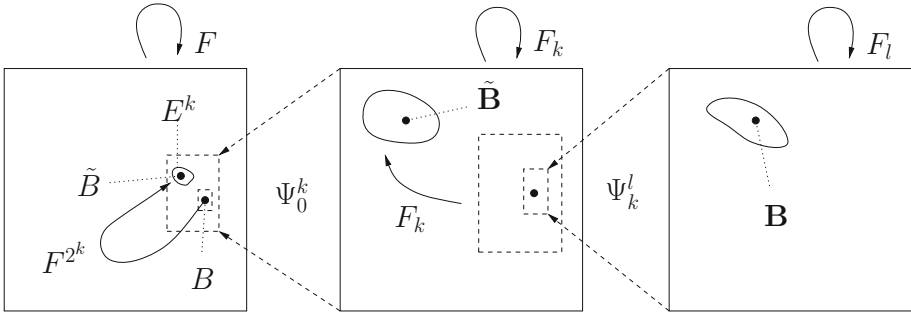


Fig. 4. Pieces B and $\tilde{B}$ viewed from appropriate scales

Let $l(k) = l_\alpha(k) = s(k) + k$. If $k \leq l \leq l(k)$ we say that the pieces $B \in \mathcal{B}^n[l]$ are *not too deep in B^k* . The choice of the control function is made so that

$$b^{2^k} \leq \alpha \sigma^{l-k} \quad \text{for } l \leq l_\alpha(k). \quad (4.3)$$

Proposition 4.1. *There exists $k^* \geq 0$ and $\alpha^* > 0$ such that for $\alpha < \alpha^*$ and $k \geq k^*$ the following holds. If $B \in \mathcal{B}^n[l]$ is regular and not too deep in B^k , $k < l \leq l_\alpha(k)$, then*

$$\tilde{B} = G_k(B) \in \mathcal{B}^n[k]$$

is regular as well.

Proof. We should view B from scale l , i.e., consider the piece $\mathbf{B} \in \mathcal{B}^{n-l}(F_l)[0]$ defined by (4.1). As the puzzle piece $\tilde{B} = F^{2^k}(B)$ has depth k , it should be viewed from this depth. So, we consider

$$\tilde{\mathbf{B}} \in \mathcal{B}^{n-k}(F_k)[0], \quad \text{such that } \Psi_0^k(\tilde{\mathbf{B}}) = \tilde{B}. \quad (4.4)$$

Observe: $\tilde{\mathbf{B}} = F_k \circ \Psi_k^l(\mathbf{B})$ (see Fig. 4).

As above, let $(\mathbf{h}, \mathbf{v})$ be the sizes of $\mathbf{B}$, and let $(\tilde{\mathbf{h}}, \tilde{\mathbf{v}})$ be the sizes of $\tilde{\mathbf{B}}$. Since B is regular, bound (4.2) holds for $\text{mod } \mathbf{B} = \mathbf{h}/\mathbf{v}$. We want to show that the same bound hold for $\text{mod } \tilde{\mathbf{B}} = \tilde{\mathbf{h}}/\tilde{\mathbf{v}}$.

The map Ψ_k^l factors into a non-linear and an affine part as described in Sect. 3:

$$\Psi_k^l = D_k^l \circ (\text{id} + \mathbf{S}_k^l).$$

Figure 5 shows⁴ details of this factorization for the map Ψ_k^l from Fig. 4. Let h_{diff} and v_{diff} be the sizes of the rectangle Q_{diff} associated with the piece $(\text{id} + \mathbf{S}_k^l)(\mathbf{B})$, see Fig. 5. Lemmas 3.9(1) and 3.10 imply for k big enough:

$$h_{\text{diff}} \leq d_0 \cdot \mathbf{h} + O(\bar{\epsilon}^{2^k}) \cdot \mathbf{v} \leq 2d_0 \cdot \mathbf{h},$$

⁴ The diagonals in the boxes in Fig. 5 refer to the sticks introduced in Sect. 5.

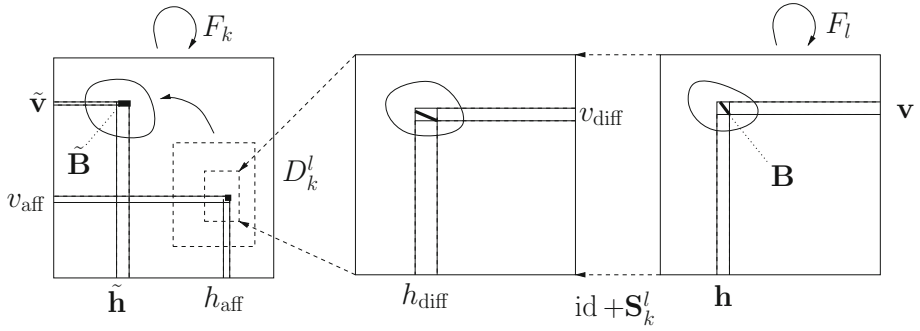


Fig. 5. Factorization of the map Ψ_k^l into horizontal and affine parts

where the last estimate takes into account (4.2). Similarly,

$$h_{\text{diff}} \geq \frac{1}{2d_0} \mathbf{h}. \quad (4.5)$$

Moreover, since the map $\text{id} + \mathbf{S}_k^l$ is horizontal, we have

$$v_{\text{diff}} = \mathbf{v} \leq C_0 \cdot \mathbf{h}. \quad (4.6)$$

Let h_{aff} and v_{aff} be the sizes of the rectangle Q_{aff} associated with the piece $B_{\text{aff}} = \Psi_k^l(\mathbf{B}) = D_k^l \circ (\text{id} + \mathbf{S}_k^l)(\mathbf{B})$ (which is the piece B viewed from scale k). Incorporating the above estimates into decomposition (3.6) and using Lemma 3.8, we obtain for large k (with $s = l - k$):

$$\begin{aligned} h_{\text{aff}} &\leq (h_{\text{diff}} \cdot \sigma^{2s} + v_{\text{diff}} \cdot |t_k| \cdot \sigma^s) \cdot (1 + O(\rho^k)) \\ &\leq [3d_0 \cdot \sigma^s + O(b^{2^k})] \cdot \sigma^s \cdot \mathbf{h}. \end{aligned}$$

Similarly, we obtain a lower bound for h_{aff} :

$$h_{\text{aff}} \geq \left[\frac{1}{3d_0} \cdot \sigma^s - O(b^{2^k}) \right] \cdot \sigma^s \cdot \mathbf{h}.$$

If B is not too deep for scale k , then $b^{2^k} \leq \alpha \sigma^s$, and we obtain:

$$h_{\text{aff}} \asymp \sigma^{2s} \cdot \mathbf{h}, \quad (4.7)$$

as long as α is small enough (depending on the geometry of F_k).

Bounds on v_{aff} are obtained similarly (in fact, easier):

$$v_{\text{aff}} = v_{\text{diff}} \cdot \sigma^{l-k} \cdot (1 + O(\rho^k)) = \mathbf{v} \cdot \sigma^s \cdot (1 + O(\rho^k)) \asymp \mathbf{v} \cdot \sigma^s. \quad (4.8)$$

Thus,

$$\text{mod } B_{\text{aff}} = \text{mod } \Psi_k^l(\mathbf{B}) \asymp \sigma^s \text{ mod } \mathbf{B}. \quad (4.9)$$

it gets roughly aligned with the parabola-like curve inside E^k , which makes its modulus of order 1. Furthermore, Theorem 3.4 and Corollary 3.5 imply, for k large enough, the following bounds on the sizes of $\tilde{\mathbf{B}}$:

$$\frac{1}{2d_1} h_{\text{aff}} - A_0 b^{2^k} \cdot v_{\text{aff}} \leq \tilde{\mathbf{h}} \leq 2d_1 \cdot h_{\text{aff}} + A_0 b^{2^k} \cdot v_{\text{aff}},$$

$$\tilde{\mathbf{v}} = h_{\text{aff}},$$

where $A_0 > 0$ is an upper bound for $a(x)(1 + O(\rho^k))$ which controls the vertical derivative of F_k . Hence

$$\text{mod } \tilde{\mathbf{B}} \leq 2d_1 + \frac{A_0 b^{2^k}}{\text{mod } \Psi_k^l(\mathbf{B})} \leq 2d_1 + \frac{A_0 b^{2^k}}{\sigma^s \text{mod } \mathbf{B}} \leq 2d_1 + A_0 C_0 \alpha \leq 3d_1,$$

as long as α is small enough. Notice that $\text{mod } \mathbf{B}$ appears only in the residual term of the last estimate. The main term ($2d_1$) depends only on the geometry of F , which makes the bound for $\text{mod } \tilde{\mathbf{B}}$ as good as that for $\mathbf{B}$.

The lower estimate, $\text{mod } \tilde{\mathbf{B}} \geq (3d_1)^{-1}$, is similar. $\square$

5. Sticks

Let us consider a piece $B \in \mathcal{B}^n[l]$ and the corresponding piece $\mathbf{B} \in \mathcal{B}^{n-l}(F_l)[0]$, see (4.1) and Fig. 3. In the degenerate case the pieces $\mathbf{B} \cap \mathcal{O}_{F_l}$ get squeezed in a narrow strip around the diagonal of the associated rectangle $\mathbf{Q}$ (except a few pieces directly near the tip where the pieces are small bended curves and the renormalization box which is a rectangle). We will show that this is also the case for many pieces of Hénon perturbations. To this end, let us quantify the thickness of the pieces in question.

Let us first introduce two standard strips of thickness δ :

$$\Delta_\delta^\pm = \{(x, y) \in [0, 1]^2 \mid |y \pm x| \leq \frac{\delta}{2}\}$$

(oriented “north-west” and “north-east” respectively.)

Given a piece $B \in \mathcal{B}^n$ and the associated rectangle $Q = Q(B)$, let $L : [0, 1]^2 \rightarrow Q$ be the orientation preserving diagonal surjective affine map. Let $\Delta(B) = L(\Delta_\delta^\pm)$, where:

- we select the “+”-sign if B comes from the upper branch of the parabola $x = f(y)$, and “-”-sign otherwise.
- $\delta = \delta_B$ is selected to be the smallest one such that $\Delta(B) \supset B \cap \mathcal{O}$.

This δ_B is called the (*relative*) *thickness* of B . The horizontal size $h\delta_B$ of $\Delta(B)$ is called the *absolute thickness* of B . $\Delta(B)$ is called the associated diagonal strip. We let $\Delta \equiv \Delta_{\mathbf{B}}$ and call it the *regular stick* associated with B , see Fig. 6.

Proposition 5.1. *There exists $k^* \geq 0$ and $\alpha^* > 0$ such that for $\alpha < \alpha^*$ and $k^* \leq k$ the following holds. If $B \in \mathcal{B}^n[l]$ is regular and not too deep in B^k , $k < l \leq l_\alpha(k)$, then*

$$\delta_{\tilde{\mathbf{B}}} \leq \frac{1}{2} \cdot \delta_{\mathbf{B}} + O(\sigma^{n-l}),$$

where $\tilde{B} = G_k(B) \in \mathcal{B}^n[k]$ and $\tilde{\mathbf{B}} = F_k(\Psi_k^l(\mathbf{B}))$.

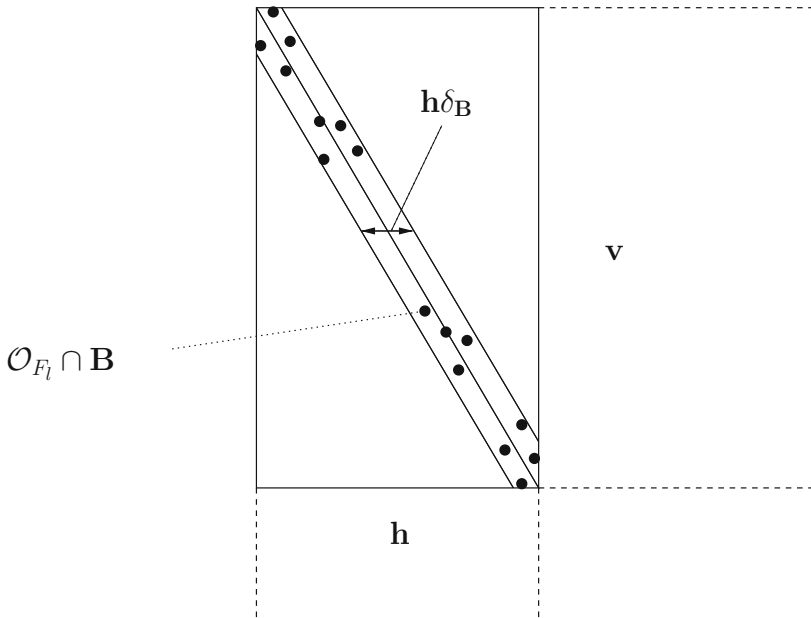


Fig. 6. Regular stick

Proof. We will use the notation from Sect. 4. Let $\delta \equiv \delta_{\mathbf{B}}$, and let $\mathbf{w} = \delta \cdot \mathbf{h}$ be the absolute thickness of $\mathbf{B}$. The relative thickness of $\tilde{\mathbf{B}}$ is denoted by $\tilde{\delta} \equiv \delta_{\tilde{\mathbf{B}}}$. To estimate $\tilde{\delta}$, we will decompose Ψ_k^l as in §4. Let w_{diff} be the absolute thickness of $B_{\text{diff}} \equiv (\text{id} + \mathbf{S}_k^l)(\mathbf{B})$. Then

$$w_{\text{diff}} = O(\mathbf{w} + \mathbf{h} \cdot \sigma^{n-l}). \quad (5.1)$$

Indeed, let Γ_y be the horizontal section of $(\text{id} + \mathbf{S}_k^l)(\Delta_{\mathbf{B}})$ on height y , i.e. the intersection with the horizontal straight line at height y , and let $\Gamma_y = (\text{id} + \mathbf{S}_k^l)^{-1}(\Gamma_y)$. Then

$$|\Gamma_y| \leq |\Gamma_y| \cdot \|\text{id} + \mathbf{S}_k^l\|_{C^1} = O(\mathbf{w}),$$

where the last estimate follows from Lemma 3.9(1).

Furthermore, let us consider a boundary curve of $(\text{id} + \mathbf{S}_k^l)(\Delta_{\mathbf{B}})$. Its horizontal deviation from any of its tangent lines is bounded by

$$\frac{1}{2} \|\text{id} + \mathbf{S}_k^l\|_{C^2} \cdot (\text{diam } \mathbf{B})^2 = O(\sigma^{n-l}) \cdot \mathbf{h}, \quad (5.2)$$

where the last estimate follows from Lemma 3.9 (2), bounded modulus of $\mathbf{B}$ (4.2), and Lemma 3.1. Hence

$$w_{\text{diff}} \leq \max_y |\Gamma_y| + O(\sigma^{n-l}) \cdot \mathbf{h},$$

and (5.1) follows. Together with (4.5), it implies:

$$\delta_{\text{diff}} = O(\delta + \sigma^{n-l}). \quad (5.3)$$

Let us now consider the piece $B_{\text{aff}} \equiv \Psi_k^l(\mathbf{B}) = D_k^l(B_{\text{diff}})$, see Fig. 5. Let $D_k^l = T \circ \Lambda$, where $\Lambda = \Lambda_k^l$ and $T = T_k^l$ are the diagonal and shear parts of D_k^l respectively, see (3.6). Let us consider a box $B_{\text{diag}} = \Lambda(B_{\text{diff}})$, and let $h_{\text{diag}} = \sigma^{2(l-k)} h_{\text{diff}}$ and $v_{\text{diag}} = \sigma^{l-k} v_{\text{diff}}$ be its horizontal and vertical sizes. Since diagonal affine maps preserve the horizontal thickness, the thickness is only effected by the shear part T_k^l , which has order $t_k \asymp b^{2k}$, see Lemma 3.8, namely:

$$\begin{aligned} \delta_{\text{aff}} &\leq \delta_{\text{diff}} \cdot \frac{1}{1 - \frac{v_{\text{diag}}}{h_{\text{diag}}} \cdot t_k} \\ &= \delta_{\text{diff}} \cdot \frac{1}{1 - \frac{v_{\text{diff}}}{h_{\text{diff}}} \cdot \sigma^{-(l-k)} \cdot t_k} \\ &= O(\delta_{\text{diff}}) = O(\delta + \sigma^{n-l}). \end{aligned} \quad (5.4)$$

where the passage to the last line comes from (4.3), (4.5), (4.6) and Lemma 3.8. Let us also consider the absolute *vertical thickness* u_{aff} of B_{aff} , i.e., the vertical size of the stick $\Delta(B_{\text{aff}})$. From triangle similarity, we have:

$$\frac{u_{\text{aff}}}{v_{\text{aff}}} = \frac{w_{\text{aff}}}{h_{\text{aff}}}$$

So

$$u_{\text{aff}} = \frac{w_{\text{aff}}}{\text{mod } B_{\text{aff}}} \asymp \frac{w_{\text{aff}}}{\sigma^s \text{ mod } \mathbf{B}} \asymp \sigma^{-s} \cdot w_{\text{aff}}, \quad s = l - k, \quad (5.5)$$

where the last estimate follows from regularity of B while the previous one comes from (4.9).

We are now prepared to apply the map $F_k : (x, y) \mapsto (f_k(x) - \epsilon_k(x, y), x)$, where $\|\epsilon_k\|_{C^2} = O(2^{b^k})$, see Theorem 3.4. Let $\tilde{\mathbf{w}}$ be the absolute thickness of $\mathbf{B}$. By (4.7)–(4.8), the rectangle Q_{aff} associated with B_{aff} has sizes

$$v_{\text{aff}} \asymp \sigma^s \mathbf{v} \quad \text{and} \quad h_{\text{aff}} \asymp \sigma^{2s} \mathbf{h}.$$

Let us use affine parametrization for the diagonal Z of B_{aff} :

$$x = x_0 + t, \quad y = y_0 + \frac{C}{\sigma^s} t, \quad 0 \leq t \leq h_{\text{aff}},$$

where x_0, y_0 is its corner where the stick Δ_{aff} begins. Restrict F_k to this diagonal:

$$F_k(x(t), y(t)) = (A + Bt + E(t), x_0 + t),$$

where $E(t)$ is the second order deviation of $F_k(Z)$ from the straight line. We obtain:

$$\begin{aligned} E(t) &= O\left(\left\|\frac{\partial^2(f_k - \epsilon_k)}{\partial x^2}\right\| + \left\|\frac{\partial^2 \epsilon_k}{\partial x y}\right\| \cdot \sigma^{-s} + \left\|\frac{\partial^2 \epsilon_k}{\partial y^2}\right\| \cdot (\sigma^{-s})^2\right) \cdot h_{\text{aff}}^2 \\ &= O(h_{\text{aff}} + b^{2k} \sigma^{-s} h_{\text{aff}} + (b^{2k} \sigma^{-s}) \cdot (\sigma^{-s} h_{\text{aff}})) \cdot h_{\text{aff}}. \end{aligned}$$

From Lemma 3.1 we have $h_{\text{aff}} = O(\sigma^{n-k})$. Hence,

$$\begin{aligned} E(t) &= O(\sigma^{n-k} + b^{2k} \sigma^{-(l-k)+n-k} + \alpha \cdot \sigma^{-(l-k)+n-k}) \cdot h_{\text{aff}} \\ &= O(\sigma^{n-l}) \cdot h_{\text{aff}} \end{aligned}$$

where we used that l is not too deep for k , i.e. $b^{2^k} \sigma^{-s} \leq \alpha$. It follows that

$$\begin{aligned}\tilde{\mathbf{w}} &= O(\sigma^{n-l} \cdot h_{\text{aff}} + b^{2^k} \cdot u_{\text{aff}}) \\ &= O(\sigma^{n-l} \cdot h_{\text{aff}} + b^{2^k} \sigma^{-s} \cdot w_{\text{aff}}) \\ &= O(\sigma^{n-l} \cdot h_{\text{aff}} + \alpha \cdot w_{\text{aff}}),\end{aligned}$$

where we used (5.5).

From the regularity of $\tilde{\mathbf{B}}$ we get $\tilde{\mathbf{h}} \asymp \tilde{\mathbf{v}} = h_{\text{aff}}$. Thus,

$$\begin{aligned}\tilde{\delta} &= O(\sigma^{n-l} + \alpha \cdot \delta_{\text{aff}}) \\ &= O(\alpha \cdot \delta + \sigma^{n-l})\end{aligned}$$

where we used (5.4). The Proposition follows as long as α is sufficiently small. $\square$

6. Scaling

Let $B \in \mathcal{B}^n[k]$ and $\hat{B} \in \mathcal{B}^{n-1}[k]$ with $B \subset \hat{B}$. Say,

$$B = B_{\omega v}^n \subset \hat{B} = B_{\omega}^{n-1} \subset E^k.$$

Let $\mathbf{B}$ and $\hat{\mathbf{B}}$ be the corresponding rescaled pieces, so $B = \Psi_0^k(\mathbf{B})$ and $\hat{B} = \Psi_0^k(\hat{\mathbf{B}})$. The horizontal and vertical sizes of the associated rectangles are called $\mathbf{h}, \mathbf{v} > 0$ and $\hat{\mathbf{h}}, \hat{\mathbf{v}} > 0$ respectively (Fig. 7).

The *scaling number* of B is

$$\sigma_{\mathbf{B}} = \frac{\mathbf{v}}{\hat{\mathbf{v}}}.$$

Remark 6.1. The scaling number can be expressed directly in terms of the original pieces B and $\hat{B}$. Indeed, since the diffeomorphism Ψ_0^k is a horizontal map, we have $\sigma_{\mathbf{B}} = v/\hat{v}$, where v and $\hat{v}$ are the vertical sizes of B and $\hat{B}$. We will use the notation σ_B when we refer to the corresponding measurement in the domain of F . This formal distinction will only play a role in (7.20).

Remark 6.2. There are many possible ways to define the scaling number. The proof of the probabilistic universality will show that the relative thickness of most pieces asymptotically vanishes. Because of this, most definitions of the scaling number become equivalent.

For $B = B_{\omega v}^n(F)$ as above, let $B^* = B_{\omega v}^n(F_*)$ be the corresponding degenerate piece for the renormalization fixed point F_* . The *proper scaling* for B is

$$\sigma_{\mathbf{B}}^* = \sigma_{B_{\omega v}^n(F_*)}.$$

The function

$$\underline{\sigma} : B \mapsto \sigma_{\mathbf{B}}$$

is called the *scaling function* of F . The universal scaling function $\underline{\sigma}^*$ of F_* is injective, as was shown in [2].

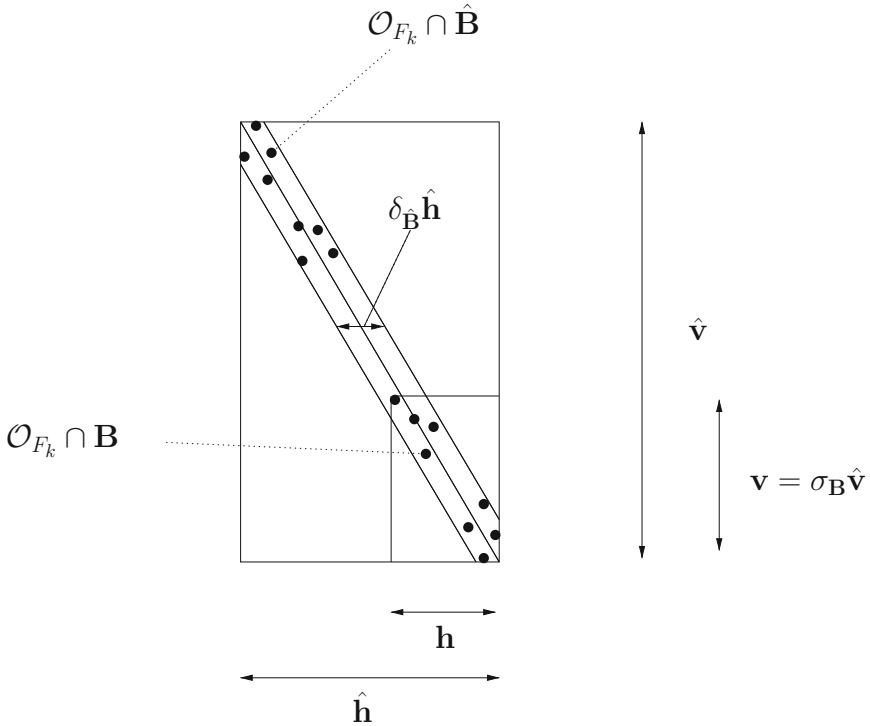


Fig. 7. The scaling number

Remark 6.3. Given a piece $B \in \mathcal{B}^{n+1}(F_*)$. Let $\hat{B}^* \in \mathcal{B}^n(F_*)$ which contains B . For some $\hat{i} < 2^n$ we have

$$\pi_1(\hat{B}^*) = I_{\hat{i}}^*(n) \in \mathcal{C}_n.$$

Similarly, $\pi_1(B^*) = I_i^*(n+1) \in \mathcal{C}_{n+1}$, for $i = \hat{i}$ or $i = \hat{i} + 2^n$. The scaling ratios σ_B are vertical measurements of pieces. Using that Hénon like maps take vertical lines to horizontal lines, $y' = x$, we have

$$\sigma_B^* = \frac{|I_{i-1}^*(n+1)|}{|I_{i-1}^*(n)|}.$$

Proposition 6.1. *There exists $k^* \geq 0$ and $\alpha^* > 0$ such that for $\alpha < \alpha^*$ and $k \geq k^*$ the following holds. If a piece $B \in \mathcal{B}^n[l]$ is regular and not too deep for E_k , i.e. $k < l \leq l_\alpha(k)$, then*

$$\sigma_{\tilde{B}} = \sigma_B + O(\delta_{\tilde{B}} + \sigma^{n-l}),$$

where $\tilde{B} = G_k(B) \in \mathcal{B}^n[k]$ and $B \subset \hat{B} = \Psi_0^l(\tilde{B}) \in \mathcal{B}^{n-1}[l]$.

Proof. As above in Sect. 4, let h_{aff} stand for the horizontal length of $B_{\text{aff}} = \Psi_k^l(\mathbf{B})$, see Fig. 5. We will use the similar notation $\hat{h}_{\text{aff}}$ and $\hat{w}_{\text{aff}}$ for the corresponding measurements of the piece $\hat{B}_{\text{aff}} := \Psi_k^l(\hat{\mathbf{B}})$.

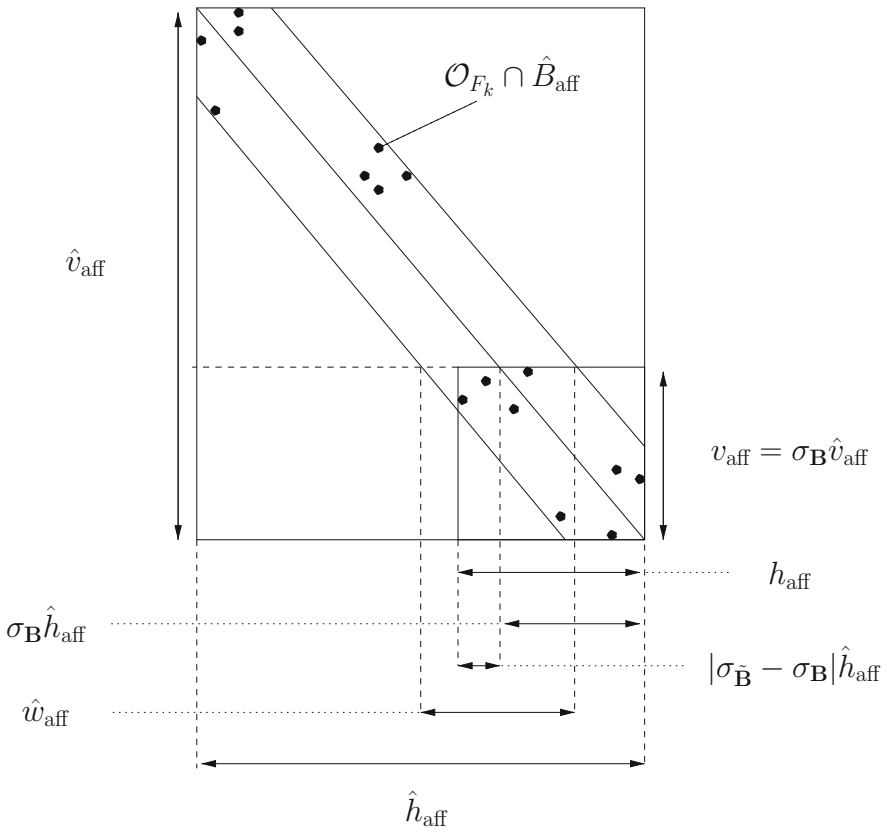


Fig. 8. Illustration for the proof of Proposition 6.1

Since F_k maps vertical lines to horizontal lines, we have

$$\sigma_{\tilde{\mathbf{B}}} = \frac{h_{\text{aff}}}{\hat{h}_{\text{aff}}}.$$

Let γ be the angle between the diagonal of $\hat{B}_{\text{aff}}$ and the vertical side, so $\text{tg } \gamma = \text{mod } \hat{B}_{\text{aff}}$. Then

$$v_{\text{aff}} \cdot \text{tg } \gamma = \hat{h}_{\text{aff}} \frac{v_{\text{aff}}}{\hat{v}_{\text{aff}}} = \hat{h}_{\text{aff}} \cdot \sigma_{\mathbf{B}},$$

Now Fig. 8 shows:

$$|h_{\text{aff}} - v_{\text{aff}} \cdot \text{tg } \gamma| \leq \hat{w}_{\text{aff}}.$$

Dividing by $\hat{h}_{\text{aff}}$ (taking into account the two previous formulas and definition of the relative thickness $\hat{\delta}_{\text{aff}} = \hat{w}_{\text{aff}}/\hat{h}_{\text{aff}}$), we obtain:

$$|\sigma_{\tilde{\mathbf{B}}} - \sigma_{\mathbf{B}}| \leq \hat{\delta}_{\text{aff}}.$$

Now the Proposition follows from (5.4). $\square$

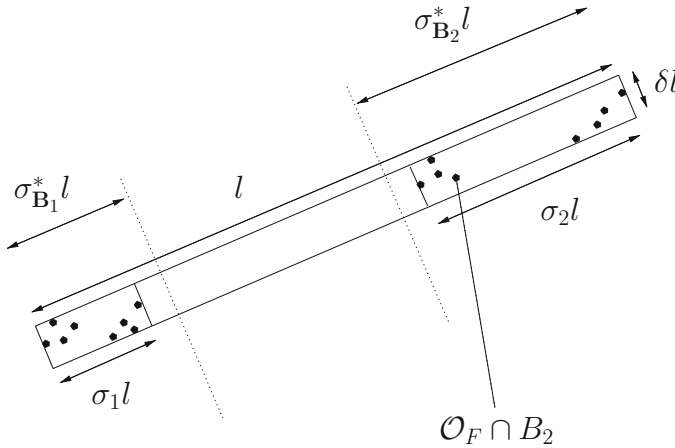


Fig. 9. ϵ -universal stick

7. Universal Sticks

7.1. Definition and statement. Let us consider a piece $B \in \mathcal{B}^n$ and the two pieces $B_1, B_2 \in \mathcal{B}^{n+1}$ of level $n+1$ contained in B . Rotate it to make it horizontal and then rescale it to horizontal size 1; denote the corresponding linear conformal map by A . Let $\delta, \sigma_{B_1}, \sigma_{B_2} \geq 0$ be the smallest numbers such that:

- (1) $A(B \cap \mathcal{O}_F) \subset [0, 1] \times [0, \delta]$,
- (2) $A(B_1 \cap \mathcal{O}_F) \subset [0, \sigma_{B_1}] \times [0, \delta]$,
- (3) $A(B_2 \cap \mathcal{O}_F) \subset [1 - \sigma_{B_2}, 1] \times [0, \delta]$,

for the optimal choice of A . The numbers σ_{B_1} , and σ_{B_2} are called *scaling factors* of B_1 and B_2 .

Remark 7.1. The scaling factor σ_B of a piece B is a measurement of the corresponding B . The scaling factor σ_B of B reverts to measurements of the actual piece in the domain of F . The difference between the scaling factors σ_B and σ_B is estimated in Proposition 7.7.

We say that B is ϵ -universal if

$$|\sigma_{B_1} - \sigma_{B_1}^*| \leq \epsilon, \quad |\sigma_{B_2} - \sigma_{B_2}^*| \leq \epsilon, \quad \text{and} \quad \delta \leq \epsilon.$$

The *precision* of the piece B is the smallest $\epsilon > 0$ for which B is ϵ -universal. The optimal $A^{-1}([0, 1] \times [0, \delta])$ is called the ϵ -stick for B (Fig. 9). We will revert to the (relative) length and (relative) height of such a stick. Let $\mathcal{S}^n(\epsilon) \subset \mathcal{B}^n$ be the collection of ϵ -universal pieces.

Definition 7.1. The Cantor attractor $\mathcal{O}_F$ of an infinitely renormalizable Hénon-like map $F \in \mathcal{H}_\Omega(\bar{\epsilon})$ is *probabilistically universal* if there is $\theta < 1$ such that

$$\mu(\mathcal{S}^n(\theta^n)) \geq 1 - \theta^n, \quad n \geq 1.$$

Now we can formulate the main result of this paper:

Theorem 7.1. *The Cantor attractor $\mathcal{O}_F$ is probabilistically universal.*

After careful choices of $\theta < 1$, $q_0 < q_1$ and $\kappa(n) = -\text{Const} + \ln n$, one distinguishes three regimes where pieces in $\mathcal{S}_n(\theta^n) \cap E^k$ are discovered by different techniques.

The *one-dimensional regime*: all the pieces in $\mathcal{B}^n[k]$ with $(1 - q_1) \cdot n \leq k \leq (1 - q_0) \cdot n$ are in $\mathcal{S}_n(\theta^n)$. These very deep pieces are controlled by the one-dimensional renormalization fixed point: they are perturbed versions the corresponding pieces of F_* and their relative displacements are exponentially small, see Lemma 7.3 and Proposition 7.2. We have to exclude the pieces in $\mathcal{B}^n[k]$ with $k > (1 - q_0) \cdot n$ because they do not have a small thickness. Viewed from their scale k , they are relatively large pieces close to the graph of f_* . The curvature of the graph of f_* causes this pieces to have a large thickness.

The *pushing-up regime*: the pieces from the one-dimensional regime can be pushed up without being distorted too much, using the Propositions 5.1 and 6.1, as long as they are not too deep. The resulting pieces have exponentially small precision, see Proposition 7.7. In this way one finds pieces in $\mathcal{S}_n(\theta^n) \cap E^k$ for $0 \leq k < (1 - q_1) \cdot n$. Unfortunately, the relative measure of these pieces in $\mathcal{S}_n(\theta^n) \cap E^k$ obtained by pushing up, is only exponentially close to 1, for $k \geq \kappa(n) \asymp \ln n$, see Proposition 8.2. That is why the pushing-up regime is restricted to $\kappa(n) \leq k < (1 - q_1) \cdot n$ where these pieces occupy E^k except for an exponential small relative part.

The *brute-force regime*: the pieces obtained in the one-dimensional and pushing-up regimes are in $B^{\kappa(n)}$. They will be spread around by brute-force iteration of the original map until returning. The time to go from $B^{\kappa(n)}$ and return by iterating the original map is $2^{\kappa(n)}$. The depth $\kappa(n)$ is the largest integer such that $2^{\kappa(n)} \leq Kn \ln 1/\theta$. The pieces in the one-dimensional and pushing-up regime have exponentially small precision. Each of the brute-force return steps used to spread around the pieces from the deeper regimes, will distort their exponential precision θ^n , see Proposition 7.8. The total distortion along such a return orbit can be bounded by $O(r^{2^{\kappa(n)}}) = O(r^{Kn \ln 1/\theta})$, with $r \gtrsim 1/b \gg 1$. However, this distortion can not destroy the exponential precision when $\theta < 1$ is chosen close enough to 1.

The pushing-up regime is split into two parts. Let $\kappa_0(n)$ be the smallest integer such that $l(\kappa_0(n)) \geq n$. As long as $\kappa_0(n) \leq k < (1 - q_1) \cdot n$ the pieces in $\mathcal{B}^n[l]$, $l > k$, are not too deep and can be pushed up into E^k . Indeed, $\kappa_0(n) \asymp \ln n$ is uniquely defined and can not be adjusted. Unfortunately, we can not use $\kappa(n) = \kappa_0(n)$ because the corresponding return time $2^{\kappa_0(n)}$ used to fill the brute-force regime might be too large. Too large in the sense that it might build up too much distortion, which is of the order $O(r_0^n)$ for some definite $r_0 > 1$. We have to choose $\kappa(n) \asymp \ln n$ much smaller than $\kappa_0(n)$ to get an arbitrarily slow growing rate for the distortion during the brute-force regime. The rate should be small enough such that the exponential decaying precision in the deeper regimes can not be destroyed. In the regime $\kappa(n) \leq k < \kappa_0(n)$ we have $l(k) < (1 - q_1) \cdot n$ which means that we can not push up all previously recovered pieces in $\mathcal{B}^n[l]$ with $l > l(k)$. This is responsible for the super-exponential loss term in Proposition 8.2.

7.2. Universal sticks created in the one-dimensional regime.

Proposition 7.2. *There exist $\rho < 1$, $q^* > 0$ with the following property. For every $0 < q_0 < q_1 \leq q^*$ there exists $n^* > 0$ such that for $n \geq n^*$ and $(1 - q_1) \cdot n \leq k \leq n$*

(1) every $B \in \mathcal{B}^n[k]$ is regular.

(2) for every $B \in \mathcal{B}^{n+1}[k]$

$$|\sigma_{\mathbf{B}} - \sigma_{\mathbf{B}}^*| = O(\rho^{q_1 \cdot n}),$$

where $B = \Psi_0^{k+1}(\mathbf{B})$.

(3) for every $B \in \mathcal{B}^n[k]$ with $(1 - q_1) \cdot n \leq k \leq (1 - q_0) \cdot n$

$$\delta_{\mathbf{B}} = O(\rho^{q_0 \cdot n}),$$

where $B = \Psi_0^k(\mathbf{B})$.

Choose, $(1 - q_1) \cdot n \leq k \leq n$ and $B \in \mathcal{B}^n[k]$. Let $\mathbf{B} \in \mathcal{B}^{n-k}(F_k)$ be such that $B = \Psi_0^k(\mathbf{B})$. Let τ_n be the tip of F_n and τ_* the tip of F_* . In the next part we will have to compare the maps Ψ_k^n related to F and the maps $(\Psi_k^n)^*$ corresponding to F_* . Let

$$\mathbf{B}_0 = B_{v^{n-k}}^{n-k}(F_k) = \Psi_k^n(\text{Dom}(F_n))$$

and

$$\mathbf{B}_0^* = B_{v^{n-k}}^{n-k}(F_*) = (\Psi_0^{n-k})^*(\text{Dom}(F_*)).$$

where $(\Psi_0^{n-k})^*$ is the change of coordinates used to construct $R^{n-k}F_*$. Then $\mathbf{B} = F_k^j(\mathbf{B}_0)$ for some $0 \leq j < 2^{n-k}$ and j is odd. Let $\mathbf{B}_j = F_k^j(\mathbf{B}_0)$ and $\mathbf{B}_j^* = F_*^j(\mathbf{B}_0^*)$ for $0 \leq j < 2^{n-k}$. We will analyse the relative positions of $\mathbf{B}_j$ and $\mathbf{B}_j^*$. Let

$$I_j = \pi_1(\mathbf{B}_j) \quad \text{and} \quad J_j = \pi_2(\mathbf{B}_j).$$

The intervals in the n th level of f_* are denoted by $I_j^*(n)$, see Sect. 3.1. Observe,

$$I_j^* \equiv I_j^*(n - k) = \pi_1(\mathbf{B}_j^*), \quad 0 \leq j < 2^{n-k}.$$

and

$$J_j^* = \pi_2(\mathbf{B}_j^*) = I_{j-1}^*(n - k), \quad 0 < j < 2^{n-k}.$$

Consider the conjugations

$$h_n : \mathcal{O}_{F_*} \rightarrow \mathcal{O}_{F_n}$$

with $h_n(\tau_*) = \tau_n$. These conjugations allow us to label the points in $\mathcal{O}_{F_n}$. Choose, $z^* \in \mathcal{O}_{F_*}$ and let $z = h_n(z^*)$. Let $(x_0, y_0) = \Psi_k^n(z) \in \mathbf{B}_0$ and $(x_0^*, y_0^*) = (\Psi_k^n)^*(z^*) \in \mathbf{B}_0^*$. The points in the orbits are

$$(x_j, y_j) = F_k^j(x_0, y_0) \quad \text{and} \quad (x_j^*, y_j^*) = F_*^j(x_0^*, y_0^*),$$

with $0 \leq j < 2^{n-k}$. The first estimates will be on the relative displacements $\frac{\Delta x_j}{|I_j^*|}$ and $\frac{\Delta y_j}{|J_j^*|}$ where $\Delta x_j = x_j - x_j^*$ and $\Delta y_j = y_j - y_j^*$.

Lemma 7.3. *There exist $\rho < 1$, $q^* > 0$ with the following property. For every $0 < q \leq q^*$ there exists $n^* > 0$ such that for $n \geq n^*$, $(1 - q) \cdot n \leq k \leq n$, and $0 \leq j < 2^{n-k}$*

$$\frac{|\Delta x_j|}{|I_j^*|} = O(\rho^{q \cdot n}), \quad \text{and} \quad \frac{|\Delta y_j|}{|J_j^*|} = O(\rho^{q \cdot n}).$$

Proof. Recall, $y_{j+1} = x_j$. Hence,

$$\frac{|\Delta y_{j+1}|}{|J_{j+1}^*|} = \frac{|\Delta x_j|}{|I_j^*|},$$

we only have to estimate the displacements Δx_j and Δy_0 . Since, $F_k \rightarrow F_*$ exponentially fast controlled by some $\rho < 1$, see Theorem 3.4, we have

$$\begin{aligned} x_{j+1} &= f_*(x_j) + O(\rho^k) \\ &= f_*(x_j^*) + Df_*(\zeta_j)\Delta x_j + O(\rho^k). \end{aligned}$$

Hence,

$$\Delta x_{j+1} = Df_*(\zeta_j)\Delta x_j + O(\rho^k).$$

There exists $K > 1$ such that

$$\frac{|\Delta x_{j+1}|}{|I_{j+1}^*|} \leq \frac{Df_*(\zeta_j)}{\frac{|I_{j+1}^*|}{|I_j^*|}} \cdot \frac{|\Delta x_j|}{|I_j^*|} + K \frac{\rho^k}{\rho_0^{n-k}}, \quad (7.1)$$

where we used the *a priori* bounds: $|I_{j+1}^*| \geq \rho_0^{n-k}$ for some $\rho_0 < 1$.

We will use (7.1) repeatedly but to do so we first need to estimate $|\Delta x_0|$. Let $\Delta z = z - z^*$ and use the Lemmas 3.11, 3.1, and 3.6 in the following estimate

$$\begin{aligned} |(x_0, y_0) - (x_0^*, y_0^*)| &\leq |\Psi_k^n(z) - (\Psi_k^n)^*(z^*)| \\ &\leq |\Psi_k^n - (\Psi_k^n)^*| + |(\Psi_k^n)^*(z) - (\Psi_k^n)^*(z^*)| \\ &\leq O(\rho^k) + |D(\Psi_k^n)^*| \cdot |\Delta z| \\ &= O(\rho^k + \sigma^{n-k} \cdot \rho^n) \\ &= O(\rho^k). \end{aligned}$$

Thus,

$$\frac{|\Delta x_0|}{|I_0^*|} = O\left(\frac{\rho^k}{\rho_0^{n-k}}\right) \quad (7.2)$$

and

$$\frac{|\Delta y_0|}{|J_0^*|} = O\left(\frac{\rho^k}{\rho_0^{n-k}}\right). \quad (7.3)$$

Let $r > 0$ and $D > 1$ be given as in Lemma 3.2 and $K > 1$ as defined above. For $q > 0$ small enough and $n \geq 1$ large enough we have

$$\frac{|\Delta x_0|}{|I_0^*|} = O\left(\frac{\rho^k}{\rho_0^{n-k}}\right) = O\left(\left(\frac{\rho^{1-q}}{\rho_0^q}\right)^n\right) = O(\rho^{q \cdot n}) \leq \frac{r}{2D}. \quad (7.4)$$

and

$$DK\left(\frac{2}{\rho_0}\right)^{n-k} \cdot \rho^k = O\left(\left(\frac{\rho^{1-q}}{(\rho_0/2)^q}\right)^n\right) = O(\rho^{q \cdot n}) \leq \frac{r}{2}. \quad (7.5)$$

One has to be careful when applying (7.1) repeatedly. The points ζ_j should not be too far from I_j^* to be able to control distortion.

Claim 7.4. For $q > 0$ small enough and $n > 1$ large enough

$$\frac{|\Delta x_j|}{|I_j^*|} \leq DK \left(\frac{2}{\rho_0}\right)^{n-k} \cdot \rho^k + D \frac{|\Delta x_0|}{|I_0^*|},$$

for $0 \leq j < 2^{n-k}$.

Proof. The proof is by induction: the statement holds for $j = 0$ because $D > 1$. Suppose it holds up to $j < 2^{n-k} - 1$. The r -neighborhoods $U_l(n) \supset I_l^*$ were introduced in Lemma 3.2. The induction hypothesis together with (7.4) and (7.5) imply that

$$\zeta_l \in U_l(n - k)$$

for $l \leq j$. Now repeatedly apply (7.1) and Lemma 3.2 to get

$$\begin{aligned} \frac{|\Delta x_{j+1}|}{|I_{j+1}^*|} &\leq \sum_{l=1}^{j+1} \left(\prod_{k=l}^j \frac{Df_*(\zeta_k)}{\frac{|I_{k+1}^*|}{|I_k^*|}} \right) \cdot K \frac{\rho^k}{\rho_0^{n-k}} + \left(\prod_{k=0}^j \frac{Df_*(\zeta_k)}{\frac{|I_{k+1}^*|}{|I_k^*|}} \right) \cdot \frac{|\Delta x_0|}{|I_0^*|} \\ &\leq (j+1)DK \frac{\rho^k}{\rho_0^{n-k}} + D \frac{|\Delta x_0|}{|I_0^*|} \\ &\leq DK \left(\frac{2}{\rho_0}\right)^{n-k} \cdot \rho^k + D \frac{|\Delta x_0|}{|I_0^*|}. \end{aligned}$$

This estimate finishes the induction step. $\square$

Now incorporate the estimates (7.4), (7.5) in the Claim and together with (7.3), Lemma 7.3 follows. $\square$

Proof of Proposition 7.2. Let $(1 - q_1) \cdot n \leq k \leq n$ and assume that the conditions of Lemma 7.3 are satisfied. Choose $B \in \mathcal{B}^n[k]$. Let $\mathbf{B} \in \mathcal{B}^{n-k}(F_k)$ be such that $B = \Psi_0^k(\mathbf{B})$, say $\mathbf{B} = \mathbf{B}_j$ with $0 < j < 2^{n-k}$ odd.

The pieces $\mathbf{B}_j^* \in \mathcal{B}^{n-k}(F_*)$, $0 < j < 2^{n-k}$ odd, are curves on the graph of f_* contained in $B_c^1(F_*)$, that is, they have a bounded slope. This bounded slope implies that

$$|I_j^*| \asymp |J_j^*|.$$

This bound and Lemma 7.3 imply that the Hausdorff distance between $\mathbf{B}_j$ and $\mathbf{B}_j^*$ is $O(\rho^{q_0 \cdot n} \cdot |I_j^*|)$. We get that $B_j = \Psi_0^k(\mathbf{B}_j)$ is regular, which proves Proposition 7.2(1).

Let $B \in \mathcal{B}^{n+1}[k]$, say $B = \Psi_0^k(\mathbf{B})$ with $\mathbf{B} \in \mathcal{B}^{n-k+1}(F_k)$ and $\mathbf{B} \subset \mathbf{B}_j \in \mathcal{B}^{n-k}(F_k)$, for some $0 < j < 2^{n-k}$. Recall that the scaling ratio of $B \in \mathcal{B}^{n+1}[k]$ is a measurement in vertical direction in the domain of F_k . The relative displacement of every point $z^* \in \mathcal{O}_{F_*}$ is estimated in Lemma 7.3. These bounds imply

$$|\sigma_{\mathbf{B}} - \sigma_{\mathbf{B}^*}| = O(\rho^{q_0 \cdot n}).$$

This finishes the proof of Proposition 7.2(2).

To control the thickness associated to $B \in \mathcal{B}^n[k]$ we have to restrict ourselves to $(1 - q_1) \cdot n \leq k \leq (1 - q_0) \cdot n$. The piece $\mathbf{B} \equiv \mathbf{B}_j$, which determines the relative thickness of $B = \Psi_0^k(\mathbf{B})$ has a Hausdorff distance $O(\rho^{q_0 \cdot n} \cdot |I_j^*|)$ to $\mathbf{B}_j^*$, Lemma 7.3.

This piece $\mathbf{B}_j^*$ is a curve in the graph of f_* contained in $B_c^1(F_*)$. This curve has a bounded slope. Hence, its relative thickness is proportional to its diameter, which is of the order $\sigma^{n-k} \leq \sigma^{q_0 \cdot n}$, see Lemmas 3.1. The control of the Hausdorff distance and the small relative thickness of $\mathbf{B}_j^*$ implies

$$\delta_{\mathbf{B}} = O(\rho^{q_0 \cdot n})$$

We finished the proof of Proposition 7.2(3). $\square$

7.3. Universal sticks created in the pushing-up regime.

Definition 7.2. Given $0 < q_0 < q_1$, the collection $\mathcal{P}_n(k; q_0, q_1)$ of (q_0, q_1) -controlled pieces consists of $B \in \mathcal{B}^n[k]$ with the following property. If $B^{(i)}$, $i = 0, 1, 2, \dots, t$, are the predecessors of $B = B^{(0)}$ with

$$k = k_0(B) < k_1(B) < k_2(B) < \dots < k_{t-1}(B) < k_t(B) < n.$$

then

- (1) $k_{i+1} \leq l(k_i)$, $i = 0, 1, 2, 3, \dots, t-1$,
- (2) there exists $0 \leq s \leq t$ such that $(1 - q_1) \cdot n \leq k_s(B) \leq (1 - q_0) \cdot n$, and
- (3) $k_{s-1}(B) \leq (1 - q_1) \cdot n$.

Remark 7.2. The definition of controlled pieces is a combinatorial definition. It does not depend on F but only on the average Jacobian b_F which is a topological invariant, [16]. If B is a (q_0, q_1) -controlled piece of F then the corresponding piece B^* is (q_0, q_1) -controlled piece of F_* .

The definition of controlled pieces implies

$$\bigcup_{k < l \leq l(k)} G_k(\mathcal{P}_n(l; q_0, q_1)) = \mathcal{P}_n(k; q_0, q_1). \quad (7.6)$$

Proposition 7.2 introduced the constants $\rho < 1$, and $q^* > 0$. The constants $\alpha^* > 0$ and $k^* > 0$ are the optimal choice given by the Propositions 4.1, 5.1 and 6.1. Now Proposition 4.1 and Proposition 7.2(1) imply

Lemma 7.5. *Let $\alpha < \alpha^*$. For every $q^* > q_1 > q_0 > 0$ there exists $n^* \geq 1$ such that every $B \in \mathcal{P}_n(k; q_0, q_1)$ and all its predecessors are regular when $n \geq n^*$ and $k \geq k^*$.*

Lemma 7.6. *Let $\alpha < \alpha^*$. For every $q^* > q_1 > q_0 > 0$ there exists $n^* \geq 1$ such that for every $\hat{B} \in \mathcal{P}_n(k; q_0, q_1)$ and $B \in \mathcal{B}^{n+1}[k]$ with $B \subset \hat{B}$*

$$\delta_{\hat{\mathbf{B}}} = O(\rho^{q_0 \cdot n})$$

and

$$|\sigma_{\mathbf{B}} - \sigma_{\hat{\mathbf{B}}}^*| = O(\rho^{q_0 \cdot n})$$

when $n \geq n^*$ and $k \geq k^*$.

Proof. Let us call the predecessors of $\hat{B}$ and B

$$B^{(i)} \subset \hat{B}^{(i)},$$

$i = 0, 1, 2, \dots, t$. Let $k_i = k_i(\hat{B}) = k_i(B)$ and δ_i the relative thickness of $\hat{\mathbf{B}}^{(i)}$, where $\hat{B}^{(i)} = \Psi_0^{k_i}(\hat{\mathbf{B}}^{(i)})$, and $\sigma_i = \sigma_{\mathbf{B}^{(i)}}$, the scaling number of $B^{(i)} \Psi_0^{k_i}(\mathbf{B}^{(i)})$, $i = 0, 1, 2, \dots, t$. Observe, the piece B might have one predecessor more than $\hat{B}$.

Apply Propositions 5.1 and 6.1. In particular,

$$\delta_{i-1} \leq \frac{1}{2}\delta_i + O(\sigma^{n-k_i}) \quad (7.7)$$

and

$$|\sigma_{i-1} - \sigma_i| = O(\delta_i + \sigma^{n-k_i}) \quad (7.8)$$

for $i = 1, 2, \dots, t$.

Iterating estimate (7.7) we obtain

$$\begin{aligned} \sum_{i=0}^s \delta_i &\leq 2\delta_s + O(\sigma^{n-k_s}) \\ &= O(\rho^{q_0 \cdot n}) + O(\sigma^{q_0 \cdot n}), \end{aligned} \quad (7.9)$$

where we used Proposition 7.2(3) and property (2) of Definition 7.2. We may assume $\sigma < \rho < 1$. The first estimate of the Lemma follows:

$$\delta_{\hat{\mathbf{B}}} = \delta_0 \leq \sum_{i=0}^s \delta_i = O(\rho^{q_0 \cdot n}).$$

To establish the second estimate of the Proposition, first observe that

$$\sigma_{\mathbf{B}^{(0)}} = \sigma_{\mathbf{B}^{(s)}} + \sum_{i=0}^{s-1} (\sigma_{\mathbf{B}^{(i)}} - \sigma_{\mathbf{B}^{(i+1)}}).$$

Hence, by using (7.8) and (7.9),

$$\begin{aligned} |\sigma_{B^{(0)}} - \sigma_{B^{(s)}}| &\leq \sum_{i=0}^{s-1} |\sigma_{B^{(i)}} - \sigma_{B^{(i+1)}}| \\ &= O\left(\sum_{i=1}^s (\delta_i + \sigma^{n-k_i})\right) \\ &= O(\rho^{q_0 \cdot n} + \sigma^{n-k_s}) = O(\rho^{q_0 \cdot n}). \end{aligned}$$

If $B \in \mathcal{P}_n(k, q_0, q_1)$ and B^* is the corresponding piece of F_* , then B^* is also controlled. Namely, each $l(k) = \infty$ because $b_{F_*} = 0$. Hence, we have the same estimate for the proper scaling

$$|\sigma_{\mathbf{B}^{(0)}}^* - \sigma_{\mathbf{B}^{(s)}}^*| = O(\rho^{q_0 \cdot n}).$$

This finishes the proof. Namely, $B^{(s)} \in \mathcal{B}^{n+1}[k_s]$ with $(1 - q_1) \cdot n \leq k_s \leq n$ and we can apply Proposition 7.2(2),

$$\begin{aligned} |\sigma_{\mathbf{B}} - \sigma_{\mathbf{B}}^*| &= |\sigma_{\mathbf{B}^{(0)}} - \sigma_{\mathbf{B}^{(0)}}^*| \\ &\leq |\sigma_{\mathbf{B}^{(0)}} - \sigma_{\mathbf{B}^{(s)}}| + |\sigma_{\mathbf{B}^{(s)}} - \sigma_{\mathbf{B}^{(s)}}^*| + |\sigma_{\mathbf{B}^{(s)}}^* - \sigma_{\mathbf{B}^{(0)}}^*| \\ &= O(\rho^{q_0 \cdot n}). \end{aligned}$$

□

The measurements of the pieces, such as scaling and thickness, are geometric quantities observed when viewing a piece from its scale, they are geometric measurements of $\mathbf{B}$ and not B itself. The next Proposition states that the actual pieces B inherit exponentially small estimates for their precision. The Proposition is also a preparation for the brute-force regime which concerns iteration of the original map.

Proposition 7.7. *Let $\alpha < \alpha^*$. For every $q^* > q_1 > q_0 > 0$ there exists $n^* \geq 1$ such that*

$$\mathcal{P}_n(k; q_0, q_1) \subset \mathcal{S}_n(O(\rho^{q_0 \cdot n}))$$

when $n \geq n^*$ and $k \geq k^*$.

The estimates in the proof of this Proposition are like the estimates used to prove the Propositions 4.1, 5.1, and 6.1.

Proof. Let $\hat{B} \in \mathcal{P}_n(k; q_0, q_1)$ and $B \in \mathcal{B}^{n+1}[k]$ with $B \subset \hat{B}$. Let $\mathbf{B}$ and $\hat{\mathbf{B}}$ be such that $B = \Psi_0^k(\mathbf{B})$ and $\hat{B} = \Psi_0^k(\hat{\mathbf{B}})$. The horizontal and vertical size of the smallest rectangle which contains $\hat{\mathbf{B}}$ are $\mathbf{h}, \mathbf{v} > 0$. Let $\delta > 0$ be the relative thickness of $\hat{\mathbf{B}}$, the absolute thickness of $\hat{\mathbf{B}}$ is $\mathbf{w} = \delta \cdot \mathbf{h}$. From Lemma 3.1 we get

$$\mathbf{h}, \mathbf{v} = O(\sigma^{n-k}).$$

Moreover, the regularity of $\hat{B}$ gives

$$\mathbf{h} \asymp \mathbf{v}.$$

The situation allows to apply Lemma 7.6:

$$\delta = O(\rho^{q_0 \cdot n}) \quad \text{and} \quad |\sigma_{\mathbf{B}} - \sigma_{\mathbf{B}}^*| = O(\rho^{q_0 \cdot n}). \quad (7.10)$$

We have to show that $\hat{B} \cap \mathcal{O}_F = \Psi_0^k(\hat{\mathbf{B}} \cap \mathcal{O}_{F_k})$ is contained in a $O(\rho^{q_0 \cdot n})$ -stick. As before we will decompose Ψ_0^k into its diffeomorphic part $(\text{id} + \mathbf{S}_0^k)$ and its affine part. Let $h_{\text{diff}}, v_{\text{diff}} > 0$ be the horizontal and vertical size of the smallest rectangle containing the image of $\hat{\mathbf{B}}$ under $(\text{id} + \mathbf{S}_0^k)$ and $w_{\text{diff}} > 0$ the absolute thickness of its stick and $\sigma_{\text{diff}} > 0$ the scaling factor of the image of $\mathbf{B}$ under the same diffeomorphism. Then we have

$$\begin{aligned} \sigma_{\text{diff}} &= \sigma_{\mathbf{B}}, \\ v_{\text{diff}} &= \mathbf{v} \end{aligned}$$

and, by recalling (5.1),

$$\begin{aligned} w_{\text{diff}} &= O(\mathbf{w} + \sigma^{n-k} \cdot \mathbf{h}), \\ h_{\text{diff}} &\asymp \mathbf{h}. \end{aligned}$$

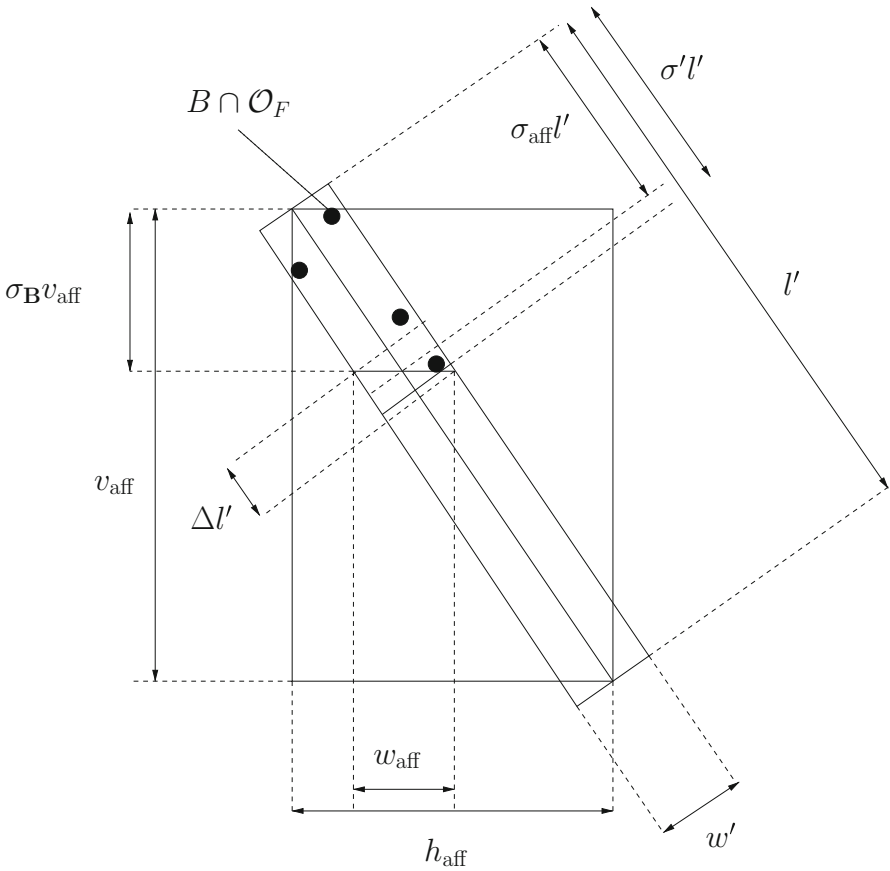


Fig. 10. Illustration for the proof of Proposition 7.7

The last two estimates rely on $\mathbf{v} \asymp \mathbf{h}$. The term $\mathbf{h} \cdot \sigma^{n-k}$ reflects the distortion of $(\text{id} + \mathbf{S}_0^k)$ on $\hat{\mathbf{B}}$ determined by the diameter of $\hat{\mathbf{B}}$ which is of the order σ^{n-k} . The next step is to apply the affine part of Ψ_0^k . Denote the measurements after this step by $h_{\text{aff}}, v_{\text{aff}}, w_{\text{aff}}, \sigma_{\text{aff}} > 0$ resp. Equation (3.6) and Lemma 3.8 yield

$$\begin{aligned} w_{\text{aff}} &\asymp \sigma^{2k} w_{\text{diff}}, \\ \sigma_{\text{aff}} &= \sigma_{\text{diff}} = \sigma_{\mathbf{B}}, \\ h_{\text{aff}} &\asymp \sigma^{2k} h_{\text{diff}} + \sigma^k v_{\text{diff}}, \end{aligned} \quad (7.11)$$

Use the above estimates in the following

$$\frac{w_{\text{aff}}}{h_{\text{aff}}} = O\left(\frac{\sigma^{2k} \cdot [\mathbf{w} + \sigma^{n-k} \cdot \mathbf{h}]}{\sigma^{2k} \cdot \mathbf{h} + \sigma^k \cdot \mathbf{v}}\right) = O(\sigma^k \cdot \delta + \sigma^n) = O(\rho^{q_0 \cdot n}). \quad (7.12)$$

Consider the smallest conformal image of a rectangle aligned along the diagonal of the rectangle containing $\hat{B} = \Psi_0^k(\hat{\mathbf{B}})$, see Fig. 10. The precision of $\hat{B}$ will be better than the precision based on the measurements of this approximation of the stick. Let $l' > 0$

be the length, $w' > 0$ be the absolute thickness and $\sigma' > 0$ be the scaling factor of $B \subset \hat{B}$ within this rectangle. Then

$$l' = \sqrt{h_{\text{aff}}^2 + v_{\text{aff}}^2}, \quad (7.13)$$

and

$$w' \leq w_{\text{aff}}. \quad (7.14)$$

First we will estimate the precision of σ' . Let γ be the angle between the diagonal of the rectangle and the horizontal. Observe,

$$\cos \gamma = \frac{h_{\text{aff}}}{\sqrt{h_{\text{aff}}^2 + v_{\text{aff}}^2}},$$

see Figure 10. The projection $\Delta l'$ of the horizontal interval of length w_{aff} onto the diagonal has length

$$\Delta l' = w_{\text{aff}} \cdot \cos \gamma.$$

Observe,

$$|\sigma' \cdot l' - \sigma_{\text{aff}} \cdot l'| \leq \Delta l' = w_{\text{aff}} \cdot \frac{h_{\text{aff}}}{\sqrt{h_{\text{aff}}^2 + v_{\text{aff}}^2}}.$$

Then, by using (7.12) and (7.13),

$$|\sigma' - \sigma_{\text{aff}}| \leq \frac{w_{\text{aff}}}{h_{\text{aff}}} \cdot \frac{h_{\text{aff}}^2}{h_{\text{aff}}^2 + v_{\text{aff}}^2} \leq \frac{w_{\text{aff}}}{h_{\text{aff}}} = O(\rho^{q_0 \cdot n}). \quad (7.15)$$

Use (7.10), (7.11), and (7.15) to estimate the precision of σ'

$$|\sigma' - \sigma_{\mathbf{B}}^*| \leq |\sigma' - \sigma_{\text{aff}}| + |\sigma_{\text{aff}} - \sigma_{\mathbf{B}}^*| = O(\rho^{q_0 \cdot n}). \quad (7.16)$$

The estimate (7.14) says that the height of the stick containing $\hat{B}$ is at most w_{aff} . The relative height is estimated by

$$\frac{w'}{l'} \leq \frac{w_{\text{aff}}}{\sqrt{h_{\text{aff}}^2 + v_{\text{aff}}^2}} \leq \frac{w_{\text{aff}}}{h_{\text{aff}}} = O(\rho^{q_0 \cdot n}), \quad (7.17)$$

where we used (7.12) and (7.13). The estimates (7.16) and (7.17) confirm that $\hat{B} \in \mathcal{S}_n(\rho^{q_0 \cdot n})$, which finishes the proof of the Proposition. $\square$

7.4. Universal sticks created in the brute-force regime. The one-dimensional regime and pushing-up regime do not create enough universal pieces. On the highest level the universal pieces are iterated around with the original map.

Proposition 7.8. *There exists $\epsilon^* > 0$, and $q^* > 0$ such that the following holds. Let $\epsilon < \epsilon^*$, and $0 < q_0 < q_1 < q^*$ then there exists $n^* \geq 1$ such that if for $0 \leq j < 2^{(1-q_1) \cdot n}$*

$$F^j(B) \in \mathcal{S}_n(\epsilon),$$

with $B \in \mathcal{B}^n[k]$, $(1 - q_1) \cdot n \leq k \leq (1 - q_0) \cdot n$, and $n \geq n^$, then*

$$F^{j+1}(B) \in \mathcal{S}_n(O(\epsilon + \rho^{q_0 \cdot n})).$$

Proof. Choose $\hat{B} \in \mathcal{B}^n[k]$ with $(1 - q_1) \cdot n \leq k \leq (1 - q_0) \cdot n$ and $B \in \mathcal{B}^{n+1}$ with $B \subset \hat{B}$. The iterates under the original map are denoted by $B_j = F^j(B)$ and $\hat{B}_j = F^j(\hat{B})$, $j \leq 2^{(1-q_1) \cdot n}$. Assume that for some $j \leq 2^{(1-q_1) \cdot n}$

$$\hat{B}_j \in \mathcal{S}_n(\epsilon).$$

The piece $\hat{B}_j$ is contained in an ϵ -stick. Say $\hat{B}_j \cap \mathcal{O}_F$ is contained in a rectangle of length $l > 0$ and height $w \leq \epsilon l$. The smaller rectangle which contains $B_j \cap \mathcal{O}_F$ has length $\sigma_j l$, where $\sigma_j = \sigma_{B_j}$ and $|\sigma_j - \sigma_{\mathbf{B}_j}^*| \leq \epsilon$. Notice that we have to estimate the scaling factor σ_{B_j} and not $\sigma_{\mathbf{B}_j}$, compare remark 6.1.

Apply F to the rectangle which contains $\hat{B}_j$. The stick which contains $\hat{B}_{j+1}$ has length $l' > 0$ and height $w' > 0$. The relevant scaling factor of B_{j+1} is $\sigma_{j+1} = \sigma_{B_{j+1}}$.

Choose, $M, m > 0$ such that

$$m|v| \leq |DF(x, y)v| \leq M|v|.$$

This is possible because F is a diffeomorphism onto its image. However, $m = O(b)$. Let $K > 0$ be the maximum norm of the Hessian of F . The diameter of $\hat{B}_j \cap \mathcal{O}_F$, which is proportional to l , is of the order σ^n , see Lemma 3.1. We can estimate the sizes l' , w' and σ' by applying the derivative of F and correcting for distortion which is bounded by Kl^2 . Let D be the absolute value of the directional derivative of F in the direction of the rectangle containing $\hat{B}_j$, measured in a corner of the rectangle. Then

$$l' \geq Dl - 2Kl^2 - 2Mw,$$

$$w' \leq Mw + 2Kl^2,$$

Observe,

$$|\sigma_{j+1} \cdot l' - D \cdot \sigma_j \cdot l| \leq 2Mw + 2Kl^2.$$

Let us first estimate the relative height of the stick of $\hat{B}_{j+1}$. Use $w \leq \epsilon l$,

$$\begin{aligned} \frac{w'}{l'} &\leq \frac{M\epsilon l + 2Kl^2}{ml - 2Kl^2 - 2M\epsilon l} \\ &\leq \frac{M}{m - 2Kl - 2M\epsilon} \cdot \epsilon + 2 \frac{K}{m - 2Kl - 2M\epsilon} \cdot l \\ &= O(\epsilon + \sigma^n) = O(\epsilon + \rho^{q_0 \cdot n}), \end{aligned} \tag{7.18}$$

when $\epsilon < \epsilon^*$, $q_0 < q_1^*$ small enough, and $n \geq n^*$ large enough. Similarly,

$$|\sigma_{j+1} - \sigma_j| = O(\epsilon + \rho^{q_0 \cdot n}). \quad (7.19)$$

Use remark 6.3 and apply Proposition 3.3 to get

$$|\sigma_{\mathbf{B}_s}^* - \sigma_{\mathbf{B}}^*| = O(\rho^{q_0 \cdot n}), \quad (7.20)$$

with $0 \leq s < 2^{(1-q_1) \cdot n}$.

We need to estimate the scaling factor σ_{j+1} of B_{j+1} . Use (7.19) and (7.20) and the notation $\sigma_j^* = \sigma_{\mathbf{B}_j}^*$. Then

$$\begin{aligned} |\sigma_{j+1} - \sigma_{j+1}^*| &\leq |\sigma_{j+1} - \sigma_j| + |\sigma_j - \sigma_j^*| + |\sigma_j^* - \sigma_{j+1}^*| \\ &\leq O(\epsilon + \rho^{q_0 \cdot n}) + \epsilon + O(\rho^{q_0 \cdot n}) \\ &= O(\epsilon + \rho^{q_0 \cdot n}), \end{aligned} \quad (7.21)$$

for $\epsilon \leq \epsilon^*$, $0 < q_0 < q^*$ small enough and $n \geq n^*$ large enough. The estimates (7.18) and (7.21) together finish the proof. $\square$

8. Probabilistic Universality

In this section we are going to estimate the measure of the pieces created in the three regimes, see Proposition 8.6. Let $\alpha = \alpha^*$, $\epsilon^* > 0$, and $0 < q_1^* < 1/3$ small and $k^* \geq 1$ large enough to allow the use of the Propositions 7.7, and 7.8.

For each $n \geq 1$, let $\kappa_0(n) \asymp \ln n$ be the smallest integer such that

$$l(\kappa_0(n)) \equiv 2^{\kappa_0(n)} \cdot \frac{\ln b}{\ln \sigma} - \frac{\ln \alpha}{\ln \sigma} + \kappa_0(n) \geq n.$$

For $n \geq 1$ large enough we have

$$\kappa_0(n) \leq \frac{\ln n}{\ln 2}. \quad (8.1)$$

Lemma 8.1. *Given $q_0 < q_1$. There exists $n^* \geq 1$ such that for $n \geq n^*$ and $\kappa_0(n) \leq k < (1 - q_0) \cdot n$,*

$$\mu(\mathcal{P}_n(k; q_0, q_1)) \geq [1 - \frac{1}{2^{(q_1 - q_0) \cdot n + 1}}] \cdot \mu(E^k).$$

Proof. Let $\beta_n(k; q_0, q_1) = \mu(E^k \setminus \mathcal{P}_n(k; q_0, q_1))$ be the measure of the uncontrolled pieces. The construction implies immediately

$$\beta_n(k; q_0, q_1) = \mu(E^k), \quad (1 - q_0) \cdot n < k \leq n, \quad (8.2)$$

and

$$\beta_n(k; q_0, q_1) = 0, \quad (1 - q_1) \cdot n \leq k \leq (1 - q_0) \cdot n, \quad (8.3)$$

every piece in the one-dimensional regime is controlled. The Lemma holds for $(1 - q_1) \cdot n \leq k \leq (1 - q_0) \cdot n$. This implies that the fraction of the uncontrolled part in $\cup_{l \geq (1-q_1) \cdot n} E^l$ is

$$\frac{\sum_{l=(1-q_1) \cdot n}^n \beta_n(l; q_0, q_1)}{\mu(B^{(1-q_1) \cdot n})} \leq \frac{1}{2^{(q_1 - q_0) \cdot n + 1}}. \quad (8.4)$$

Observe,

$$\begin{aligned} l((1 - q_1) \cdot n - 1) &= 2^{(1-q_1) \cdot n - 1} \cdot \frac{\ln b}{\ln \sigma} - \frac{\ln \alpha}{\ln \sigma} + (1 - q_1) \cdot n - 1 \\ &\gtrsim 2^{(1-q_1) \cdot n - 1} \geq n, \end{aligned}$$

holds when $n \geq 1$ is large enough. All pieces in $\mathcal{B}^n[k]$, with $k \geq (1 - q_1) \cdot n$ are not too deep for level $(1 - q_1) \cdot n - 1$. Hence, equation (7.6) reduces to

$$\mathcal{P}_n((1 - q_1) \cdot n - 1; q_0, q_1) = \bigcup_{(1-q_1) \cdot n \leq l \leq n} G_{(1-q_1) \cdot n - 1}(\mathcal{P}_n(l; q_0, q_1)).$$

Hence, using (8.4),

$$\begin{aligned} \beta_n((1 - q_1) \cdot n - 1; q_0, q_1) &= \sum_{l=(1-q_1) \cdot n}^n \beta_n(l; q_0, q_1) \\ &\leq \frac{1}{2^{(q_1 - q_0) \cdot n + 1}} \cdot \mu(B^{(1-q_1) \cdot n}) \\ &= \frac{1}{2^{(q_1 - q_0) \cdot n + 1}} \cdot \mu(E^{(1-q_1) \cdot n - 1}). \end{aligned}$$

Now we finish the proof by induction. The Lemma is proved for $k = (1 - q_1) \cdot n - 1$. Assume the Lemma holds from $(1 - q_1) \cdot n - 1$ down to $k + 1 \leq (1 - q_1) \cdot n - 1$. Because $k \geq \kappa_0(n)$ we have $l(k) \geq n$. Hence, again by using (7.6), (8.2), (8.3), and $\mu(E^l) = \frac{1}{2^{l+1}}$, $l \geq 0$, we get

$$\begin{aligned} \mu(\mathcal{P}_n(k; q_0, q_1)) &= \mu\left(\bigcup_{l=k+1}^n G_k(\mathcal{P}_n(l; q_0, q_1))\right) \\ &= \sum_{l=k+1}^{(1-q_1) \cdot n - 1} \mu(\mathcal{P}_n(l; q_0, q_1)) + \sum_{l=(1-q_1) \cdot n}^{(1-q_0) \cdot n} \mu(E^l) \\ &\geq \left(1 - \frac{1}{2^{(q_1 - q_0) \cdot n + 1}}\right) \cdot \left[\sum_{l=k+1}^{(1-q_1) \cdot n - 1} \mu(E^l) + \frac{1}{2^{(1-q_1) \cdot n}}\right] \\ &= \left(1 - \frac{1}{2^{(q_1 - q_0) \cdot n + 1}}\right) \cdot \left[\sum_{l=k+1}^{(1-q_1) \cdot n - 1} \mu(E^l) + \sum_{l=(1-q_1) \cdot n}^{\infty} \mu(E^l)\right] \\ &= \left(1 - \frac{1}{2^{(q_1 - q_0) \cdot n + 1}}\right) \cdot \mu(E^k). \end{aligned}$$

□

Proposition 8.2. *Given $q_0 < q_1 < q_1^*$. There exists $n^* \geq 1$ such that for $n \geq n^*$ and $k \leq (1 - q_0) \cdot n$*

$$\mu(\mathcal{P}_n(k; q_0, q_1)) \geq \left[1 - \frac{1}{2^{(q_1 - q_0) \cdot n + 1}} - 2^{\frac{\ln \alpha \sigma}{\ln \sigma}} \sum_{l=k}^{\infty} 2^l (b^\gamma)^{2^l}\right] \cdot \mu(E^k),$$

where $\gamma = -\frac{\ln 2}{\ln \sigma} \in (0, 1)$.

Proof. According to Lemma 8.1, the Proposition holds for $\kappa_0(n) \leq k \leq (1 - q_0) \cdot n$. The proof for the lower values of $k < \kappa_0(n)$ is by induction. Assume by induction

$$\beta_n(k; q_0, q_1) \leq \left[\frac{1}{2^{(q_1 - q_0) \cdot n + 1}} + 2^{\frac{\ln \alpha \sigma}{\ln \sigma}} \sum_{l=k}^{\kappa_0(n)-1} 2^l (b^\gamma)^{2^l} \right] \cdot \mu(E^k),$$

which holds for $k = \kappa_0(n)$. Suppose it holds from $\kappa_0(n)$ down to $k+1 \leq \kappa_0(n)$. Observe,

$$\frac{1}{2^{l(k)}} = 2^{\frac{\ln \alpha}{\ln \sigma}} \cdot 2^{-\lfloor \frac{k}{2^k} + \frac{\ln b}{\ln \sigma} \rfloor \cdot 2^k} \leq 2^{\frac{\ln \alpha}{\ln \sigma}} \cdot 2^{-\frac{\ln b}{\ln \sigma} \cdot 2^k}.$$

Hence,

$$\frac{1}{2^{l(k)}} \leq 2^{\frac{\ln \alpha}{\ln \sigma}} \cdot (b^\gamma)^{2^k}. \quad (8.5)$$

Use (8.1) and observe,

$$\begin{aligned} l_{\kappa_0(n)-1} &= 2^{\kappa_0(n)-1} \cdot \frac{\ln b}{\ln \sigma} - \frac{\ln \alpha}{\ln \sigma} + \kappa_0(n) - 1 \\ &= \frac{1}{2} \left(n + \frac{\ln \alpha}{\ln \sigma} - \kappa_0(n) \right) - \frac{\ln \alpha}{\ln \sigma} + \kappa_0(n) - 1 \\ &\leq \frac{1}{2} n \left(1 + \frac{\kappa_0(n)}{n} \right) + O(1) \\ &\leq \frac{1}{2} n \left(1 + \frac{\ln n}{n \ln 2} \right) + O(1) \\ &< (1 - q_1) \cdot n, \end{aligned}$$

holds when $n \geq n^*$ large enough because $q_1^* < \frac{1}{3}$. Hence, for $n \geq 1$ large enough, we have

$$l(k) \leq l(\kappa_0(n) - 1) < (1 - q_1) \cdot n. \quad (8.6)$$

Use (7.6), the induction hypothesis, (8.5), and (8.6) in the following estimates.

$$\begin{aligned} \beta_n(k; q_0, q_1) &\leq \sum_{l=k+1}^{l(k)} \beta_n(l; q_0, q_1) + \mu(B^{l(k)+1}) \\ &\leq \left[\frac{1}{2^{(q_1 - q_0) \cdot n + 1}} + 2^{\frac{\ln \alpha \sigma}{\ln \sigma}} \sum_{l=k+1}^{\kappa_0(n)-1} 2^l (b^\gamma)^{2^l} \right] \cdot \sum_{l=k+1}^{l(k)} \mu(E^l) + \frac{1}{2^{l(k)}} \\ &\leq \left[\frac{1}{2^{(q_1 - q_0) \cdot n + 1}} + 2^{\frac{\ln \alpha \sigma}{\ln \sigma}} \sum_{l=k+1}^{\kappa_0(n)-1} 2^l (b^\gamma)^{2^l} \right] \cdot \mu(E^k) + 2^{\frac{\ln \alpha}{\ln \sigma}} (b^\gamma)^{2^k} \\ &= \left[\frac{1}{2^{(q_1 - q_0) \cdot n + 1}} + 2^{\frac{\ln \alpha \sigma}{\ln \sigma}} \sum_{l=k}^{\kappa_0(n)-1} 2^l (b^\gamma)^{2^l} \right] \cdot \mu(E^k), \end{aligned}$$

where the last equality uses $\mu(E^k) = \frac{1}{2^{k+1}}$. $\square$

For each $K > 0$ and $\theta < 1$, let $\kappa(n)$ be the largest integer such that

$$2^{\kappa(n)} \leq K n \ln 1/\theta.$$

Lemma 8.3. *There exists $K > 0$ such that for every $\theta < 1$ there exists $n^* \geq 1$ such that $\kappa(n) \geq k^*$ for $n \geq n^*$ and*

$$2^{\frac{\ln \alpha \sigma}{\ln \sigma}} \sum_{l=\kappa(n)}^{\infty} 2^l (b^\gamma)^{2^l} \leq \frac{1}{3} \theta^n.$$

Proof. Observe,

$$\sum_{l=\kappa(n)}^{\infty} 2^l (b^\gamma)^{2^l} = O(2^{\kappa(n)} (b^\gamma)^{2^{\kappa(n)}}).$$

To achieve the property of the Lemma it suffices to satisfy

$$\ln 2^{\kappa(n)} + 2^{\kappa(n)} \ln b^\gamma + O(1) \leq n \ln \theta.$$

In turn, this holds when

$$n \ln 1/\theta \cdot [\frac{1}{2} K \ln b^\gamma + 1] + O(1) \leq 0.$$

This holds for large $n \geq 1$ when $K > 0$ is chosen large enough. $\square$

In the sequel we will fix $K > 0$ according to the previous Lemma. For each $Q > 0$ and $\theta < 1$, define q_0 by

$$q_0 = Q \ln 1/\theta.$$

and

$$q_1 = [Q + \frac{3}{2 \ln 2}] \cdot \ln 1/\theta.$$

Lemma 8.4. *For every $\theta < 1$ there exists $n^* \geq 1$ such that for $Q > 0$ and $n \geq n^*$*

$$\frac{1}{2^{(q_1 - q_0) \cdot n + 1}} \leq \frac{1}{3} \theta^n.$$

The brute-force regime consists of iterates of $\bigcup_{l=\kappa(n)}^{(1-q_0) \cdot n} \mathcal{P}_n(l; q_0, q_1)$ up to just one step before the moment of return to $B^{\kappa(n)} \equiv \bigcup_{l=\kappa(n)}^{\infty} E^l$. The return uses exactly $2^{\kappa(n)}$ steps. Thus we obtain for each choice $Q > 0$ and $\theta < 1$, the collection

$$\mathcal{P}_n = \bigcup_{j=0}^{2^{\kappa(n)} - 1} F^j \left(\bigcup_{l=\kappa(n)}^{(1-q_0) \cdot n} \mathcal{P}_n(l; q_0, q_1) \right) \quad (8.7)$$

Proposition 8.5. *There exist $Q > 0$ and $\theta^* < 1$ such that the following holds. For $\theta^* \leq \theta < 1$ there exists $n^* \geq 1$ such that for $n \geq n^*$*

$$\mathcal{P}_n \subset \mathcal{S}_n(\theta^n).$$

Proof. Take $B \in \bigcup_{l=\kappa(n)}^{(1-q_0) \cdot n} \mathcal{P}_n(l; q_0, q_1)$. According to Proposition 7.7 there exists $C > 0$ such that

$$B \in \mathcal{S}_n(C\rho^{q_0 \cdot n}), \quad (8.8)$$

when $\theta < 1$ close enough to 1 and $n \geq 1$ large enough (Recall that q_0 depends on θ). Now consider an image $F^j(B)$ with $j \leq 2^{\kappa(n)} - 1 < 2^{(1-q_1) \cdot n}$. Denote its precision by ϵ_j . This is a piece in the brute-force regime. If $\theta < 1$ close enough to 1 and $n \geq 1$ large enough we can apply Proposition 7.8: there exists $r > 1$ such that if $\epsilon_j \leq \epsilon^*$ then

$$\epsilon_{j+1} \leq r \cdot (\epsilon_j + \rho^{q_0 \cdot n}). \quad (8.9)$$

Choose $Q > 0$ large enough such that

$$Q \ln \rho + K \ln r + \frac{3}{2} \leq 0.$$

This choice implies

$$\rho^{q_0 \cdot n} \cdot r^{2^{\kappa(n)}} \leq (\theta^{\frac{3}{2}})^n. \quad (8.10)$$

Now we can repeatedly apply (8.9): for $n \geq 1$ large enough and $0 \leq j < 2^{\kappa(n)}$

$$\begin{aligned} \epsilon_j &\leq C\rho^{q_0 \cdot n} \cdot r^j + \rho^{q_0 \cdot n} \cdot \sum_{i=0}^{j-1} r^{j-i} \\ &\leq (C + \frac{r}{r-1}) \cdot \rho^{q_0 \cdot n} \cdot r^{2^{\kappa(n)}} \leq \theta^n \leq \epsilon^*. \end{aligned}$$

Every piece in $\mathcal{P}_n$ is θ^n -universal. $\square$

In the sequel we will fixed $Q > 0$ according to the previous Proposition.

Proposition 8.6. *There exists $\theta^* < 1$ such that the following holds. For $\theta^* \leq \theta < 1$ there exists $n^* \geq 1$ such that for $n \geq n^*$*

$$\mu(\mathcal{P}_n) \geq 1 - \theta^n.$$

Proof. For $\theta < 1$ close enough to 1 we have

$$\frac{1}{2^{\frac{1}{2}-Q \ln 1/\theta}} \leq \theta.$$

Hence, for $n \geq 1$ large enough

$$\frac{Kn \ln 1/\theta}{2^{(1-Q \ln 1/\theta) \cdot n+1}} \leq \frac{1}{3} \cdot \theta^n. \quad (8.11)$$

For $\theta < 1$ close enough to 1, and $n \geq 1$ large enough we can apply Proposition 8.2, Lemmas 8.3, 8.4, and (8.11) to obtain

$$\begin{aligned}\mu(\mathcal{P}_n) &= 2^{\kappa(n)} \cdot \mu\left(\bigcup_{l=\kappa(n)}^{(1-q_0)\cdot n} \mathcal{P}_n(l; q_0, q_1)\right) \\ &\geq 2^{\kappa(n)} \cdot \left(1 - \frac{2}{3}\theta^n\right) \cdot \sum_{l=\kappa(n)}^{(1-q_0)\cdot n} \mu(E^l) \\ &= \left(1 - \frac{2}{3}\theta^n\right) \cdot \left(1 - \frac{2^{\kappa(n)}}{2^{(1-q_0)\cdot n+1}}\right) \\ &\geq \left(1 - \frac{2}{3}\theta^n\right) \cdot \left(1 - \frac{Kn \ln 1/\theta}{2^{(1-Q \ln 1/\theta)\cdot n+1}}\right) \\ &\geq 1 - \theta^n.\end{aligned}$$

□

The Propositions 8.6 and 8.5 confirm probabilistic universality, Theorem 7.1.

9. Recovery

The pieces in $\mathcal{B}^n$ which are contained in θ^n -sticks can be determined by pure combinatorial methods. In [7], it has been shown that there are pieces which are not contained in θ^n -sticks. Probabilistic universality says that these *bad* spots will be filled on deeper levels with pieces contained in sticks with exponential precision. Although a bad spot might have very non-universal geometry, the dynamics on deeper levels forgets this bad geometry and recovers, filling the bad spot with mostly universal pieces on deeper levels. This recovery process has a combinatorial description.

A piece $B \in \mathcal{B}^n$ has an associated word $\omega = w_1 w_2 \dots w_n$, with letters $w_k \in \{c, v\}$, such that

$$B = \text{Im } \psi_{w_1}^1 \circ \psi_{w_2}^2 \circ \dots \circ \psi_{w_n}^n$$

where ψ_v^k is the non-affine rescaling used to renormalize $R^k F$, and to obtain $R^{k+1} F$ and $\psi_c^k = R^k F \circ \psi_v^k$. If $B_1, B_2 \in \mathcal{B}^{n+1}$ are the two pieces contained in B then the associated words for B_1 and B_2 are wc and wv . This discussion defines a homeomorphism

$$w : \mathcal{O}_F \rightarrow \{c, v\}^{\mathbb{N}}.$$

The relation between the $k_i(B)$, $i = 0, 1, 2, \dots, t$, which define the predecessors of $B \in \mathcal{B}^n$ and the word $\omega = w_1 w_2 \dots w_n$ is as follows. If $i \in \{k_0(B), k_1(B), \dots, k_t(B)\}$ then $w_i = c$, otherwise $w_i = v$.

In the previous section we constructed the collection $\mathcal{P}_n \subset \mathcal{S}_n(\theta^n)$, see (8.7). The word $\omega = w_1 w_2 \dots w_n$ of a piece $B \in \mathcal{P}_n$ is characterized by

- (1) If $k \geq \kappa(n)$ and $w_k = c$ then there exists $k < i \leq l(k)$ with $w_i = c$.
- (2) There exists $n - q_1 \cdot n \leq k \leq n - q_0 \cdot n$ with $w_k = c$.

Remark 9.1. Recall, q_0, q_1 , and the function $l(k)$, depend only on the average Jacobian, which is a topological invariant, see [16]. The characterization of the pieces in $\mathcal{P}_n$ is purely topological.

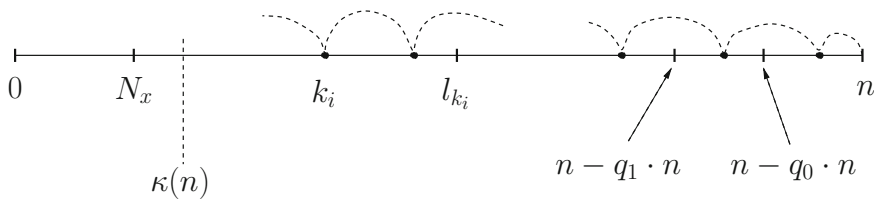


Fig. 11. Controlled points

Definition 9.1. A point $x \in \mathcal{O}_F$ is *eventually controlled* if there exists $N_x \geq 1$ such that for all $n \geq N_x$ there exists $n - q_1 \cdot n \leq k \leq n - q_0 \cdot n$ with

$$w_k = c,$$

where $w(x) = w_1 w_2 w_3 \dots$. The collection of eventually controlled points is denoted by $C_F \subset \mathcal{O}_F$.

Lemma 9.1. The set of eventually controlled points satisfies $\mu(C_F) = 1$ and

$$C_F = \bigcup_{N \geq 1} \bigcap_{n \geq N} \mathcal{P}_n.$$

Proof. There exists $k^* \geq 1$ such that $(1 - q_1) \cdot l(k) > k$ for $k \geq k^*$. Let $x \in C_F$. Choose $n \geq 1$ large enough such that $n \geq \kappa(n) \geq N_x$ and $\kappa(n) \geq k^*$. The piece $B_n(x) \in \mathcal{B}^n$ contains x . Then $B_n(x)$ satisfies property (2).

Choose $k \geq \kappa(n)$. Then $l(k) > (1 - q_1) \cdot l(k) > k \geq \kappa(n) \geq N_x$. Hence, there exists $w_i = c$ with $(1 - q_1) \cdot l(k) \leq i \leq (1 - q_0) \cdot l(k)$. Now, $i \geq (1 - q_1) \cdot l(k) > k$. Moreover, $i \leq (1 - q_0) \cdot l(k) < l(k)$. The piece $B_n(x)$ satisfies property (1). We proved,

$$x \in \bigcap_{\kappa(n) \geq \max\{N_x, k^*\}} \mathcal{P}_n. \quad (9.1)$$

Choose $x \in \bigcap_{n \geq N} \mathcal{P}_n$. Then property (2) implies that for every $n \geq N$ there exists $n - q_1 \cdot n \leq k \leq n - q_0 \cdot n$ with

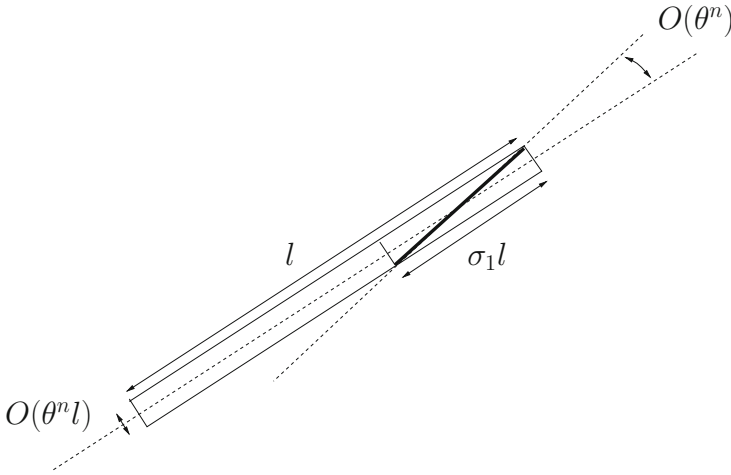
$$w_k = c.$$

We proved that $\bigcap_{n \geq N} \mathcal{P}_n \subset C_F$, for $N \geq 1$. The statement on the measure of C_F follows from Proposition 8.6. This finishes the proof of Lemma 9.1 $\square$

The recovery process can be described by using Proposition 8.5 and (9.1)

Proposition 9.2. If $x \in \mathcal{O}_F$ is controlled then and $\kappa(n) \geq N_x$ then $B_n(x) \in \mathcal{S}_n(\theta^n)$.

Remark 9.2. Given a conjugation $h : \mathcal{O}_{F_1} \rightarrow \mathcal{O}_{F_2}$ then $b_{F_1} = b_{F_2}$, see [16], and $h(C_{F_1}) = C_{F_2}$. The set of controlled points is a topological invariant (Fig. 11).

Fig. 12. θ^n -stick

10. Probabilistic Rigidity

The geometry of large parts of $\mathcal{O}_F$ resemble that of the geometry of $\mathcal{O}_{F_*}$, see Theorem 7.1, probabilistic universality. The large parts are

$$X_N = \bigcap_{k \geq N} \mathcal{S}_k(\theta^k), \quad (10.1)$$

where $\theta < 1$ is given by Theorem 7.1, with

$$\mu(X_N) \geq 1 - O(\theta^N).$$

Let

$$X = \bigcup_{N \geq 1} X_N$$

and note $\mu(X) = 1$.

As a consequence of a result from [7] we know that there is no continuous line field on $\mathcal{O}_F$ consisting of tangent lines to $\mathcal{O}_F$. However, the first step towards describing the geometry of $\mathcal{O}_F$ will be the construction of tangent lines to $\mathcal{O}_F$ in all points of $X \subset \mathcal{O}_F$. Choose $N \geq 1$ and define for $n \geq N$

$$T_n : X_N \rightarrow \mathbb{P}^1$$

as follows. Let $x \in X_N$ and let $B^n(x) \in \mathcal{B}^n$, $n \geq N$, be the piece with $x \in B_n(x)$. The part $\mathcal{O}_F \cap B^n(x)$ is contained in a θ^n -stick see Fig. 12. The direction of the longest edge of this stick is denoted by $T_n(x) \in \mathbb{P}^1$.

The *a priori* bounds give that the scaling σ_1 of $B^{n+1}(x)$ is strictly away from zero. Namely, $\sigma_1 = \sigma_{B_{n+1}(x)} \geq \sigma_{B_{n+1}(x)}^* - \theta^n \geq a > 0$.

The angle between $T_n(x)$ and $T_{n+1}(x)$ is of the order θ^n , see Fig. 12. The piecewise constant functions T_n form a Cauchy sequence,

$$\text{dist}(T_{n+1}(x), T_n(x)) = O(\theta^n). \quad (10.2)$$

for $n \geq N$ and $x \in X_N$. The limit is denoted by

$$T = \lim_{n \rightarrow \infty} T_n : X_N \rightarrow \mathbb{P}^1.$$

The construction implies that we get in fact a map

$$T : X \rightarrow \mathbb{P}^1.$$

The actual line through $x \in X \subset \mathcal{O}_F$ with direction $T(x)$ is denoted by $T_x \subset \mathbb{R}^2$.

Definition 10.1. The Cantor set $\mathcal{O}_F$ is *almost everywhere* $(1 + \beta)$ -differentiable if for each $N \geq 1$ there exists $C_N > 0$ such that

$$\text{dist}(x, T_{x_0}) \leq C_N |x - x_0|^{1+\beta}$$

when $x \in \mathcal{O}_F$, $x_0 \in X_N$.

The tangent line field of $\mathcal{O}_F$ is *weakly β -Hölder* if for each $N \geq 1$ there exists $C_N > 0$ such that

$$\text{dist}(T(x_0), T(x_1)) \leq C_N |x_0 - x_1|^\beta,$$

with $x_0, x_1 \in X_N$.

Remark 10.1. The objects we consider have Hölder estimates on the growing sets X_N . Although, the increasing sequence of sets $X_1 \subset X_2 \subset X_3 \subset \dots$ is intrinsically related to the notion of being almost everywhere Hölder we will suppress it in the notation, instead of using *almost everywhere Hölder with respect to the sequence* $\{X_N\}$.

Theorem 10.1. The Cantor set $\mathcal{O}_F$ is almost everywhere $(1 + \beta)$ -differentiable, where $\beta > 0$ is universal. The tangent line field is weakly β -Hölder.

Proof. Choose $N \geq 1$. Let

$$d_N = \min_{B \in \mathcal{B}^N} \text{diam}(B \cap \mathcal{O}_F) > 0.$$

Choose, $x_0, x_1 \in X_N$. We will find a uniform Hölder estimate for the function $T|_{X_N}$ in these two points. Let $n \geq 1$ such that $x_1 \in B^n(x_0)$ and $x_1 \notin B^{n+1}(x_0)$. To prove a Hölder estimate we may assume that $n \geq N$. The *a priori* bounds for the Cantor set of the one-dimensional map f_* and the probabilistic universality of $\mathcal{O}_F$ observed in the sets X_N , see (10.1), give a $\rho < 1$ such that

$$|x_1 - x_0| \geq \rho^{n-N} \cdot d_N.$$

Estimate (10.2) implies

$$\begin{aligned} \text{dist}(T(x_1), T(x_0)) &\leq \text{dist}(T(x_1), T_n(x_1)) + \text{dist}(T_n(x_0), T(x_0)) \\ &= O(\theta^n) \\ &\leq C_N |x_1 - x_0|^\beta. \end{aligned}$$

where $C_N = O(\frac{\theta^N}{(d_N)^\beta})$ and $\beta > 0$ is such that

$$\rho^\beta = \theta. \tag{10.3}$$

The estimate only holds when x_0 and x_1 are in the same piece of $\mathcal{B}^N$. To get a global estimate we might have to increase the constant to obtain

$$\text{dist}(T(x_1), T(x_0)) \leq C_N |x_1 - x_0|^\beta,$$

for any pair $x_0, x_1 \in X_N$.

Choose $x \in \mathcal{O}_F$ to prove that $T_{x_0}, x_0 \in X_N$, is a β -Hölder tangent line to $\mathcal{O}_F$. Again let $n \geq 1$ such that $x \in B^n(x_0)$ and $x \notin B^{n+1}(x_0)$. The distance between x_0 and x is bounded from below when $n < N$. To find the Hölder estimate for the distance between x and T_{x_0} we may assume that $n \geq N$. Recall, $\text{dist}(T(x_0), T_n(x_0)) = O(\theta^n)$ and $|x - x_0| \geq \rho^{n-N} \cdot d_N$. Denote the length of the stick which contains $\mathcal{O}_F \cap B_n(x_0)$ by $l > 0$. The a priori bounds imply

$$l = O(|x - x_0|).$$

Then

$$\begin{aligned} \text{dist}(x, T_{x_0}) &= O(\theta^n) \cdot l \\ &= O((\rho^n)^\beta |x - x_0|) \\ &\leq C_N |x - x_0|^{1+\beta}. \end{aligned} \tag{10.4}$$

This estimate holds when x_0, x are in the same piece of $\mathcal{B}^N$. We might have to increase the constant C_N to get a global Hölder estimate. $\square$

In [7] it has been shown that the Cantor attractors $\mathcal{O}_F$, with $b_F > 0$, can not be part of a smooth curve.

Theorem 10.2. *Each set $X_N \subset \mathcal{O}_F$ is contained in a $C^{1+\beta}$ -curve.*

Proof. The proof will not use the specific structure of the set X_N described by the pieces in $\mathcal{B}^n$. The proof holds for every closed set in the plane with tangents line to each point with Hölder dependence on the point.

We will construct a $C^{1+\beta}$ -curve through every set $X_N \cap B$ with $B \in \mathcal{B}^{N+K}$ and $K \geq 0$ large enough. This suffices to prove the Theorem.

Choose $B \in \mathcal{B}^{N+K}$ with $X_N \cap B \neq \emptyset$. For each $x_0 \in X_N \cap B$ consider the cusps

$$S_{x_0} = \{x \in B | \text{dist}(x, T_{x_0}) < C_N |x - x_0|^{1+\beta}\}.$$

Note $X_N \cap B \subset S_{x_0}$. Thus

$$S \equiv \bigcap_{x \in X_N} S_x \supset X_N \cap B.$$

Fix $K \geq 0$ large enough such that each $S_x \setminus \{x\}$ has two components. This defines already an order on $X_N \cap B$. Write

$$S_x \setminus \{x\} = S_x^+ \cup S_x^-,$$

where $S_x^\pm$ are the connected components. We may assume that the assignment of connected components preserves the order in the following sense. If $x_1 \in S_{x_0}^+$ then

$$S_{x_1}^+ \cap X_N \subset S_{x_0}^+.$$

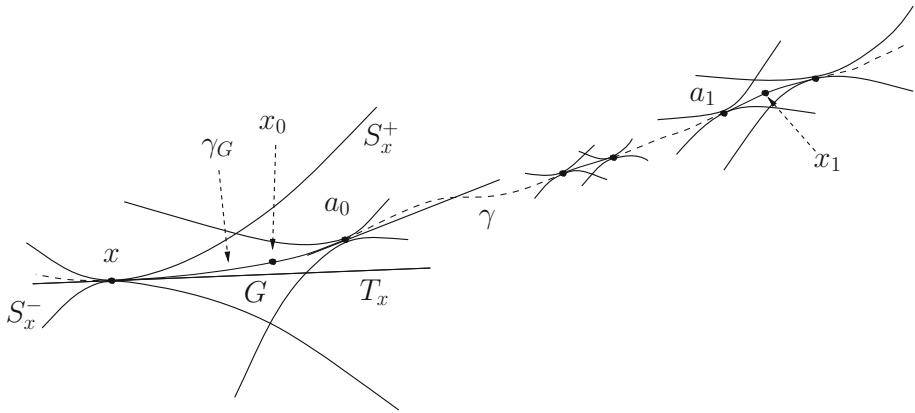


Fig. 13. Illustration for the proof of Theorem 10.1

A point $x \in X_N$ is a boundary point of X_N if $S_x^+ \cap X_N = \emptyset$ or $S_x^- \cap X_N = \emptyset$. A connected component $G \subset S \setminus X_N$, see Fig. 13, is called a gap of X_N . For every gap there exist two boundary points $x_0, x_1 \in X_N$ such that

$$G \subset S_{x_0}^+ \cap S_{x_1}^-.$$

Consider a gap between two boundary points x_0 and x_1 and the graph over the tangent line T_{x_0} of a cubic polynomial γ_G which passes through x_0 and x_1 and is tangent to the tangent lines T_{x_0} and T_{x_1} . Denote the graph of γ_G also by γ_G . A calculation shows that

$$|\gamma_G|_0 \leq 7C_N |x_0 - x_1|^{1+\beta}$$

and if $D\gamma_G(x) \in \mathbb{P}^1$ is the direction of the tangent line to the graph γ_G at a point $x \in \gamma_G$ then

$$|D\gamma_G(y) - D\gamma_G(x)| \leq 21C_N |y - x|^\beta.$$

In particular, the distance between the tangent directions along the curve and the direction at the boundary points shrink to zero as the diameter of the gap shrinks. This implies that the closure of the union of the curves γ_G

$$\gamma = X_N \cup \bigcup_G \gamma_G$$

is a C^1 curve.

Left is to show that the tangent direction $D\gamma$ is C^β . Choose $x_0, x_1 \in \gamma$. Let $a_0 \in \gamma \cap X_N$ be the closest point to x_0 on the line segment between x_0 and x_1 . Similarly, let a_1 be the closest point to x_1 . If $x_0 \in G$ then a_0 is a boundary point of the gap G . See Fig. 13. For $K \geq 0$ large enough, the distances between these points are, up to a factor close to 1, equal to the corresponding distances of the projections of these points to the tangent line through a_0 . We may assume that $|x_1 - a_1|, |a_1 - a_0|, |a_0 - x_0| \leq 2|x_1 - x_0|$. Then

$$\begin{aligned} |D\gamma(x_1) - D\gamma(x_0)| &\leq C_N \cdot \{21|x_1 - a_1|^\beta + |a_1 - a_0|^\beta + 21|a_0 - x_0|^\beta\} \\ &\leq 86C_N |x_1 - x_0|^\beta. \end{aligned}$$

The curve γ is $C^{1+\beta}$ and contains $X_N \cap B$. $\square$

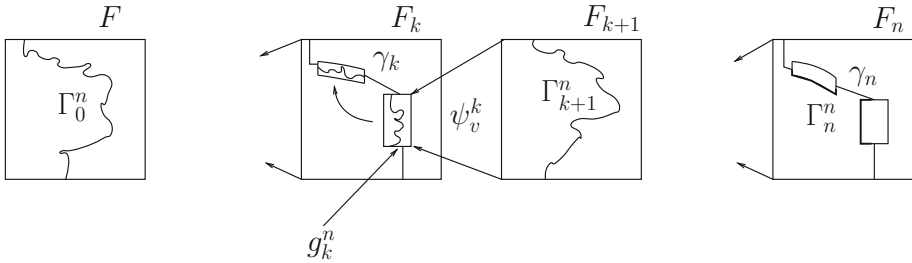


Fig. 14. Illustration for the proof of Theorem 10.3

The following Theorem is an answer to a question posed by J.C. Yoccoz.

Theorem 10.3. *The Cantor attractor $\mathcal{O}_F$ is contained in a rectifiable curve without self-intersections.*

Proof. Let $F_n : [0, 1]^2 \rightarrow [0, 1]^2$ be the n th-renormalization of F . The piece $B_v^1(F_n) \subset \text{Dom}(F_n)$ is strip bounded between two horizontal line segments and $B_c^1(F_n) \subset \text{Dom}(F_n)$ is strip bounded between two vertical line segments. Let γ_n be a collection of three line segments which connects the two pieces and each piece with the horizontal boundaries of $\text{Dom}(F_n) = [0, 1]^2$, see Fig. 14.

For each $n \geq 1$ we will construct inductively a curve Γ^n in the domain of F which passes through all pieces $B \in \mathcal{B}^n$ of the n th-cycle of F . Let Γ_n^n consists of γ_n and curves in the boundaries of $B_v^1(F_n)$ and $B_c^1(F_n)$ connecting the end points of γ_n , see Fig. 14.

Suppose Γ_{k+1}^n is defined and its end point are in the two horizontal boundary part of the domain of F_{k+1} , see Fig. 14. Let Γ_k^n be the curve connecting the top and bottom of the domain of F_k consists of the curves

$$\Gamma_k^n = \psi_v^k(\Gamma_{k+1}^n) \cup \psi_c^k(\Gamma_{k+1}^n) \cup \gamma_k \cup g_k^n,$$

where g_k^n consists of the two shortest horizontal line segments connecting the endpoints of $\psi_v^k(\Gamma_{k+1}^n)$ with the end points of γ_k and the two vertical line segments connecting the endpoints of $\psi_c^k(\Gamma_{k+1}^n)$ with the end points of γ_k , see Fig. 14. Let $\Gamma^n = \Gamma_0^n$.

The curve Γ^{n+1} is obtained from Γ^n by changing it inside the pieces of $\mathcal{B}^n$. Hence,

$$\Gamma^{n+1} \setminus \mathcal{B}^n = \Gamma^n \setminus \mathcal{B}^n.$$

This refinement process induces natural parametrizations of the curves Γ^n where the parametrization of Γ^{n+1} is obtained from the one of Γ^n by only adjusting only inside the pieces of $\mathcal{B}^n$. In each piece $B \in \mathcal{B}^n$, the curve Γ^{n+1} is partitioned into five sub-curves, see Fig. 14. The refinement of the parametrization of Γ^n spends equal time in each of these five sub-curves. The diameter of the pieces in $\mathcal{B}^n$ decay exponentially fast, $\sup_{B \in \mathcal{B}^n} \text{diam}(B) = O(\sigma^n)$. The construction and this decay imply that the parametrization have a uniform Hölder bound. This bound allows us to take a limit. Let Γ be the limiting Hölder curve. It contains $\mathcal{O}_F$.

The maps ψ_v^k and ψ_c^k are contracting distance by at least $\frac{1}{2.5}$, for $k \geq 1$ large enough, see Lemma 3.1. Denote the length of Γ_k^n by $|\Gamma_k^n|$. Then,

$$\begin{aligned} |\Gamma_k^n| &\leq \frac{2}{2.5} \cdot |\Gamma_{k+1}^n| + |\gamma_k| + |g_k^n| \\ &\leq \frac{2}{2.5} \cdot |\Gamma_{k+1}^n| + 4. \end{aligned}$$

The curves Γ_k^n have a bounded length. In particular, the limiting curve Γ is rectifiable.

Outside the pieces $B \in \mathcal{B}^n$ the curve Γ coincides with Γ^n which consists of non-intersecting curves. A self-intersection has to be a point $x \in \mathcal{O}_F$. Let $B^n(x) \in \mathcal{B}^n$ the piece which contains this self-intersection. The interval of parameter values which correspond to points in $B^n(x)$ is an interval of length $O(1/5^n)$. This means that the parametrization is injective. There are no self-intersections. $\square$

Remark 10.2. The curve Γ for the degenerate maps follows the same combinatorial construction as for a non-degenerate maps. This implies that the order of the pieces $B \in \mathcal{B}^n$ in the curve Γ is the same order as observed in one-dimensional maps.

Remark 10.3. The relative height (or thickness) of a piece $B \in \mathcal{B}^n$ coincides with the number $\beta(B) \leq 1$ introduced by P. Jones. In [13], Jones characterizes sets which are contained in rectifiable curves. A set $\mathcal{O}$ is contained in a rectifiable curve if and only if its diadic covers $\mathcal{B}^n$ satisfy the summability condition

$$\sum_{n \geq 1} \sum_{B \in \mathcal{B}^n} \beta^2(B) \cdot \text{diam } B < \infty.$$

In the present case of $\mathcal{O}_F$, one can use the dynamical covers $\mathcal{B}^n$ instead of the diadic ones. Since $\text{diam}(B) = O(\sigma^n)$ with $2\sigma < 1$, the set $\mathcal{O}_F$ satisfies the summability condition with respect to these covers. The diameter of the pieces decay fast enough so that we do not have to consider actual geometric information of the pieces: the bound $\beta(B) \leq 1$ suffices. For completeness we include a direct proof for rectifiability using the strongly contracting rescalings ψ_c^k and ψ_v^k .

The sets X_N have better geometric properties. The relative height (or thickness) of the pieces covering X_N and the corresponding numbers $\beta(B)$ decay exponentially fast. This is responsible for the smooth curves containing these sets.

The *tangent bundle* over $\mathcal{O}_F$ is defined by

$$TX = \{(x, v) \in X \times \mathbb{R}^2 | v \in T_x\}.$$

If $Y \subset X$ then the tangent bundle over Y is denoted by

$$TY = \{(x, v) \in TX | x \in Y\}.$$

We identify $T_x \subset \mathbb{R}^2$, $\{x\} \times T(x) \subset TX$ with the *tangent space* at $x \in X \subset \mathcal{O}_F$. Let $\pi_x : \mathbb{R}^2 \rightarrow T_x$ be the orthogonal projection.

Let $Y \subset \mathcal{O}_{F_1}$. A map $h : Y \rightarrow h(Y) \subset \mathcal{O}_{F_2}$ is differentiable at $x_0 \in Y$ if x_0 and $h(x_0)$ have a tangent line, and there exists a linear $Dh(x_0) : T_{x_0} \rightarrow T_{h(x_0)}$ such that for $x \in Y$

$$h(x) = h(x_0) + Dh(x_0)(\pi_{x_0}(x) - x_0) + o(|x - x_0|).$$

We will identify $Dh(x_0)$ with a number.

A bijection $h : X \rightarrow h(X) \subset \mathcal{O}_{F_2}$ is *almost everywhere* a $(1 + \beta)$ -diffeomorphism if for each $N \geq 1$ the restriction $h|X_N$ is differentiable at each $x \in X_N$ and

$$Dh : TX_N \rightarrow Th(X_N)$$

and its inverse are β -Hölder homeomorphisms.

Let $\mathcal{O}_{F_*}$ be the Cantor attractor of the fixed point of renormalization, the degenerate map F_* . Its invariant measure is denoted by μ_* . In [16] it has been shown that every

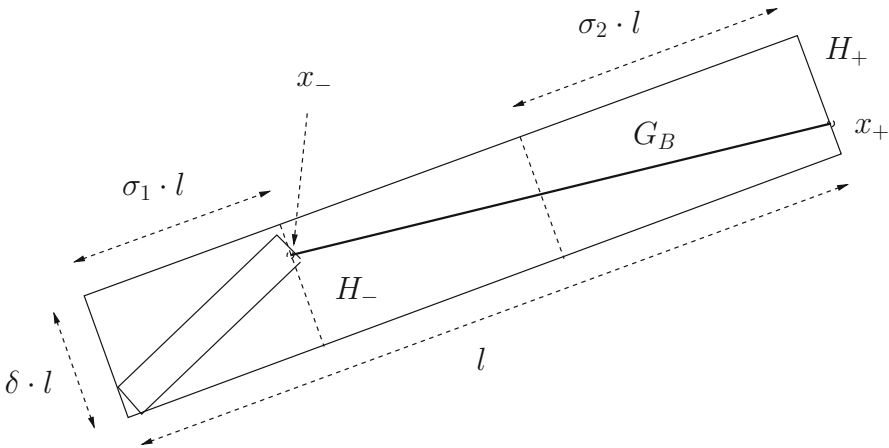


Fig. 15. Illustration for the proof of Theorem 10.4

conjugation which extends to a homeomorphism between neighborhoods of $\mathcal{O}_F$ and $\mathcal{O}_{F_*}$ respects the orbits of the tips. We will only consider conjugations

$$h : \mathcal{O}_F \rightarrow \mathcal{O}_{F_*}$$

with $h(\tau_F) = \tau_{F_*}$.

Definition 10.2. The attractor $\mathcal{O}_F$ of an infinitely renormalizable Hénon map $F \in \mathcal{H}_\Omega(\bar{\epsilon})$ is *probabilistically rigid* if there exists $\beta > 0$ such that the restriction $h : X \rightarrow h(X)$ of the conjugation $h : \mathcal{O}_F \rightarrow \mathcal{O}_{F_*}$, is almost everywhere a $(1+\beta)$ -diffeomorphism.

Theorem 10.4. The Cantor attractor $\mathcal{O}_F$ is probabilistically rigid.

Proof. Fix $N \geq 1$ and choose $B^0 \in \mathcal{S}_N(\theta^N)$ which intersects X_N . Consider the stick which contains B^0 . Call one of the long edges of this stick the bottom and choose an orientation of this line segment. It suffices to show the differentiability of the conjugation restricted to such a piece.

We will construct a curve containing $X_N \cap B^0$. This curve will be the closure of a countable collection of pairwise disjoint line segments. These line segments are called gaps. This piecewise affine curve is better adapted to the problem at hand than the curve of Theorem 10.2. Let

$$\mathcal{X}_N(k) = \{B \in \mathcal{S}_k(\theta^k) \mid B \cap X_N \neq \emptyset \text{ and } B \subset B^0\}.$$

Given $B \in \mathcal{X}_N(k)$. Let $\delta > 0$ be the relative height of the stick of B and $\sigma_1, \sigma_2 > 0$ the scaling factors of the two pieces $B_1, B_2 \in \mathcal{B}^{k+1}$ contained in B . The stick of B has three parts. Two rectangles of relative length σ_1 and σ_2 containing respectively B_1 and B_2 and the complement within the stick. This last part does not intersect X_N . It could be that one of the other parts also does not intersect X_N . At least one of the parts does intersect X_N . Let E be the union of the parts which do not intersect X_N and H_- and H_+ be the vertical boundaries of E , see Fig. 15.

The gap of B will be a line segment G_B connecting H_- with H_+ . Let $B_l \in \mathcal{X}_N(l)$ which intersect H_+ , $l = k, \dots, L$. Choose

$$x_B^+ \in H_+ \cap \mathcal{O}_F \cap \bigcap_{l=k}^L B_l.$$

The point x_B^+ is uniquely defined when $L = \infty$. In fact, it will be a point of X_N . When $L < \infty$ we have some freedom choosing x_B^+ . Choose it to be the closest point to the bottom of B_0 . Similarly, choose a point $x_B^- \in H_-$. The gap of B , denoted by G_B , is the line segment (x_B^-, x_B^+) .

The length of a gap is defined by

$$|G_B| \equiv |x_B^+ - x_B^-|.$$

Remark 10.4. The gaps are pairwise disjoint. For $B_1 \in \mathcal{X}_N(k+1)$ and $B \in \mathcal{X}_N(k)$ it might happen that G_{B_1} and G_B have a common endpoint. The angle between the gap G_B and the bottom of $B \in \mathcal{X}_N(k)$ is of order θ^k . This is a consequence of $\delta = O(\theta^k)$ and the *a priori* bounds on σ_1 and σ_2 .

There is a natural order on $X_N \cap B^0$ and the collection of gaps. It coincides with the order of the projections of X_N and the gaps onto the bottom of B^0 . Let us define the order between some $x \in X_N \cap B^0$ and a gap G_{B_1} . Let $k \geq N$ be maximal such that there is $B \in \mathcal{X}_N(k)$ with $x \in B$ and $G_{B_1} \cap B \neq \emptyset$. The stick of B has three parts as described above. Observe, x and G_{B_1} cannot be in the same part of the stick of B . The angle of the axis of B with the bottom of B^0 is of order θ^N . This defines an order on the three parts of this stick. Accordingly, this defines whether $x > G_{B_1}$, or, $x < G_{B_1}$.

The *gap-distance* between $x, y \in X_N \cap B^0$ is

$$|x - y|_g = \sum_{x < G_B < y} |G_B|.$$

The gaps between $x, y \in X_N \cap B^0$ form a curve

$$[x, y]_g \equiv \overline{\bigcup_{x < G_B < y} G_B}.$$

It is a graph over the tangent line of x .

Claim 10.5. If $x, y \in B \cap X_N$ with $B \in \mathcal{X}_N(k)$ then

$$\frac{|x - y|_g}{|x - y|} = 1 + O(\theta^k).$$

Proof. Let π_x be the projection onto the tangent line T_x of x . Then

$$|x - \pi_x(y)| = \sum_{x < G_{B'} < y} |\pi_x(G_{B'})|. \quad (10.5)$$

The angle between each gap $G_{B'}$ between x and y , and the tangent line of x is of order θ^k , see (10.2) and remark 10.4. This implies that

$$\frac{|\pi_x(G_{B'})|}{|G_{B'}|} = 1 + O(\theta^k). \quad (10.6)$$

The Cantor set $\mathcal{O}_F$ is almost everywhere differentiable, see Theorem 10.1. In particular, use (10.4) to obtain

$$\frac{|x - \pi_x(y)|}{|x - y|} = 1 + O(\theta^k). \quad (10.7)$$

The estimates (10.5), (10.6), and (10.7) prove the Claim. $\square$

Given a piece B of F , the corresponding piece of F_* is denoted by $B^* = h(B)$.

Claim 10.6. Let $B_l \in \mathcal{X}_N(l)$ with $B_l \subset B_k \in \mathcal{X}_N(k)$. Then

$$\ln \frac{|G_{B_l}|}{|G_{B_k}|} \cdot \frac{|G_{B_k^*}|}{|G_{B_l^*}|} = O(\theta^k).$$

Proof. The Claim holds for $l = k + 1$ because the relevant pieces are in $\mathcal{S}_k(\theta^k)$ and $\mathcal{S}_{k+1}(\theta^{k+1})$. In general, there is a unique sequence of pieces $B_j \in \mathcal{X}_N(j)$, $k \leq j \leq l$ with $B_l \subset B_{l-1} \subset \dots \subset B_{k+1} \subset B_k$. Then

$$\ln \frac{|G_{B_l}|}{|G_{B_k}|} \cdot \frac{|G_{B_k^*}|}{|G_{B_l^*}|} = \sum_{j=k}^{l-1} \ln \frac{|G_{B_{j+1}}|}{|G_{B_j}|} \cdot \frac{|G_{B_j^*}|}{|G_{B_{j+1}^*}|} = \sum_{j=k}^{l-1} O(\theta^j) = O(\theta^k).$$

□

Claim 10.7. Let $x, y, z \in X_N \cap B$ with $B \in \mathcal{X}_N(k)$ and $x^*, y^*, z^* \in h(X_N)$ the corresponding images under h . Then

$$\ln \frac{|x - y|_g}{|x - z|_g} \cdot \frac{|x^* - z^*|_g}{|x^* - y^*|_g} = O(\theta^k).$$

Proof. Claim 10.6 gives for every piece $\tilde{B} \subset B$

$$|G_{\tilde{B}}| = |G_{\tilde{B}^*}| \cdot \frac{|G_{\tilde{B}}|}{|G_{\tilde{B}^*}|} \cdot (1 + O(\theta^k)).$$

This implies

$$\begin{aligned} \frac{|x - y|_g}{|x - z|_g} &= \frac{\sum_{x < \tilde{B} < y} |G_{\tilde{B}}|}{\sum_{x < \tilde{B} < z} |G_{\tilde{B}}|} \\ &= \frac{\sum_{x^* < \tilde{B}^* < y^*} |G_{\tilde{B}^*}|}{\sum_{x^* < \tilde{B}^* < z^*} |G_{\tilde{B}^*}|} \cdot (1 + O(\theta^k)) \\ &= \frac{|x^* - y^*|_g}{|x^* - z^*|_g} \cdot (1 + O(\theta^k)). \end{aligned}$$

This finishes the proof of the Claim. □

A reformulation of this Claim is the following. Let $x, y, z \in X_N \cap B$ with $B \in \mathcal{X}_N(k)$. Then

$$\left| \ln \frac{|h(y) - h(x)|_g}{|y - x|_g} - \ln \frac{|h(z) - h(x)|_g}{|z - x|_g} \right| = O(\theta^k). \quad (10.8)$$

This implies that for $x, y \in X_N \cap B^0$ the following limit exists.

$$Dh(x) = \lim_{y \rightarrow x} \frac{|h(y) - h(x)|_g}{|y - x|_g}.$$

Moreover, the limit depends continuously on x .

Claim 10.8. There exists a universal $\beta > 0$, independent of N , such that $Dh : X_N \rightarrow \mathbb{R}$ is β -Hölder.

Proof. Choose $x_0, x \in X_N \cap B^0$ to prove a Hölder estimate for $\ln Dh$. Let $k \geq N$ be maximal such that $x \in B_k(x_0)$. Observe, as before in the proof of Theorem 10.1,

$$|x - x_0| \geq \rho^{k-N} \cdot \text{diam}(B^0)$$

where $\rho < 1$. Choose $\beta > 0$ such that $\rho^\beta = \theta$. Then

$$\theta^k = O(|x - x_0|^\beta). \quad (10.9)$$

Hence, using (10.8) and (10.9),

$$|\ln Dh(x) - \ln Dh(x_0)| = O(\theta^k) = O(|x - x_0|^\beta).$$

This suffices to show the Hölder bound for Dh . $\square$

We will identify $Dh(x)$ with a linear map $Dh(x) : T_x \rightarrow T_{h(x)}$. The positive function Dh is bounded. This bound, (10.8), and Claim 10.5, imply that for $x, x_0 \in X_N$

$$|h(x) - h(x_0)| = O(|x - x_0|). \quad (10.10)$$

Claim 10.9. For $x, y \in X_N \cap B^0$

$$|h(y) - h(x)| = Dh(x) \cdot |x - y| \cdot (1 + O(|x - y|^\beta)).$$

Proof. Let $k \geq N$ be maximal such that $y \in B^k(x)$. Apply Claim 10.5, (10.8), and (10.9), in the following estimate

$$\begin{aligned} |h(y) - h(x)| &= \frac{|h(y) - h(x)|}{|h(y) - h(x)|_g} \cdot |h(y) - h(x)|_g \\ &= (1 + O(\theta^k)) \cdot Dh(x) \cdot |y - x|_g \\ &= (1 + O(|y - x|^\beta)) \cdot Dh(x) \cdot |y - x|. \end{aligned}$$

$\square$

Now we are prepared to show the differentiability of h . Choose $x, x_0 \in X_N \cap B^0$. Let $k \geq N$ be maximal such that $x \in B_k(x_0)$. Let $\Delta = Dh(x_0)(\pi_{x_0}(x) - x_0) \in T_{h(x_0)}$. Claim 10.9, (10.7), and (10.9), imply

$$|\Delta| = |h(x) - h(x_0)| \cdot (1 + O(|x - x_0|^\beta)). \quad (10.11)$$

Let $J = \pi_{h(x_0)}(h(x)) - h(x_0) \in T_{h(x_0)}$ and $V = h(x) - \pi_{h(x_0)}(h(x))$. The image $h(\mathcal{O}_F)$ is contained in a smooth curve, the image of the degenerate map F_* . Hence,

$$\begin{aligned} |J| &= |h(x) - h(x_0)| \cdot (1 + O(|h(x) - h(x_0)|^2)) \\ &= |h(x) - h(x_0)| \cdot (1 + O(|x - x_0|^\beta)) \end{aligned} \quad (10.12)$$

and

$$|V| = O(|h(x) - h(x_0)|^2). \quad (10.13)$$

Apply (10.11), (10.12), (10.13), and (10.10), in the following estimate

$$\begin{aligned} h(x) &= h(x_0) + \Delta + (J - \Delta) + V \\ &= h(x_0) + \Delta + O(|h(x) - h(x_0)| \cdot |x - x_0|^\beta) + O(|h(x) - h(x_0)|^2) \\ &= h(x_0) + Dh(x_0)(\pi_{x_0}(x) - x_0) + O(|x - x_0|^{1+\beta}). \end{aligned}$$

This finishes the proof of the differentiability and the Theorem. $\square$

Remark 10.5. The conjugation $h : \mathcal{O}_F \rightarrow \mathcal{O}_{F_*}$ satisfies

$$h(x) = h(x_0) + Dh(x_0)(\pi_{x_0}(x) - x_0) + O(|x - x_0|^{1+\beta})$$

in almost every point $x_0 \in \mathcal{O}_F$. Observe, that the Hölder exponent is universal. The Hölder constant tends to infinity when h is restricted to larger and larger sets X_N , when $N \rightarrow \infty$.

The Cantor attractor $\mathcal{O}_F$ has two characteristic exponents, [21]. One is zero the other is $\ln b_F$, see [7]. The function $T : X \rightarrow \mathbb{P}^1$ constructed before defines a measurable line field, with respect to μ , on $\mathcal{O}_F$.

Proposition 10.10. *The line field*

$$T : \mathcal{O}_F \rightarrow \mathbb{P}^1$$

is the invariant line field of zero characteristic exponent.

Proof. For each point $x_0 \in X$ we have, see Theorem 10.1,

$$\text{dist}(x, T_{x_0}) \leq C_{x_0}|x - x_0|^{1+\beta}$$

with $x \in \mathcal{O}_F$. The map F is a diffeomorphism which preserves $\mathcal{O}_F$. Hence,

$$\text{dist}(x, DF(x_0)T_{x_0}) = O(|x - F(x_0)|^{1+\beta})$$

with $x \in \mathcal{O}_F$. For almost every $x_0 \in X$ we have $F(x_0) \in X$. Hence, T is an invariant line field, i.e. for almost every $x_0 \in \mathcal{O}_F$ we have

$$DF(x_0)T_{x_0} = T_{F(x_0)}.$$

The map F has only two invariant lines fields, the two characteristic directions, [21]. Left is to show that $T(x)$ corresponds to the zero exponent.

Choose $N \geq 1$. For almost every $x_0 \in X_N$ there are $t_n \rightarrow \infty$ such that

$$F^{t_n}(x_0) \in X_N.$$

This is because the ergodic measure μ assigns positive measure to X_N . Let $v \in T_{x_0}$ and $v_* \in T_{h(x_0)}$ be unit vectors. Apply the chain rule

$$\begin{aligned} |DF^{t_n}(x_0)v| &= |Dh^{-1}(F_*(h(x_0)))| \cdot |DF_*^{t_n}(h(x_0))Dh(x_0)v| \cdot |Dh(x_0)| \\ &\asymp |DF_*^{t_n}(h(x_0))v_*|. \end{aligned}$$

Observe, $v_* \in T_{h(x_0)}$ which is a tangent line to the graph of f_* . The degenerate Hénon map F_* has zero exponential contraction along this curve. Hence,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \ln |DF^t(x_0)v| = \lim_{n \rightarrow \infty} \frac{1}{t_n} \ln |DF^{t_n}(x_0)v| = 0$$

On a set of full measure in X_N there is no exponential contraction along the direction $T(x)$. The line field T has exponent zero. $\square$

The Hausdorff dimension of a measure μ on a metric space $\mathcal{O}$ is defined as

$$HD_\mu(\mathcal{O}) = \inf_{\mu(X)=1} HD(X).$$

Theorem 10.11. *The Hausdorff dimension of the invariant measure is universal*

$$HD_\mu(\mathcal{O}_F) = HD_{\mu_*}(\mathcal{O}_{F_*}).$$

Proof. Let $h : \mathcal{O}_F \rightarrow \mathcal{O}_{F_*}$ be a conjugation which exchanges the orbits of the tips. According to Theorem 10.4 there are sets $X_N \subset \mathcal{O}_F$ with $\mu(X_N) \geq 1 - O(\theta^N)$ and on which h is a $(1 + \beta)$ -diffeomorphism. The continuity of the derivative gives upper and lower bounds of the derivative. This implies

$$HD(h(X_N)) = HD(X_N).$$

Hence, for $X = \bigcup_{N \geq 1} X_N$ and every $Z \subset \mathcal{O}_F$

$$HD(h(X \cap Z)) = HD(X \cap Z).$$

Let $Z_N \subset \mathcal{O}_F$ with $\mu(Z_N) = 1$ and $\lim_{N \rightarrow \infty} HD(Z_N) = HD_\mu(\mathcal{O}_F)$ then

$$\begin{aligned} HD_\mu(\mathcal{O}_F) &\geq \lim_{N \rightarrow \infty} HD(Z_N \cap X) \\ &= \lim_{N \rightarrow \infty} HD(h(Z_N \cap X)) \\ &\geq HD_{\mu_*}(\mathcal{O}_{F_*}), \end{aligned}$$

where the last inequality holds because $\mu_*(h(Z_N \cap X)) = \mu(Z_N \cap X) = 1$. The opposite inequality $HD_{\mu_*}(\mathcal{O}_{F_*}) \geq HD_\mu(\mathcal{O}_F)$ is obtained in the same way. $\square$

Remark 10.6. We can identify the Hausdorff dimension of the measure on the Cantor attractor. Namely,

$$HD_\mu(\mathcal{O}_F) = \frac{\ln 2}{\int \ln |Dr_*| d\mu_*}.$$

where r_* is the analytic expanding one dimensional map constructed such that $\pi_1(\mathcal{O}_{F_*})$ is its invariant Cantor set, see for example [2] and references therein. The measure μ_* is the projected measure from $\mathcal{O}_{F_*}$.

Acknowledgements We thank all the institutions and foundations that have supported us in the course of this work: Simons Mathematics and Physics Endowment, Fields Institute, NSF, NSERC, University of Toronto. In fall 2005, when M. Feigenbaum saw the negative results of [7], he made computer experiments that suggested that the universal scaling of the attractor is violated very rarely. Our paper provides a rigorous justification of Feigenbaum's experiments and conjectures. We also thank C. Tresser for many valuable renormalization discussions, and R. Schul for interesting comments on [13].

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix: Open Problems

Let us finish with some questions related to the previous discussion.

Problem I: The collections $\mathcal{P}_n$, see (8.7), of good pieces that we have constructed are determined by the average Jacobian of the map. Observe that $\mathcal{S}_n(\theta^n)$ might be slightly larger than $\mathcal{P}_n$. It was suggested by Feigenbaum's experiment, mentioned in the introduction, that the statistics of the remaining *bad* pieces, might be governed by some universality law. This problem is also related to one of the open problems in [7] on the regularity of the conjugation $h : \mathcal{O}_F \rightarrow \mathcal{O}_G$ when $b_F = b_G$.

Problem II: Do wandering domains exist? This question was already formulated in [16]. It is included again because its solution might be obtained by using the techniques developed in this paper.

List of Symbols

b_F	Average Jacobian
B_ω^n	A piece of the n th-renormalization level
$\mathcal{B}^n$	Collection of pieces in the n th-renormalization level
$\mathcal{B}^n[k]$	Pieces of $\mathcal{B}^n$ in E^k
$B_n(x)$	The piece in $\mathcal{B}^n$ containing $x \in \mathcal{O}_F$
$\mathbf{B}$	The piece B viewed from its proper scale
$\text{Dist}(\phi)$	Distortion
D_k	Derivative of ψ_v^k at the tip
δ_B	Thickness of B
Δ_B	Absolute thickness of B
E^k	Part of a dynamical partition
f_*	Unimodal renormalization fixed point
G_k	Return map related to the partition by E^k
$k_i(B)$	Depth of the i th-predecessor of B
$\kappa_0(n)$	Minimal depth to safely push-up
$\kappa(n)$	Upper bound of the brute-force regime
$l(k)$	Maximal allowable depth
η_ϕ	Nonlinearity
$\mathcal{O}_F$	Invariant Cantor set of F
$\psi_{c,v}^k$	Coordinate changes related to the renormalization $R(R^k F)$
ψ_ω^n	Coordinate change
Ψ_k^n	Coordinate change relating $R^{n-k}(R^k F)$ to R^n
$\mathcal{P}_n(k; q_0, q_1)$	Collection of q_0, q_1 -controlled pieces
$\mathcal{P}_n$	Pieces obtained by applying the three regimes
q_0, q_1	Boundary one-dimensional regime
σ	Scaling factor of the unimodal renormalization fixed point
σ_B	Scaling factor of B
$\mathcal{S}^n(\epsilon)$	Collection of pieces in $\mathcal{B}^n$ with ϵ precision

t_k	Tilt of the derivative of ψ_v^k at the tipnoalign
T	Tangent line field to $\mathcal{O}_F$ noalign
τ_F	Tipnoalign
X	The differentiable part of $\mathcal{O}_F$ noalign

References

- [1] Benedicks, M., Carleson, L.: The dynamics of the Hénon map. *Ann. Math.* **133**(1), 73–169 (1991)
- [2] Birkhoff, C., Martens, M., Tresser, C.P.: On the scaling structure for period doubling. *Asterisque* **286**, 167–186 (2003)
- [3] Collet, P., Eckmann, J.P., Koch, H.: Period doubling bifurcations for families of maps on $\mathbb{R}^n$. *J. Stat. Phys.* **25**, 1–15 (1980)
- [4] Chandramouli, V.V.M.S., Martens, M., de Melo, W., Tresser, C.P.: Chaotic period doubling. *Ergod. Theory Dyn. Syst.* **29**, 381–418 (2009)
- [5] Cvitanovic, P.: *Universality in Chaos*, 2nd edn. Adam Hilger, Bristol (1989)
- [6] Coulet, P., Tresser, C.: Itération d'endomorphismes et groupe de renormalisation. *J. Phys. Colloque C* **539**, C5-25 (1978)
- [7] de Carvalho, A., Lyubich, M., Martens, M.: Renormalization in the Hénon family, I: universality but non-rigidity. *J. Stat. Phys.* **121**(5/6), 611–669 (2005)
- [8] de Faria, E., de Melo, W., Pinto, A.: Global hyperbolicity of renormalization for C^r unimodal mappings. *Ann. Math.* **164**(3), 731–824 (2006)
- [9] de Melo, W., van Strien, S.: *One-Dimensional Dynamics*. Springer, Berlin (1993)
- [10] Epstein, H.: New proofs of the existence of the Feigenbaum functions. *Commun. Math. Phys.* **106**, 395–426 (1986)
- [11] Feigenbaum, M.J.: Quantitative universality for a class of non-linear transformations. *J. Stat. Phys.* **19**, 25–52 (1978)
- [12] Gambaudo, J.-M., van Strien, S., Tresser, C.: Hénon-like maps with strange attractors: there exist C^∞ Kupka–Smale diffeomorphisms on S^2 with neither sinks nor sources. *Nonlinearity* **2**, 287–304 (1989)
- [13] Jones, P.: Rectifiable sets and the traveling salesman problem. *Invent. Math.* **102**(1), 1–16 (1990)
- [14] Lanford, O.E., III.: A computer assisted proof of the Feigenbaum conjectures. *Bull. Am. Math. Soc. New Ser.* **6**, 427–434 (1982)
- [15] Lyubich, M.: Feigenbaum–Coulet–Tresser Universality and Milnor's Hairiness Conjecture. *Ann. Math.* **149**, 319–420 (1999)
- [16] Lyubich, M., Martens, M.: Renormalization in the Hénon family, II. The Heteroclinic Web, IMS Stony Brook preprint 08-2 and accepted for publication in *Invent. Math.*
- [17] Lyubich, M., Martens, M.: Renormalization of Hénon maps. In: M.M. Peixoto, A.A. Pinto, D.A. Rand (Eds.) *Dynamics, Games and Science I*, Springer Proceedings in Mathematics 1, pp. 597–618 (2011)
- [18] Martens, M.: Distortion results and invariant cantor sets for unimodal maps. *Ergod. Theory Dyn. Syst.* **14**, 331–349 (1994)
- [19] Martens, M.: The periodic points of renormalization. *Ann. Math.* **147**, 543–584 (1998)
- [20] McMullen, C.: Renormalization and 3-manifolds which fiber over the circle. *Annals of Mathematics Studies*, vol. 135. Princeton University Press, Princeton (1996)
- [21] Oseledec, V.I.: A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow. Math. Soc.* **19**, 197–231 (1969)
- [22] Sullivan, D.: Bounds, quadratic differentials, and renormalization conjectures. *AMS Centennial Publications. 2: Mathematics into Twenty-first Century* (1992)

Communicated by C. Liverani