# PROBABILISTIC UNIVERSALITY IN TWO-DIMENSIONAL DYNAMICS

M. LYUBICH, M. MARTENS

Abstract. In this paper we continue to explore infinitely renormalizable Hénon maps with small Jacobian. It was shown in [CLM] 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 weight assigned by the canonical invariant measure to these bad spots tends to zero on microscopic scales. This phenomenon is called *Probabilistic Universality*. It implies, in particular, that the Hausdorff dimension of the canonical measure is universal. In this way, universality and rigidity phenomena of one-dimensional dynamics assume a probabilistic nature in the two-dimensional world.

Contents



*Date*: June 24, 2011.



Stony Brook IMS Preprint #2011/2
*June 2011*

## 1. Introduction

Renormalization ideas have played a central role in Dynamics since the discovery of the Universality and Rigidity phenomena by Feigenbaum [F], and independently by Coullet and Tresser [CT], 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 physcis, geometric function theory, Teichmüller theory, hyperbolic geometry, infinite-dimensional complex geometry) and by now has been fully and rigorously justified, see [Ep], [FMP], [L], [Lan], [Ma2], [McM], [S] and references therein.

In [CT] Coullet 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 [CLM] 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 canonical invariant measure. *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 dissipative Hénon-like maps on the unit box $B^0 = [0,1]\times[0,1]$ of form

$$F(x,y) = (f(x) - \varepsilon(x,y), x), \tag{1.1}$$

where $f(x)$ is a unimodal map with non-degenerate critical point and $\varepsilon$ 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 box $B^1 \subset B^0$ around the tip of 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 $\Psi : B^0 \to B^1$ 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 renormalizble*, etc. In this paper we will be concerned with *infinitely renormalizable* Hénon-like maps. Such a map admits a nest of $2^n$-periodic boxes $B^0 \supset B^1 \supset B^2 \supset \dots$ shrinking to the *tip* $\tau$ of $F$. The $n^{th}$-renormalization cycle is the orbit $\mathcal{B}^n = \{B_i^n = f^i(B^n), i = 0, 1, \dots 2^n - 1\}$. We obtain a hierarchy of such cycles 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 canonical invariant measure $\mu$ supported on $\mathcal{O}_F$. We define the *average Jacobian* of $F$ as follows:

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

Consider a strongly dissipative infinitely renormalizable Hénon-like map. 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 $\epsilon-$*precision* if after one simultaneous rescaling and translation $A : \mathbb{R}^2 \to \mathbb{R}^2$ 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 $\epsilon \cdot \mathrm{diam}(I)$. The triples $B_1, B_2 \subset B$ and $I_1, I_2 \subset I$ are geometrical almost the same.

Collect the pieces of the $n^{th}-$cycle with $\epsilon-$precision in

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

**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 $h : \mathcal{O}_F \to \mathcal{O}_{f_*}$ 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 $X_1 \subset X_2 \subset X_3 \subset \cdots \subset \mathcal{O}_F$ such that $h : X_N \to h(X_N) \subset \mathcal{O}_{f_*}$ is $(1+\beta)$-differentiable, and $\mu(X_N) \to 1$.

**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 [CLM]. However, large parts of it, the sets

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

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 $\mu_*$ 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 geometrical 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 cycle 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 cycle 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 §2. Also more special notations are collected in the Nomenclature. For a survey on Hénon renormalization see [LM2]. For early experiments and results on Hénon renormalization see [CEK], [Cv], and [GST].

**Acknowledgment.** 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 [CLM], 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 [J].

## 2. Outline

2.1. **Infinitely renormalizable Hénon-like maps.** We will start with outlining the set-up developed in [CLM, LM1] – see §3 for details.

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

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

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 box around the *tip*, a point which plays the role of the "critical value", and $\Psi_0^1 : \mathrm{Dom}(F_1) \to B^1$ 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 box*, $\Psi_0^n : \mathrm{Dom}(F_n) \to B^n$ is a non-linear change of variables.

These boxes $B^n$ form a nest around the *tip* of $F$:

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

Taking the iterates $F^k B^n$, $k = 0, 1, \ldots, 2^n - 1$, we obtain a family $\mathcal{B}^n$ of $2^n$ pieces $\{B^n_\omega\}$, 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 $B^n_{v^n} \equiv B^n$. See §3 for details. Then

$$\mathcal{O}_F = \bigcap_n \bigcup_\omega B^n_\omega$$

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

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

Usually, we will denote the average Jacobian with $b$.

The size of the boxes decays exponentially:

$$\operatorname{diam} B^n_\omega \le C\sigma^n, \tag{2.1}$$

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

A surprising phenomenon discovered in [CLM] is that unlike its one-dimensional counterpart, the Cantor set $\mathcal{O}_F$ *does not have universal geometry*: it essentially depends on the average Jacobian $b$. However, the difference appears only in scale of order $b$: if all the pieces $B^n_\omega$ of level $n$ are much bigger than $b$ then the geometry of the pieces $B^n_\omega$ is controlled by one-dimensional dynamics: the pieces are aligned along the parabola $x = f(y)$ with thickness of order $b$. According to (2.1), this happens whenever

$$\alpha\sigma^n \ge b \tag{2.2}$$

with sufficienty small (absolute) $\alpha > 0$, i.e., when

$$n \le c|\log b| - A, \quad \text{where } c = \frac{1}{|\log \sigma|}, \; A = \frac{\log \alpha}{\log \sigma}. \tag{2.3}$$

We will call these levels *safe*.

## 2.2. **Random walk model.**

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

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

Here the tip is the only point of infinite depth. If $\operatorname{depth}(x) = k$ then $x \in E^k \equiv B^k \setminus B^{k+1}$ (see Figure 2.1 and 4.1). We say that a point $x \in \mathcal{O}$ is *combinatorially closer to $\tau$ 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 $\tau$ 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 $\tau$ than all previous points $F^{-t}x$, $t = 0, 1, \dots, t_m - 1$. Since the dynamics of $F$ on $\mathcal{O}$ is the adding machine, this is an infinite sequence of moments for any $x \notin \operatorname{orb}(\tau)$. If $x = F^t(\tau)$, we terminate the code at the moment $t$. Let

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

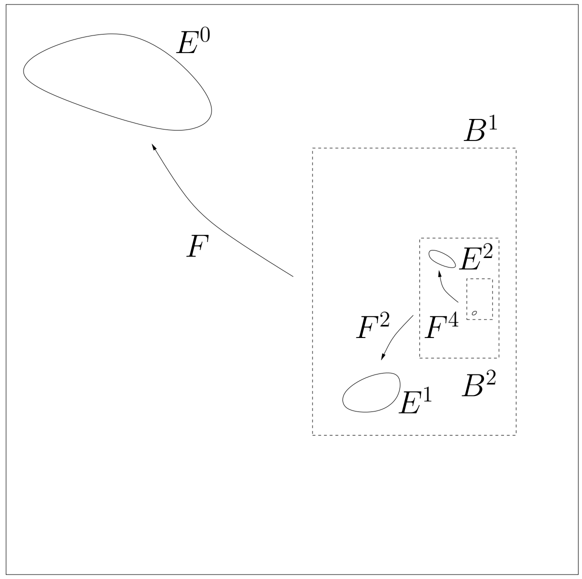


FIGURE 2.1

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 Figure 2.1 and 4.1).

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

and the *first landing map in backward time*

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

Then we have by definition:

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

Let $\Sigma$ 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 $\Sigma$ with a weak topology and the measure $\nu$ corresponding to the following *random walk* on $\mathbb{N}$: the probability of jumping from $k \in \mathbb{N}$ to $l \in \mathbb{N}$ is equal to $1/2^{l-k}$ if $l > k$, and it vanishes otherwise. The initial distribution on $\mathbb{N}$ is given by $\nu\{k\} = 1/2^{k+1}$. 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 $\Sigma$ and a measure-theoretic isomorphism between $(\mathcal{O}, \mu)$ and $(\Sigma, \nu)$.*

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 $k_m < n$ $(m = 0, 1 \ldots, T-1)$, 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* $s : \mathbb{N} \to \mathbb{Z}_+$. We say that a sequence $\bar{k} = \{k_m\}_{m=0}^\infty$ is *s-controlled after a moment* $N$ if $j_m \le s(k_m)$ for all $k_m \ge N$. 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 $K_N$.

**Lemma 2.2.** *Under the summability assumption*

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

*we have*

$$\nu(K_N) \ge 1 - O(\sum_{k=N}^{\infty} \frac{1}{2^{s(k)}}).$$

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

$$\nu(K_N) \ge \prod_{k=N}^{\infty} (1 - \frac{1}{2^{s(k)}}),$$

which implies the Lemma. □

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

The renormalizations have the following nearly *universal shape*:

$$R^n F = (f_n(x) - \ b^{2^n}\, a(x)\, y\, (1 + O(\rho^n)),\ x\ ), \tag{2.4}$$

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

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

$$\Psi_k^l = D_k^l \circ (\mathrm{id} + \mathbf{S}_k^l), \tag{2.5}$$

where

$$D_k^l = \begin{pmatrix} 1 & t_k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} (\sigma^2)^{l-k} & 0 \\ 0 & (-\sigma)^{l-k} \end{pmatrix} (1 + O(\rho^k)). \tag{2.6}$$

is a linear map with $t_k \asymp b_F^{2^k}$, while $\mathrm{id} + \mathbf{S}_k^l : (x, y) \mapsto (x + S_k^l(x, y), y)$ 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 boxes.** In this section we outline the results of §4.

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

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

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

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

We can view the box $B$ in the renormalization coordinates on various scales. Namely, to view $B$ from scale $k \le n$ means that we consider its preimage $\mathbf{B}$ under the (nonlinear) rescaling $\Psi_0^k : \mathrm{Dom}(F_k) \to B^k$. 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 box (see Figure 4.2). This seemingly minor ingredient plays a crucial role in the estimates.

A box $B$ 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. In other words, $\mathrm{mod}\,\mathbf{B}$ (the ratio of the the vertical and horizontal sizes of $\mathbf{B}$) is of order 1.

We will control depth by the control function

$$s(k) = a2^k - A \quad \text{where } a = \frac{\log b}{\log \sigma}, \ A = \frac{\log \alpha}{\log \sigma}, \tag{2.7}$$

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

$$\alpha \sigma^{l-k} \ge b^{2^k}. \tag{2.8}$$

Comparing it to (2.2) and (2.3), we see that the level $l - k$ controlled in this way is safe for the renormalization $F_k$.

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

$$l - k \le s(k).$$

There are a number of constant which have to be chosen appropriately, for example $\alpha$ 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.

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

*Outline of the proof.* Let $B \in \mathcal{B}^n[l]$, $n > l > k$. We should view $B$ from scale $l$, i.e., consider the piece $\mathbf{B}$ of level $n-l$ for the renormalization $F_l$, see Figure 4.3. As the piece $\tilde{B} = G_k(B)$ has depth $k$, it should be viewed from this depth. So, we consider the corresponding piece $\tilde{\mathbf{B}}$ of level $n-k$ for the renormalization $F_k$. Then $\tilde{\mathbf{B}} = F_k \circ \Psi^l_k(\mathbf{B})$.

Using geometric estimates for factorization (2.5) we show that

$$\operatorname{mod} \Psi^l_k(\mathbf{B}) \asymp \sigma^{l-k} \operatorname{mod} \mathbf{B},$$

provided $\mathbf{B}$ is regular. So $\Psi^l_k(\mathbf{B})$ is highly stretched in the vertical direction. The nearly Universal map $F_k$, see (2.4), will contract the vertical size by a factor of order $b^{2^k} << \sigma^{l-k}$ 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. □

2.5. **Universal sticks.** Given a box $B \in \mathcal{B}^n[l]$ of a map $F$, let $\mathcal{O}(B) := \mathcal{O}_F \cap B$ be the part of the postcritical 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 box corresponding to $B$.

We say that a box $B \in \mathcal{B}^n[l]$ is a $\delta$-stick if the postcritical set $O(\mathbf{B})$ is contained in a diagonal strip $\Pi$ of thickness $\delta$, relatively the horizontal size of $\mathbf{B}$. The minimal thickness is denoted by $\delta_{\mathbf{B}}$. See Figure 5.1.

Let us consider the pieces $B_1$ and $B_2$ of level $n+1$ contained in $B$. The corresponding pieces $\mathbf{B}_1$ and $\mathbf{B}_2$ occupy fractions $\sigma_{\mathbf{B}_1}$ and $\sigma_{\mathbf{B}_2}$ of $\mathbf{B}$, called *scaling ratios*, see Figure 6.1. Let $\sigma^*_{\mathbf{B}_1}$ and $\sigma^*_{\mathbf{B}_2}$ be the scaling ratios of the corresponding pieces for the degenerate renormalization fixed point $F_*$. Let $\Delta\sigma_{\mathbf{B}}$ be the maximal difference between the corresponding scaling ratios.

A piece $B \in \mathcal{B}^n$ is called $\varepsilon$-*universal* if $\delta_{\mathbf{B}} \le \epsilon$ and $\Delta\sigma_{\mathbf{B}} \le \epsilon$.

Consider very deep pieces $B \in \mathcal{B}^n[k]$, with $(1-q_0)\cdot n \le k \le n$, at scale $n-k$. Then we are watching pieces of $\mathcal{B}^{n-k}(F_k)$ which can be obtained by following the orbit of $B^{n-k}_v(F_k)$ for $2^{n-k}$ steps. $F_k$ is at a distance

$O(\rho^k)$ to the degenerate renormalization fixed point $F_*$. When $q_0 > 0$ is small, 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 geometrical model. We call this the *one-dimensional regime.*

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

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

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

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

*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 $\tilde{\mathbf{B}} = F_k(\Psi^l_k(\mathbf{B}))$.*

*Outline of the proof.* We consecutively estimate, using geometric estimates of §2.3, the relative thickness of the pieces $B_{\text{diff}} = (\text{id} + \mathbf{S}^l_k)(\mathbf{B})$, $B_{\text{aff}} = D^l_k(B_{\text{diff}})$ and $\tilde{\mathbf{B}} = F_k(B_{\text{aff}})$, see Figure 4.4. The first one is comparable with the thickness of $\mathbf{B}$, up to an error of order $\sigma^{n-l}$, since the horizontal map $\text{id} + \mathbf{S}^l_k$ has bounded geometry (where the error $\sigma^{n-l} \ge \operatorname{diam} \mathbf{B}$ comes from the second order correction).

Let us now represent the affine map $D^l_k$ as a composition of the diagonal part $\Lambda$ and the the sheer part $T$, see (2.6). The diagonal map $\Lambda$ preserves the horizontal thickness, so the thickness is only effected by the sheer 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 maps $\Psi^l_k$ do not distort the scaling ratios at all as a consequence of the precise defintion of scaling ratios. The piece $\tilde{\mathbf{B}}$ is the image under $F_k$ of $B_{\text{aff}} = \Psi^l_k(\mathbf{B})$. This map is exponentially close to the degenerate renormalization fixed point. It will not distort the scaling ratios too much. □

Starting with pieces obtained during the one-dimensional regime, we apply repeatedly the maps $G_k$ as long as the new pieces are not too deep. 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 ticknesses 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 exponential 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 $\mathcal{B}^n$ generated by combining these three regimes by $\mathcal{P}_n$. These pieces are $\theta^n$-universal.

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

**Theorem 2.3.** *The geometry of $O$ is probabilistically universal.*

*Proof.* Let $n \geq 1$. The pieces in $\mathcal{P}_n$ are $\theta^n$-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 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 $\mathcal{B}^n$ except for an exponential small fraction. Let $R_n$ be the corresponding set of paths of the random walk. These are the paths which hit the interval

$[(1-q_1)\cdot n, \leq (1-q_0)\cdot n]$ at least once but are not necessarily $s$-controlled. So

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

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

$$\nu(K_{\kappa(n)}) = 1 - O(\sum_{k=\kappa(n)}^{\infty} \frac{1}{2^{s(k)}}) = 1 - O\left(\frac{1}{2^{a2^{\kappa(n)}}}\right) = 1 - O(\rho^n)$$

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

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

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

for some $\theta \in (0,1)$. □

## 3. Preliminaries

A complete discussion of the following definitions and statements can be found in part I and part II, see [CLM], [LM1], 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 $\mathcal{H}_\Omega(\overline{\epsilon})$ consists of maps $F : [-1,1]^2 \to [-1,1]^2$ of the following form.

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

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

A unimodal map $f : [-1,1] \to [-1,1]$ with critical point $c \in [-1,1]$ is *renormalizable* if $f^2 : [f^2(c), f^4(c)] \to [f^2(c), f^4(c)]$ 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 $f_* : [-1,1] \to [-1,1]$. The introduction of [FMP] 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 $F^2 : D \to D$. The construction of the domain $D$ is inspired by renormalization of unimodal maps. In particular, it should be a topological construction. However, for small $\overline{\epsilon} > 0$ the actual domain $A \subset [-1,1]$, used to renormalize as was done in [CLM], has an analytical definition. The precise definition can be found in §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 $A \subset [-1,1]$, an essentially vertical strip, is bounded by two curves of the form

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

These curves are graphs over the $y$-axis with a slope of the order $\overline{\epsilon} > 0$. 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

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

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 $s$-dilation $\Lambda$, 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 $\mathcal{H}^n_\Omega(\overline{\epsilon}) \subset \mathcal{H}_\Omega(\overline{\epsilon})$. If $F \in \mathcal{H}^n_\Omega(\overline{\epsilon})$ we use the notation

$$F_n = R^n F.$$

The set of infinitely renormalizable maps is denoted by

$$\mathcal{I}_\Omega(\overline{\epsilon}) = \bigcap_{n\geq 1} \mathcal{H}^n_\Omega(\overline{\epsilon}).$$

The renormalization operator acting on $\mathcal{H}^1_\Omega(\overline{\epsilon})$, $\overline{\epsilon} > 0$ small enough, has a unique fixed point $F_* \in \mathcal{I}_\Omega(\overline{\epsilon})$. 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,

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

If we want to emphasize that some set, say $A$, 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

$$\psi_v^{k+1} = (\Lambda_k \circ H_k)^{-1} : \mathrm{Dom}(F_{k+1}) \to A(F_k). \tag{3.1}$$

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

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

where

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

Notice, that each $A_k \subset [-1,1]$ is of full height and $\Psi_0^k$ conjugates $R^kF$ to $F^{2^k}|A_k$. Furthermore, $A_{k+1} \subset A_k$.

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 $\psi_v^2$ and $\psi_c^2$ 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

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

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 [CLM, Lemma 5.1].

**Lemma 3.1.** *Let $F \in \mathcal{I}^c_\Omega(\bar{\varepsilon})$. There exist $C > 0$ such that for $w \in W^n$, $\|D\psi^n_w\| \le C\sigma^n$, $n \ge 1$.*

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

$$B^n_\omega = \operatorname{Im}\psi^n_\omega.$$

The first return maps to the domains

$$B^n_{v^n} = \operatorname{Im}\Psi^n_0 = \operatorname{Im}\psi^n_{v^n}$$

correspond to the different renormalizations. Notice,

$$B^{n+1}_{v^{n+1}} \subset B^n_{v^n}.$$

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

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

The dynamics on this Cantor set is conjugate to an adding machine. Its unique invariant measure is denoted by $\mu$. The *average Jacobian*

$$b_F = \exp \int \log \operatorname{Jac} F d\mu$$

with respect to $\mu$ is an important parameter that influences the geometry of $\mathcal{O}_F$, see [CLM, 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*

$$\{\tau_F\} = \bigcap_{n\ge 1} B^n_{v^n}.$$

3.1. **One-dimensional maps.** Recall that $f_*\colon [-1,1] \to [-1,1]$ 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 $\mathcal{C}_n$, $n \ge 1$. 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 $j = 0, 1, 2, \cdots, 2^n - 1$. 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 $\phi : I \to \phi(I) \subset \mathbb{R}$, $I \subset \mathbb{R}$, is

$$\eta_\phi = D \ln D\phi. \tag{3.3}$$

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

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

If $\eta$ is the nonlinearity of $\phi$ then

$$\mathrm{Dist}(\phi) \le \|\eta\|_{C^0} \cdot |I|. \tag{3.4}$$

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 $n \ge 1$ the following holds. Let $\zeta_j \in T_j(n)$ then*

$$\prod_{l=k_1}^{k_2-1} \frac{|Df_*(\zeta_j)|}{\frac{|I^*_{j+1}(n)|}{|I^*_j(n)|}} \le D.$$

*with $0 \le k_1 < k_2 < 2^n$.*

*Proof.* The a priori bounds on the cycles $\mathcal{C}_n$ are described in [MS], see also [CMMT]. 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

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

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

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

Let $\eta_j(n)$ be the nonlinearity, see (3.3), of the rescaling of $f_* : I^*_j(n) \to I^*_{j+1}(n)$. The rescaling turns domain and range into $[-1, 1]$. Lemma 3.1 in [Ma2] says that

$$\|\eta_j(n+1)\|_{C_0} \le \frac{|T_j(n+1)|}{|T_j(n)|} \cdot \|\eta_j(n)\|_{C_0},$$

$$\|\eta_{j+2^{n+1}}(n+1)\|_{C^0} \le \frac{|T_{j+2^{n+1}}(n+1)|}{|T_j(n)|} \cdot \|\eta_j(n)\|_{C_0}.$$

Hence,

$$\|\eta_j(n+1)\|_{C_0} + \|\eta_{j+2^{n+1}}(n+1)\|_{C^0} \le (1-r)\cdot\|\eta_j(n)\|_{C_0},$$

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

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

Inductively, this gives a universal bound

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

Use (3.4) and observe,

$$\log \frac{|Df_*(\zeta_j)|}{\frac{|I^*_{j+1}(n)|}{|I^*_j(n)|}} = O(\|\eta_j(n)\|_{C_0}).$$

The Lemma follows. □

**Proposition 3.3.** *There exists* $\rho < 1$ *such that the following holds. Let* $0 < q_0$ *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 to* $U_k$. *Then for every* $j \le t_I$

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

*Proof.* Let $s_I \ge t_I$ 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 $f_*^{s_I} : J_0 \to U_{(1-q_0)\cdot n}$ diffeomorphically, [Ma1]. 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 $f_* : J_k \to J_{k+1}$ which has the unit interval as its domain and range, is denoted by $\eta_k$. As in the proof of Lemma 3.2 we obtain

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

The nonlinearity of the rescaled version of the map $f_* : I_k \to I_{k+1}$ which has the unit interval as its domain and range, is denoted by $\eta^I_k$. Lemma 3.1 in [Ma2] says that the nonlinearity of the restriction $f_* : I_k \to I_{k+1}$ of $f_* : J_k \to J_{k+1}$ satisfies

$$\|\eta^I_k\|_{C^0} \le \frac{|I_k|}{|J_k|}\cdot\|\eta_k\|_{C^0}.$$

Hence,

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

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

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

This finishes the proof of the Lemma. □

3.2. **Geometrical 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 $f_n \to f_*$ 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 $k \ge 1$ large enough*

$$\frac{1}{d_1} \le |\frac{\partial F_k}{\partial x}(z)| \le d_1,$$

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

Let $\tau_n$ be the tip of $F_n = R^n F$ and $\tau^*$ be the tip of $F_*$.

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

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

*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 $k \ge n$

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

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

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 [CLM] gives that $(\psi_w^{k+1})^* = \psi_w^*$ are contractions, $|D\psi_w^*| \le \sigma < 1$. Then for $k \ge n$

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

Then for every $m > n$ we have

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

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

3.3. **Analytical properties of the coordinate changes.** Fix an infinitely renormalizable Hénon-like map $F \in \mathcal{I}_\Omega(\bar\varepsilon)$ to which we can apply the results of [CLM] and [LM1], $\bar\varepsilon > 0$ is small enough. For such an $F$, we have a well defined *tip*:

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

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 $\Psi_k$ at 0 by $D_k$ and decompose it into the unipotent and diagonal factors:

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

Let us factor this derivative out from $\Psi_k$:

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

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

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

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

(3) $|\partial^2_{xx} s_k| = O(1), \quad |\partial^2_{xy} s_k| = O(\bar\varepsilon^{2^k}), \quad |\partial^2_{yy} s_k| = O(\bar\varepsilon^{2^k})$.

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

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

We will use the following notation

$$B^{n-k}_{v^{n-k}}(F_k) = \operatorname{Im} \Psi^n_k.$$

Consider the derivatives of the maps $\Psi^n_k$ at the origin:

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

We can reshuffle this composition and obtain:

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

Factoring the derivatives $D^n_k$ out from $\Psi^n_k$, we obtain:

$$\Psi^n_k = D^n_k \circ (\operatorname{id} + \mathbf{S}^n_k), \tag{3.7}$$

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

The following Lemma is Lemma 7.6 in [CLM]

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

(1) $|\partial_x S^n_k| = O(1), \quad |\partial_y S^n_k| = O(\bar\varepsilon^{2^k})$;
(2) $|\partial^2_{xx} S^n_k| = O(1), \quad |\partial^2_{yy} S^n_k| = O(\bar\varepsilon^{2^k}), \quad |\partial^2_{xy} S^n_k| = O(\bar\varepsilon^{2^k}\, \sigma^{n-k})$.

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

$$\frac{1}{d_0} \leq \left|\frac{\partial(\operatorname{id} + \mathbf{S}^n_k)}{\partial x}\right| \leq d_0$$

*Proof.* According to proposition 7.8 in [CLM], the diffeomorphisms $\operatorname{id} + \mathbf{S}^n_k$ stay wihin a compact family of diffeomorphisms. This gives the upperbound on the derivative. The partial derivative $\frac{\partial(\operatorname{id} + \mathbf{S}^n_k)}{\partial x}$ 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. □

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

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

*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. □

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

Let $\subset \mathbb{R}^2$ 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 tip piece $B^k \equiv B^k_{v^k}$ of level $k \in \mathbb{N}$ contains two pieces of level $k+1$, the tip one, $B^{k+1}$, and the lateral one,

$$E^k = B^{k+1}_{v^k c}.$$

They are illustrated in Figure 2.1, and more schematically, in Figure 4.1.

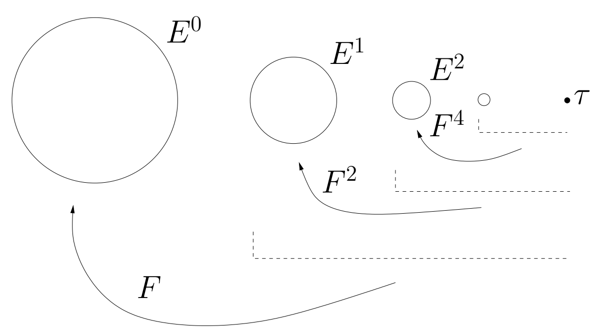


Figure 4.1

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

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

We call $k$ the *depth* of any piece $B \in \mathcal{B}^n[k]$. A piece $B^n_\omega$ 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 $\mu$ is the invariant measure on $\mathcal{O}_F$. Let

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

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

$$\mathbf{B} = B^{n-k}_{c\omega_{k+2}\cdots\omega_n}(F_k) \in \mathcal{B}^{n-k}(F_k)[0], \quad \text{so } B = \Psi^k_0(\mathbf{B}), \tag{4.1}$$

see Figure 4.2.

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

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

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

$$\frac{1}{C_0} \leq \operatorname{mod} \mathbf{B} \leq C_0, \tag{4.2}$$

with $C_0 = 3d_1$, where $d_1 > 0$ is the bound on $\partial F_k/\partial x$ from Corollary 3.5 (see Figure 4.2).

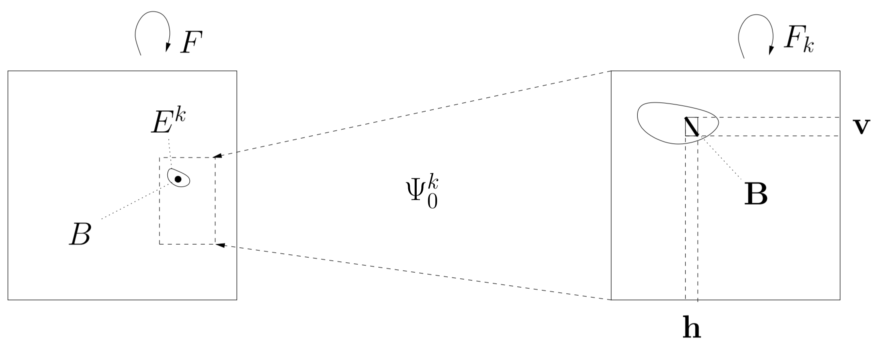


FIGURE 4.2. A regular piece

Notice that in the degenerate case, $F(x, y) = (f(x), x)$, 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 (2.7). 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 $\alpha = \alpha(\bar{\varepsilon}) > 0$ depending only on the geometry of $F$ will be made in the cause of the paper (see Propositions 4.1, 5.1, etc.).

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

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

**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

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

Observe: $\tilde{\mathbf{B}} = F_k \circ \Psi_k^l(\mathbf{B})$ (see Figure 4.3).

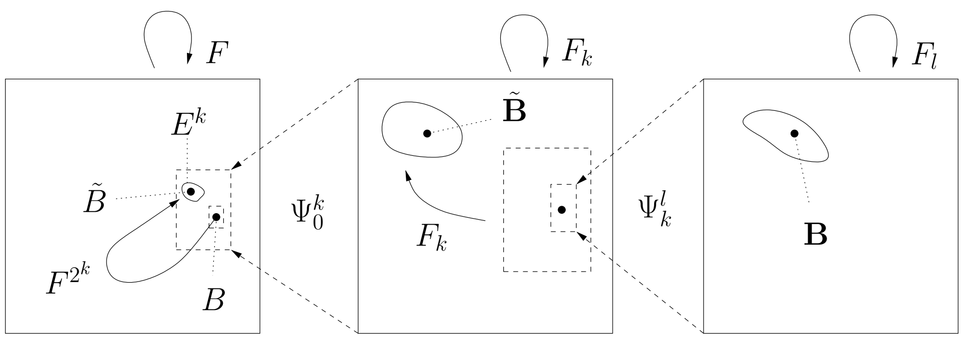


FIGURE 4.3. Pieces $B$ and $\tilde{B}$ viewed from appropriate scales.

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) hold for $\operatorname{mod} \mathbf{B} = \mathbf{h}/\mathbf{v}$. We want to show that the same bound hold for $\operatorname{mod} \tilde{\mathbf{B}} = \tilde{\mathbf{h}}/\tilde{\mathbf{v}}$.

The map $\Psi_k^l$ factors into a non-linear and an affine part as described in §3:

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

Figure 4.4 shows details of this factorization for the map $\Psi_k^l$ from Figure 4.3. Let $h_{\text{diff}}$ and $v_{\text{diff}}$ be the sizes of the rectangle $Q_{\text{diff}}$ associated with the piece $(\operatorname{id} + \mathbf{S}_k^l)(\mathbf{B})$, see Figure 4.4. Lemmas 3.9(1) and 3.10 imply for $k$ big enough:

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

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

$$h_{\text{diff}} \ge \frac{1}{2d_0} \mathbf{h}. \tag{4.5}$$

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

$$v_{\text{diff}} = \mathbf{v} \le C_0 \cdot \mathbf{h}. \tag{4.6}$$

Let $h_{\text{aff}}$ and $v_{\text{aff}}$ be the sizes of the rectangle $Q_{\text{aff}}$ associated with the piece $B_{\text{aff}} = \Psi_k^l(\mathbf{B}) = D_k^l \circ (\operatorname{id} + \mathbf{S}_k^l)(\mathbf{B})$ (which is the piece $B$ viewed

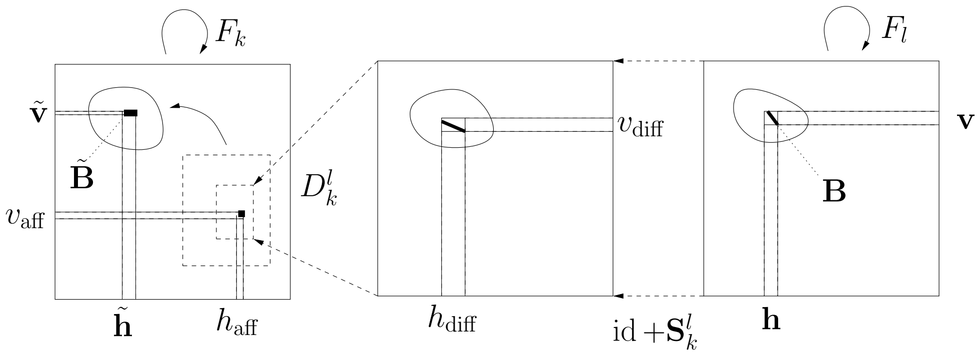


FIGURE 4.4. Factorization of the map $\Psi_k^l$ into horizontal and affine parts.

from scale $k$). Incorporating the above estimates into decomposition (3.6) and using Lemma 3.8, we obtain for large $k$ (with $s = l - k$) :

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

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

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

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

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

as long as $\alpha$ is small enough (depending on the geometry of $F_k$).

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

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

Thus,

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

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 $\tilde{\mathbf{B}}$:

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

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

$$\operatorname{mod}\tilde{\mathbf{B}} \le 2d_1 + \frac{A_0 b^{2^k}}{\operatorname{mod}\Psi^l_k(\mathbf{B})} \le 2d_1 + \frac{A_0 b^{2^k}}{\sigma^s \operatorname{mod}\mathbf{B}} \le 2d_1 + A_0 C_0 \alpha \le 3d_1,$$

as long as $\alpha$ is small enough.

*Remark* 4.2. Notice that $\operatorname{mod}\mathbf{B}$ 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 $\operatorname{mod}\tilde{\mathbf{B}}$ as good as that for $\mathbf{B}$.

The lower estimate, $\operatorname{mod}\tilde{\mathbf{B}} \ge (3d_1)^{-1}$, is similar. □

## 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 Figure 4.2. In the degenerate case, most of the pieces $\mathbf{B} \cap \mathcal{O}_{F_l}$ get squeezed in a narrow strip around the diagonal of the associated rectangle $\mathbf{Q}$. 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 $\delta$:

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

(oriented "north-west" and "north-east" respectively.)

Given a piece $B \in \mathcal{B}^n$ and the associated rectangle $Q = Q(B)$, let $L : [0,1]^2 \to Q$ be the diagonal 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 $\delta_B$ 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 $\mathbf{\Delta} \equiv \Delta_{\mathbf{B}}$ and call it the *regular stick* associated with $B$, see Figure 5.1.

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

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

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

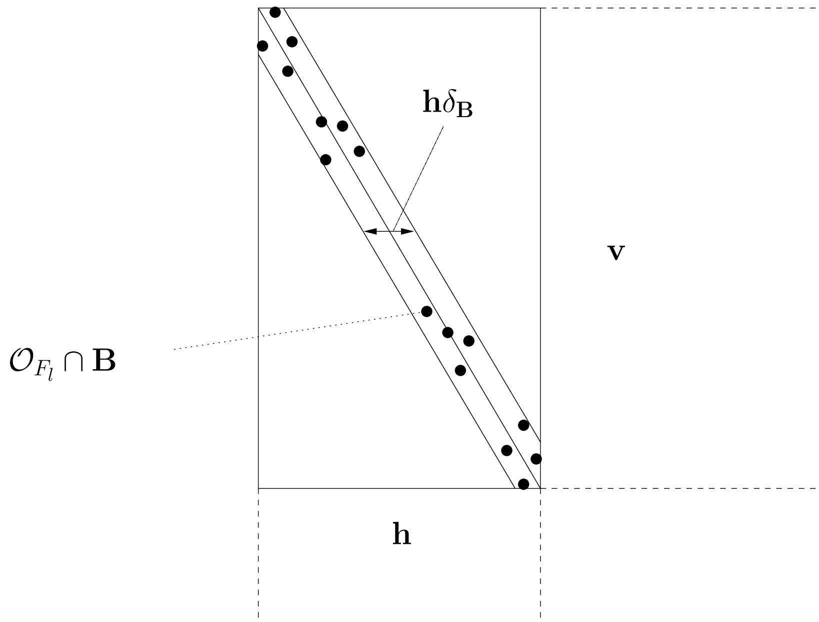


FIGURE 5.1. Regular stick

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

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

Indeed, let $\Gamma_y$ be the horizontal section of $(\text{id} + \mathbf{S}^l_k)(\Delta_{\mathbf{B}})$ on height $y$, and let $\boldsymbol{\Gamma}_y = (\text{id} + \mathbf{S}^l_k)^{-1}(\Gamma_y)$. Then

$$|\Gamma_y| \leq |\boldsymbol{\Gamma}_y| \cdot \| \text{id} + \mathbf{S}^l_k \|_{C^1} = O(\mathbf{w}),$$

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

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

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

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:

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

Let us now consider the piece $B_{\mathrm{aff}} \equiv \Psi_k^l(\mathbf{B}) = D_k^l(B_{\mathrm{diff}})$, see Figure 4.4. Let $D_k^l = T \circ \Lambda$, where $\Lambda = \Lambda_k^l$ and $T = T_k^l$ are the diagonal and sheer parts of $D_k^l$ respectively, see (3.6). Let us consider a box $B_{\mathrm{diag}} = \Lambda(B_{\mathrm{diff}})$, and let $h_{\mathrm{diag}} = \sigma^{2(l-k)} h_{\mathrm{diff}}$ and $v_{\mathrm{diag}} = \sigma^{l-k} v_{\mathrm{diff}}$ be its horizontal and vertical sizes. Since diagonal affine maps preserve the horizontal thickness, the thickness is only effected by the sheer part $T_k^l$, which has order $t_k \asymp b^{2^k}$, see Lemma 3.8, namely:

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

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* $u_{\mathrm{aff}}$ of $B_{\mathrm{aff}}$, i.e., the vertical size of the stick $\Delta(B_{\mathrm{aff}})$. From triangle similarity, we have:

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

So

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

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 $\|\varepsilon_k\|_{C^2} = O(2^{b^k})$, see Theorem 3.4. Let $\tilde{\mathbf{w}}$ be the absolute thickness of $\tilde{\mathbf{B}}$. By (4.7)–(4.8), the rectangle $Q_{\mathrm{aff}}$ associated with $B_{\mathrm{aff}}$ has sizes

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

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

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

where $x_0, y_0$ is its corner where the stick $\Delta_{\mathrm{aff}}$ 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:

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

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

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

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

*Remark* 5.1. This was the moment where the thickness improves.

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

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

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

## 6. Scaling

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

$$B = B^n_{\omega\nu}\subset\hat{B} = B^{n-1}_{\omega}\subset E^k.$$

Let $\mathbf{B}$ and $\hat{\mathbf{B}}$ be the corresponding rescaled pieces, so $B = \Psi^k_0(\mathbf{B})$ and $\hat{B} = \Psi^k_0(\hat{\mathbf{B}})$. 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.

The *scaling number* of $B$ is

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

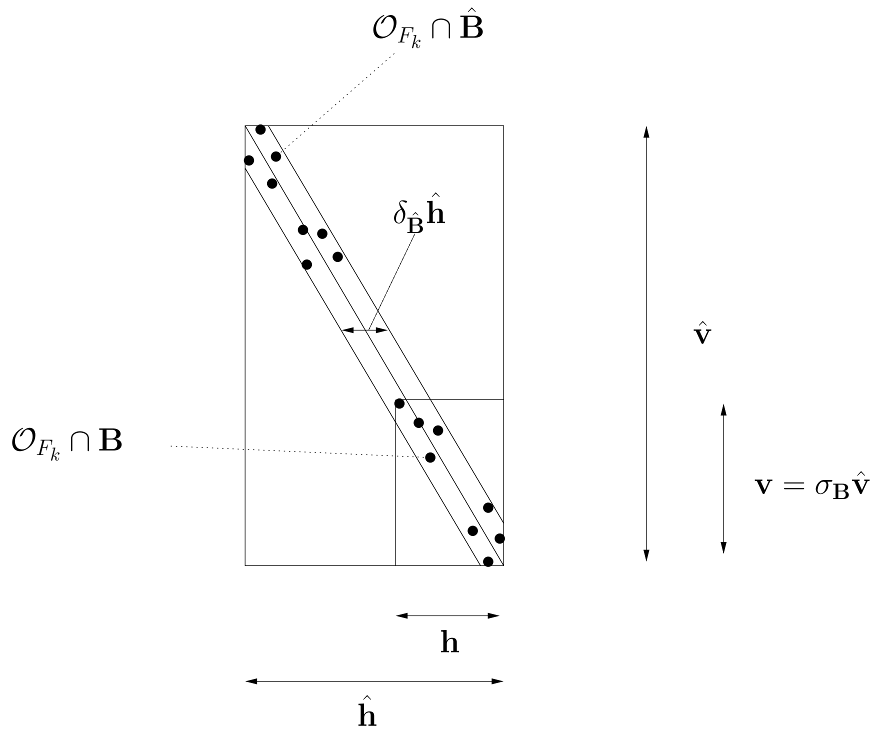


FIGURE 6.1

*Remark* 6.1. The scaling number can be expressed directly in terms of the original pieces $B$ and $\hat{B}$. Indeed, since the diffeomorphism $\Psi_0^k$ 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 $\sigma_B$ 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 $B = B^n_{\omega\nu}(F)$ as above, let $B^* = B^n_{\omega\nu}(F_*)$ be the corresponding degenerate piece for the renormalization fixed point $F_*$. The *proper scaling* for $B$ is

$$\sigma^*_{\mathbf{B}} = \sigma_{B^n_{\omega\nu}(F_*)}.$$

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 [BMT].

*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 $\sigma_{\mathbf{B}}$ are vertical measurements of pieces. Using that Hénon like maps take vertical lines to horizontal lines, $y' = x$, we have

$$\sigma^*_{\mathbf{B}} = \frac{|I^*_{i-1}(n+1)|}{|I^*_{\hat{i}-1}(n)|}.$$

**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*

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

*where $\tilde{B} = G_k(B) \in \mathcal{B}^n[k]$ and $B \subset \hat{B} = \Psi^l_0(\hat{\mathbf{B}}) \in \mathcal{B}^{n-1}[l]$.*

*Proof.* As above in §4, let $h_{\text{aff}}$ stand for the horizontal length of $B_{\text{aff}} = \Psi^l_k(\mathbf{B})$, see Figure 4.4. We will use the similar notation $\hat{h}_{\text{aff}}$ and $\hat{w}_{\text{aff}}$ for the corresponding measurements of the piece $\hat{B}_{\text{aff}} := \Psi^l_k(\hat{\mathbf{B}})$.

Since $F_k$ maps vertical lines to horizontal lines, we have

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

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

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

Now Figure 6.2 shows:

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

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). □

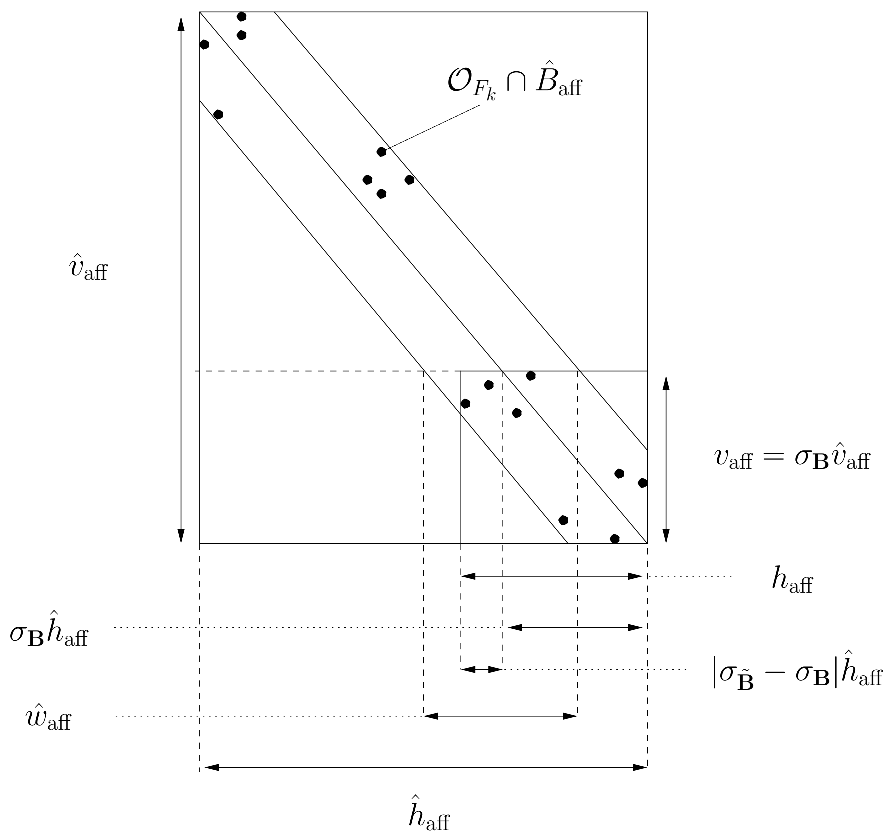


FIGURE 6.2

## 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 $\sigma_{B_1}$, and $\sigma_{B_2}$ are called *scaling factors* of $B_1$ and $B_2$.

*Remark* 7.1. The scaling factor $\sigma_{\mathbf{B}}$ of a piece $B$ is a measurement of the corresponding $\mathbf{B}$. The scaling factor $\sigma_B$ of $B$ reveres to measurements of the actual piece in the domain of $F$. The difference between the scaling factors $\sigma_B$ and $\sigma_{\mathbf{B}}$ is estimated in Proposition 7.7.

We say that $B$ is *$\epsilon$-universal* if

$$|\sigma_{B_1} - \sigma^*_{\mathbf{B}_1}| \le \epsilon, \quad |\sigma_{B_2} - \sigma^*_{\mathbf{B}_2}| \le \epsilon, \quad \text{and} \quad \delta \le \epsilon.$$

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

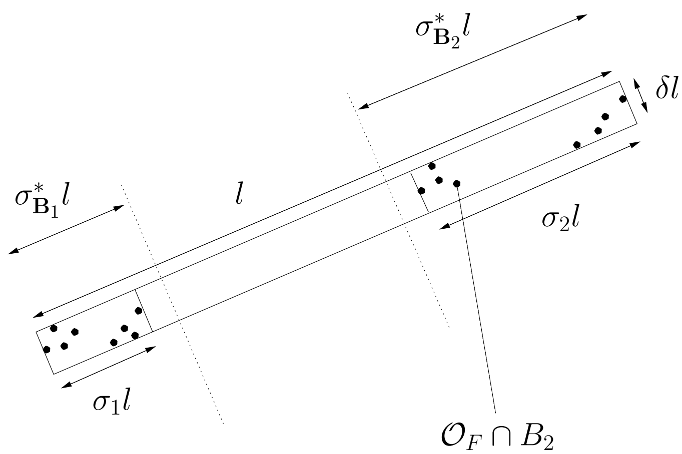


FIGURE 7.1

**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

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

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 $(1-q_1)\cdot n \le k \le (1-q_0)\cdot n$ 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 $0 \le k < (1-q_1) \cdot n$. 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 $k \ge \kappa(n) \asymp \ln n$, see Proposition 8.2 . That is why the pushing-up regime is restricted to $\kappa(n) \le k < (1-q_1) \cdot n$ 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 $2^{\kappa(n)} \le Kn \ln 1/\theta$. 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 $\theta^n$, 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 $r \gtrsim 1/b >> 1$. 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 $l(\kappa_0(n)) \ge n$. As long as $\kappa_0(n) \le k < (1-q_1) \cdot n$ 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 $\kappa(n) \le k < \kappa_0(n)$ we have $l(k) < (1-q_1) \cdot n$ which means that we can not push up all previously recovered pieces in $B^n[l]$ 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 $0 < q_0 < q_1 \le q^*$ there exists $n^* > 0$ such that for $n \ge n^*$ and $(1 - q_1) \cdot n \le k \le n$*

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 $(1 - q_1) \cdot n \le k \le (1 - q_0) \cdot n$*
$$\delta_{\mathbf{B}} = O(\rho^{q_0 \cdot n}),$$
*where $B = \Psi_0^k(\mathbf{B})$.*

Choose, $(1 - q_1) \cdot n \le k \le n$ 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 $\tau_n$ be the tip of $F_n$ and $\tau_*$ the tip of $F_*$. In the next part we will have to compare the maps $\Psi_k^n$ related to $F$ and the maps $(\Psi_k^n)^*$ corresponding to $F_*$. Let

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

and

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

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 $0 \le j < 2^{n-k}$ and $j$ is odd. Let $\mathbf{B}_j = F_k^j(\mathbf{B}_0)$ and $\mathbf{B}^*_j = F_*^j(\mathbf{B}^*_0)$ for $0 \le j < 2^{n-k}$. 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 $n^{th}$ cycle of $f_*$ are denoted by $I^*_j(n)$, see §3.1. Observe,

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

and

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

Consider the conjugations

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

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 $0 \le j < 2^{n-k}$. 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 $0 < q \le q^*$ there exists $n^* > 0$ such that for $n \ge n^*$, $(1-q)\cdot n \le k \le n$, and $0 \le j < 2^{n-k}$*

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

*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 $\Delta x_j$ and $\Delta y_0$. Since, $F_k \to F_*$ 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

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

where we used the *a priori* bounds: $|I_{j+1}^*| \ge \rho_0^{n-k}$ 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

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

Thus,

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

and

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

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

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

and

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

One has to be careful when applying (7.1) repeatedly. The points $\zeta_j$ 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

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

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

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

This estimate finishes the induction step. □

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

*Proof of Proposition 7.2.* Let $(1-q_1)\cdot n \le k \le n$ 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 $(1-q_1)\cdot n \le k \le (1-q_0)\cdot n$. 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 $\sigma^{n-k} \le \sigma^{q_0 \cdot n}$, 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). □

## 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 $B^{(i)}$, $i = 0, 1, 2, \cdots, t$, are the predecessors of $B = B^{(0)}$ with

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

then

1. $k_{i+1} \le l(k_i)$, $i = 0, 1, 2, 3, \ldots, t-1$,
2. there exists $0 \le s \le t$ such that $(1-q_1)\cdot n \le k_s(B) \le (1-q_0)\cdot n$, 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, [LM1]. 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

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

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*

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

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

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

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

$i = 0, 1, 2, \ldots, t$. Let $k_i = k_i(\hat{B}) = k_i(B)$ and $\delta_i$ 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 $B^{(i)}\Psi_0^{k_i}(\mathbf{B}^{(i)})$, $i = 0, 1, 2, \ldots, t$. Observe, the piece $B$ might have one predecessor more than $\hat{B}$.

Apply Propositions 5.1 and 6.1. In particular,

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

and

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

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

Iterating estimate (7.7) we obtain

$$\sum_{i=0}^{s}\delta_i \le 2\delta_s + O(\sigma^{n-k_s}) = O(\rho^{q_0\cdot n}) + O(\sigma^{q_0\cdot n}), \tag{7.9}$$

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:

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

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),

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

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 $(1-q_1)\cdot n \le k_s \le n$ and we can apply Proposition 7.2(2),

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

□

The measurements of the pieces, such as scaling and thickness, are geometrical quantities observed when viewing a piece from its scale, they are geometrical measurements of **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 $\boldsymbol{\delta} > 0$ be the relative thickness of $\hat{\mathbf{B}}$, the absolute thickness of $\hat{\mathbf{B}}$ is $\mathbf{w} = \boldsymbol{\delta} \cdot \mathbf{h}$. 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 :

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

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 $\Psi_0^k$ into its diffeomorphic part $(\mathrm{id} + \mathbf{S}_0^k)$ 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 $(\mathrm{id} + \mathbf{S}_0^k)$ 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

$$\sigma_{\text{diff}} = \sigma_{\mathbf{B}},$$
$$v_{\text{diff}} = \mathbf{v}$$

and, by recalling (5.1),

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

The last two estimates rely on $\mathbf{v} \asymp \mathbf{h}$. The term $\mathbf{h} \cdot \sigma^{n-k}$ reflects the distortion of $(\mathrm{id} + \mathbf{S}_0^k)$ on $\hat{\mathbf{B}}$ determined by the diameter of $\hat{\mathbf{B}}$ which is of the order $\sigma^{n-k}$. The next step is to apply the affine part of $\Psi_0^k$. 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

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

Use the above estimates in the following

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

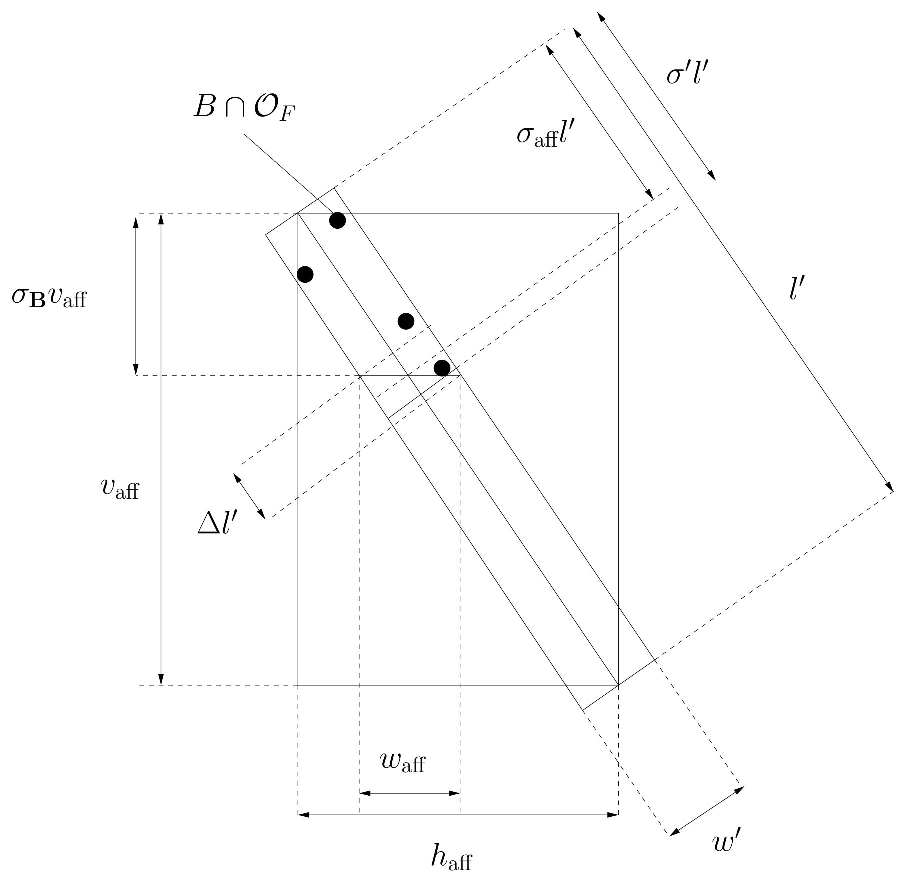


FIGURE 7.2

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 Figure 7.2. 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

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

and

$$w' \le w_{\text{aff}}. \tag{7.14}$$

First we will estimate the precision of $\sigma'$. Let $\gamma$ 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 7.2. The projection $\Delta l'$ of the horizontal interval of length $w_{\text{aff}}$ onto the diagonal has length

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

Observe,

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

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

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

Use (7.10), (7.11), and (7.15) to estimate the precision of $\sigma'$

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

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

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

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.**

**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 $n^* \ge 1$ such that if for $0 \le j < 2^{(1-q_1)\cdot n}$*

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

*with* $B \in \mathcal{B}^n[k]$, $(1-q_1)\cdot n \le k \le (1-q_0)\cdot n$, *and* $n \ge n^*$, *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 $(1-q_1)\cdot n \le k \le (1-q_0)\cdot n$ 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 $\hat{B}_j = F^j(\hat{B})$, $j \le 2^{(1-q_1)\cdot n}$. Assume that for some $j \le 2^{(1-q_1)\cdot n}$

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

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

Apply $F$ to this rectangle. 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

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

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 $\sigma^n$, see Lemma 3.1. We can estimate the sizes $l', w'$ and $\sigma'$ 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

$$\begin{aligned} l' &\ge Dl - 2Kl^2 - 2Mw, \\ w' &\le Mw + 2Kl^2, \end{aligned}$$

Observe,

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

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

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

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

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

Use remark 6.3 and apply Proposition 3.3 to get

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

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

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

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

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

## 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 $k^* \ge 1$ large enough to allow the use of the Propositions 7.7, and 7.8.

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

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

For $n \ge 1$ large enough we have

$$\kappa_0(n) \le \frac{\ln n}{\ln 2}. \tag{8.1}$$

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

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

*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

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

and

$$\beta_n(k; q_0, q_1) = 0, \quad (1 - q_1) \cdot n \le k \le (1 - q_0) \cdot n, \tag{8.3}$$

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

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

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

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

□

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

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

*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

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

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

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

Hence,

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

Use (8.1) and observe,

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

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

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

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

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

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

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

$$2^{\kappa(n)} \le Kn \ln 1/\theta.$$

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

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

*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

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

In turn, this holds when

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

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

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

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

**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

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

when $\theta < 1$ close enough to 1 and $n \geq 1$ large enough (Recall that $q_0$ depends on $\theta$). Now consider an image $F^j(B)$ with $j \leq 2^{\kappa(n)} - 1 < 2^{(1-q_1)\cdot n}$. Denote its precision by $\epsilon_j$. 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

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

Choose $Q > 0$ large enough such that

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

This choice implies

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

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 $\theta^n$-universal. □

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 $\theta^* \le \theta < 1$ there exists $n^* \ge 1$ such that for $n \ge n^*$*

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

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

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

Hence, for $n \ge 1$ large enough

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

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

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

□

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 $\theta^n$-sticks can be determined by pure combinatorial methods. In [CLM], it has been shown that there are pieces which are not contained in $\theta^n$-sticks. Probabilistic universality says that these *bad* spots will be filled on deeper levels with pieces contained in sticks with exponential precision. 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

$$B = \operatorname{Im} \psi^1_{w_1} \circ \psi^2_{w_2} \circ \dots \psi^n_{w_n}$$

where $\psi^k_v$ is the non-affine rescaling used to renormalize $R^k F$, and to obtain $R^{k+1} F$ and $\psi^k_c = R^k F \circ \psi^k_v$. 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

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

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 exits $n - q_1 \cdot n \leq k \leq n - q_0 \cdot n$ 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 [LM1]. The characterization of the pieces in $\mathcal{P}_n$ is purely topological.

**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,

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

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 □

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 $h : \mathcal{O}_{F_1} \to \mathcal{O}_{F_2}$ then $b_{F_1} = b_{F_2}$, see [LM1], and $h(C_{F_1}) = C_{F_2}$. The set of controlled points is a topological invariant.

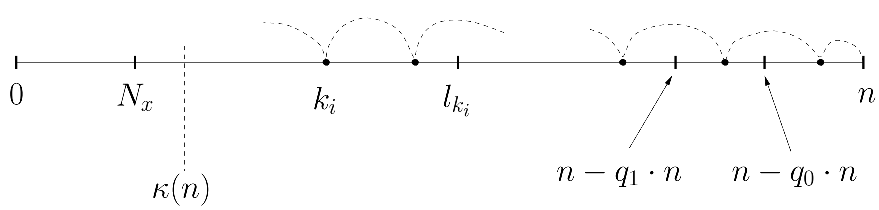


FIGURE 9.1

## 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

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

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 [CLM] we known 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 $n \geq N$

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

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 $\theta^n$-stick see Figure 10.1. 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 $\sigma_1$ 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

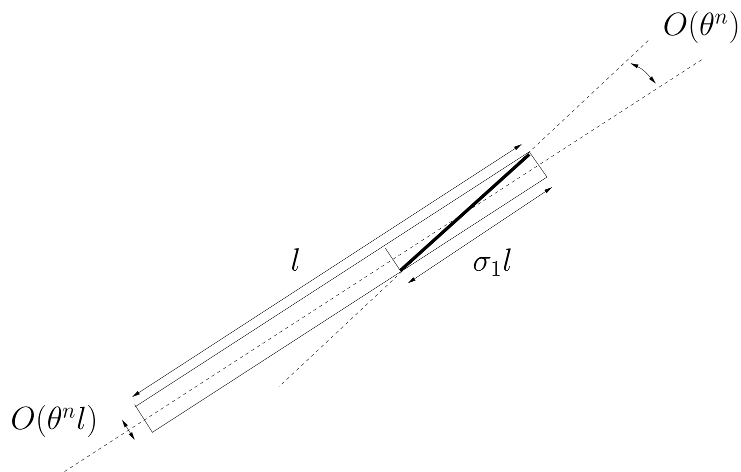


FIGURE 10.1

angle between $T_n(x)$ and $T_{n+1}(x)$ is of the order $\theta^n$, see Figure 10.1. The piecewise constant functions $T_n$ form a Cauchy sequence,

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

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

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

The construction implies that we get in fact a map

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

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

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

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

The tangent line field of $\mathcal{O}_F$ is *weakly* $\beta$*-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 $X_1 \subset X_2 \subset X_3 \subset \cdots$ 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* $\beta$*-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 $\beta-$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. □

In [CLM] 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

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

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 Figure 10.2, 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

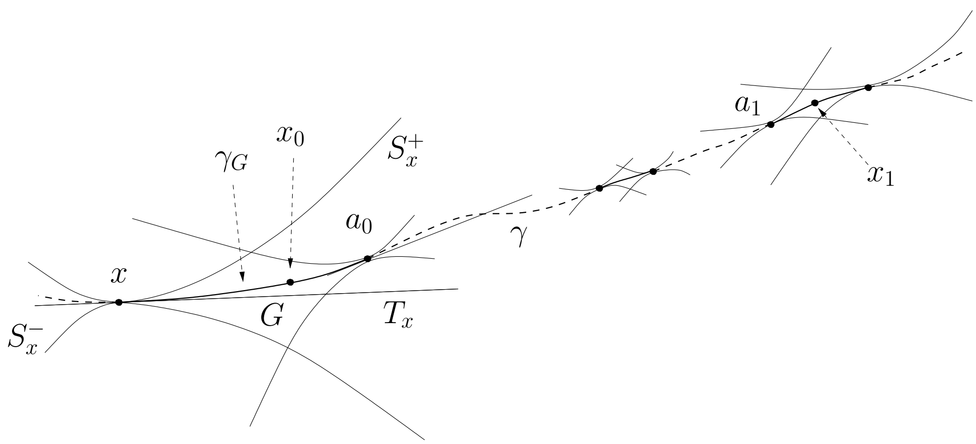


FIGURE 10.2

over the tangent line $T_{x_0}$ of a cubic polynomial $\gamma_G$ 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 $\gamma_G$ also by $\gamma_G$. 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 $\gamma_G$ at a point $x \in \gamma_G$ then

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

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 $\gamma_G$

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

is a $C^1$ curve.

Left is to show that the tangent direction $D\gamma$ is $C^\beta$. 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 Figure 10.2. 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

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

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

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 $F_n : [0,1]^2 \to [0,1]^2$ be the $n^{th}$-renormalization of $F$. The piece $B^1_v(F_n) \subset \mathrm{Dom}(F_n)$ is strip bounded between two horizontal line segments and $B^1_c(F_n) \subset \mathrm{Dom}(F_n)$ is strip bounded between two vertical line segments. Let $\gamma_n$ be a collection of three line segments which connects the two pieces and each piece with the horizontal boundaries of $\mathrm{Dom}(F_n) = [0,1]^2$, see Figure 10.3.

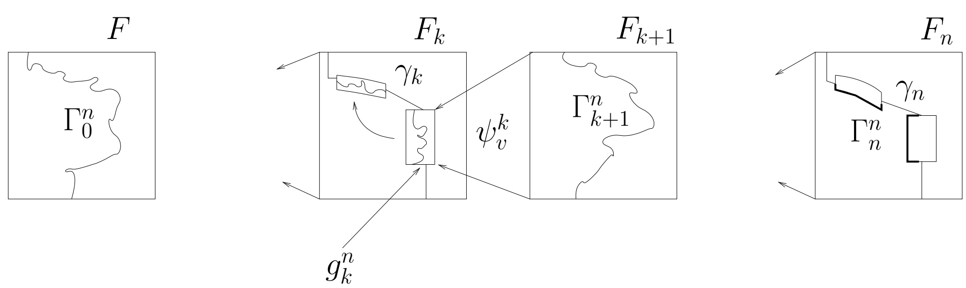


FIGURE 10.3

For each $n \geq 1$ we will construct inductively a curve $\Gamma^n$ in the domain of $F$ which passes through all pieces $B \in \mathcal{B}^n$ of the $n^{th}$-cycle of $F$. Let $\Gamma^n_n$ consists of $\gamma_n$ and curves in the boundaries of $B^1_v(F_n)$ and $B^1_c(F_n)$ connecting the end points of $\gamma_n$, see Figure 10.3.

Suppose $\Gamma^n_{k+1}$ is defined and its end point are in the two horizontal boundary part of the domain of $F_{k+1}$, see Figure 10.3. Let $\Gamma^n_k$ be the curve connecting the top and bottom of the domain of $F_k$ consists of the curves

$$\Gamma^n_k = \psi^k_v(\Gamma^n_{k+1}) \cup \psi^k_c(\Gamma^n_{k+1}) \cup \gamma_k \cup g^n_k,$$

where $g^n_k$ consists of the two shortest horizontal line segments connecting the endpoints of $\psi^k_v(\Gamma^n_{k+1})$ with the end points of $\gamma_k$ and the two vertical line segments connecting the endpoints of $\psi^k_v(\Gamma^n_{k+1})$ with the end points of $\gamma_k$, see Figure 10.3. Let $\Gamma^n = \Gamma^n_0$.

The curve $\Gamma^{n+1}$ is obtained from $\Gamma^n$ 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 $\Gamma^n$ where the parametrization of $\Gamma^{n+1}$ is obtained from the one of $\Gamma^n$ by only adjusting only inside the pieces of $\mathcal{B}^n$. In each piece $B \in \mathcal{B}^n$, the curve $\Gamma^{n+1}$ is partitioned into five sub-curves, see Figure 10.3. The refinement of the parametrization of $\Gamma^n$ spends equal time in each of these five sub-curves. The diameter of the pieces in $\mathcal{B}^n$ decay exponentially fast, $\sup_{B\in\mathcal{B}_n} \operatorname{diam}(B) = O(\sigma^n)$. The construction and this decay imply that the parametrization have a uniform Hölder bound. This bound allows us to take a limit. Let $\Gamma$ be the limiting Hölder curve. It contains $\mathcal{O}_F$.

The maps $\psi^k_v$ and $\psi^k_c$ are contracting distance by at least $\frac{1}{2.5}$, for $k \geq 1$ large enough, see Lemma 3.1. Denote the length of $\Gamma^n_k$ by $|\Gamma^n_k|$. Then,

$$\begin{aligned}|\Gamma^n_k| &\leq \frac{2}{2.5} \cdot |\Gamma^n_{k+1}| + |\gamma_k| + |g^n_k| \\ &\leq \frac{2}{2.5} \cdot |\Gamma^n_{k+1}| + 4.\end{aligned}$$

The curves $\Gamma^n_k$ have a bounded length. In particular, the limiting curve $\Gamma$ is rectifiable.

Outside the pieces $B \in \mathcal{B}^n$ the curve $\Gamma$ coincides with $\Gamma^n$ 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. □

*Remark* 10.2. The curve $\Gamma$ 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 $\Gamma$ 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 $\beta(B) \le 1$ introduced by P. Jones. In [J], 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

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

In the present case of $\mathcal{O}_F$, one can use the dynamical covers $\mathcal{B}^n$ instead of the diadic ones. Since $\operatorname{diam}(B) = O(\sigma^n)$ 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 geometrical information of the pieces: the bound $\beta(B) \le 1$ suffices. For completeness we include a direct proof for rectifiability using the strongly contracting rescalings $\psi^k_c$ and $\psi^k_v$.

The sets $X_N$ have better geometrical 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 $\pi_x : \mathbb{R}^2 \to T_x$ be the orthogonal projection.

Let $Y \subset \mathcal{O}_{F_1}$. A map $h : Y \to h(Y) \subset \mathcal{O}_{F_2}$ is differentiable at $x_0 \in Y$ if $x_0$ and $h(x_0)$ have a tangent line, and there exists a linear $Dh(x_0) : T_{x_0} \to T_{h(x_0)}$ 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 $h : X \to h(X) \subset \mathcal{O}_{F_2}$ is *almost everywhere a* $(1+\beta)$-*diffeomorphism* if for each $N \ge 1$ the restriction $h|X_N$ is differentiable at each $x \in X_N$ and

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

and its inverse are $\beta$-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 $\mu_*$. In

[LM1] it has been shown that every 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

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

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

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

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

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 $\sigma_1$ and $\sigma_2$ containing respectively $B_1$ and $B_2$ and the 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 Figure 10.4.

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 $(x_B^-, x_B^+)$.

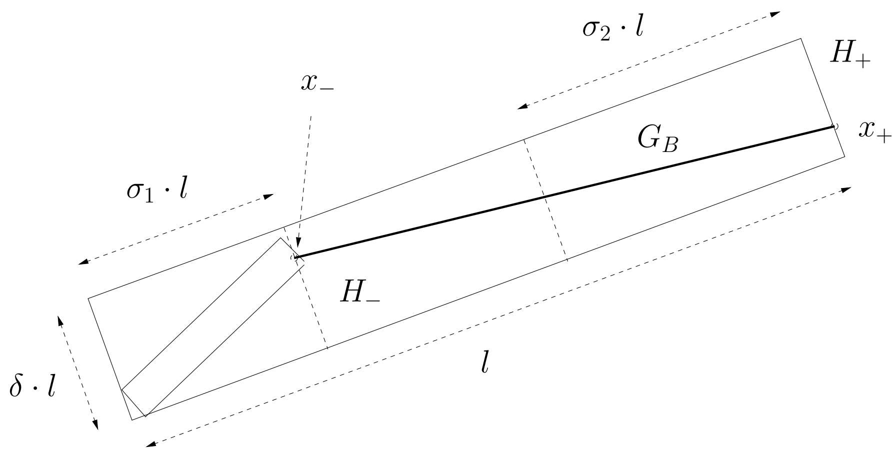


FIGURE 10.4

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 $\theta^k$. This is a consequence of $\delta = O(\theta^k)$ and the *a priori* bounds on $\sigma_1$ and $\sigma_2$.

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 $\theta^N$. 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 $\pi_x$ be the projection onto the tangent line $T_x$ of $x$. Then

$$|x - \pi_x(y)| = \sum_{x<G_{B'}<y} |\pi_x(G_{B'})|. \tag{10.5}$$

The angle between each gap $G_{B'}$ between $x$ and $y$, and the tangent line of $x$ is of order $\theta^k$, see (10.2) and remark 10.4. This implies that

$$\frac{|\pi_x(G_{B'})|}{|G_{B'}|} = 1 + O(\theta^k). \tag{10.6}$$

The Cantor set $\mathcal{O}_F$ is almost everywhere differentiable, see Theorem 10.1. In particular, use (10.4) to obtain

$$\frac{|x - \pi_x(y)|}{|x-y|} = 1 + O(\theta^k). \tag{10.7}$$

The estimates (10.5), (10.6), and (10.7) prove the Claim. □

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 $B_j \in \mathcal{X}_N(j)$, $k \le j \le l$ with $B_l \subset B_{l-1} \subset \dots B_{k+1} \subset B_k$. 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 $\tilde{B} \subset B$

$$|G_{\tilde{B}}| = |G_{\tilde{B}^*}| \cdot \frac{|G_B|}{|G_{B^*}|} \cdot (1 + O(\theta^k)).$$

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

$$|\ln\frac{|h(y)-h(x)|_g}{|y-x|_g} - \ln\frac{|h(z)-h(x)|_g}{|z-x|_g}| = O(\theta^k). \tag{10.8}$$

This implies that for $x, y \in X_N \cap B^0$ the following limit exists.

$$Dh(x) = \lim_{y\to x}\frac{|h(y)-h(x)|_g}{|y-x|_g}.$$

Moreover, the limit depends continuously on $x$.

*Claim* 10.8. There exists a universal $\beta > 0$, independent of $N$, such that $Dh : X_N \to \mathbb{R}$ is $\beta$-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,

$$|x - x_0| \geq \rho^{k-N}\cdot \operatorname{diam}(B^0)$$

where $\rho < 1$. Choose $\beta > 0$ such that $\rho^\beta = \theta$. Then

$$\theta^k = O(|x-x_0|^\beta). \tag{10.9}$$

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 $Dh$. □

We will identify $Dh(x)$ with a linear map $Dh(x) : T_x \to T_{h(x)}$. The positive function $Dh$ is bounded. This bound, (10.8), and Claim 10.5, imply that for $x, x_0 \in X_N$

$$|h(x) - h(x_0)| = O(|x-x_0|). \tag{10.10}$$

*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

$$\begin{aligned} |h(y)-h(x)| &= \frac{|h(y)-h(x)|}{|h(y)-h(x)|_g} \cdot |h(y)-h(x)|_g \\ &= (1+O(\theta^k)) \cdot Dh(x) \cdot |y-x|_g \\ &= (1+O(|y-x|^\beta)) \cdot Dh(x) \cdot |y-x|. \end{aligned}$$

□

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

$$|\Delta| = |h(x)-h(x_0)| \cdot (1+O(|x-x_0|^\beta)). \tag{10.11}$$

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,

$$\begin{aligned} |J| &= |h(x)-h(x_0)| \cdot (1+O(|h(x)-h(x_0)|^2)) \\ &= |h(x)-h(x_0)| \cdot (1+O(|x-x_0|^\beta)) \end{aligned} \tag{10.12}$$

and

$$|V| = O(|h(x)-h(x_0)|^2). \tag{10.13}$$

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. □

*Remark* 10.5. The conjugation $h : \mathcal{O}_F \to \mathcal{O}_{F_*}$ 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 $N \to \infty$.

The Cantor attractor $\mathcal{O}_F$ has two characteristic exponents, [O]. One is zero the other is $\ln b_F$, see [CLM]. The function $T : X \to \mathbb{P}^1$ constructed before defines a measurable line field, with respect to $\mu$, on $\mathcal{O}_F$.

**Proposition 10.10.** *The line field*

$$T : \mathcal{O}_F \to \mathbb{P}^1$$

*is the invariant line field of zero characteristic exponent.*

*Proof.* For each point $x_0 \in X$ we have, see Theorem 10.1,

$$dist(x, T_{x_0}) \le C_{x_0}|x - x_0|^{1+\beta}$$

with $x \in \mathcal{O}_F$. The map $F$ is a diffeomorphism which preserves $\mathcal{O}_F$. Hence,

$$dist(x, DF(x_0)T_{x_0}) = O(|x - F(x_0)|^{1+\beta})$$

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, [O]. Left is to show that $T(x)$ corresponds to the zero exponent.

Choose $N \ge 1$. For almost every $x_0 \in X_N$ there are $t_n \to \infty$ such that

$$F^{t_n}(x_0) \in X_N.$$

This is because the ergodic measure $\mu$ 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,

$$\lim_{t\to\infty} \frac{1}{t} \ln |DF^t(x_0)v| = \lim_{n\to\infty} \frac{1}{t_n} \ln |DF^{t_n}(x_0)v| = 0$$

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. □

The Hausdorff dimension of a measure $\mu$ 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 $h : \mathcal{O}_F \to \mathcal{O}_{F_*}$ 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 $\lim_{N\to\infty} HD(Z_N) = HD_\mu(\mathcal{O}_F)$ then

$$\begin{aligned} HD_\mu(\mathcal{O}_F) &\geq \lim_{N\to\infty} HD(Z_N \cap X) \\ &= \lim_{N\to\infty} HD(h(Z_N \cap X)) \\ &\geq HD_{\mu_*}(\mathcal{O}_{F_*}), \end{aligned}$$

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. □

*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 [BMT] and references therein. The measure $\mu_*$ is the projected measure from $\mathcal{O}_{F_*}$.

## 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 [CLM] on the regularity of the conjugation $h : \mathcal{O}_F \to \mathcal{O}_G$ when $b_F = b_G$.

Problem II: Do wandering domains exist? This question was already formulated in [LM1]. It is included again because its solution might be obtained by using the techniques developed in this paper.

## Nomenclature

| | |
|---|---|
| $b_F$ | average Jacobian, §3 |
| $B^n_\omega$ | a piece of the $n^{th}$-renormalization cycle, §3 |
| $\mathcal{B}^n$ | collection of pieces in the $n^{th}$-renormalization cycle, §3 |
| $\mathcal{B}^n[k]$ | pieces of $\mathcal{B}^n$ in $E^k$, §4 |
| $B_n(x)$ | the piece in $\mathcal{B}^n$ containing $x \in \mathcal{O}_F$ |
| $\mathbf{B}$ | the piece $B$ viewed from its proper scale, §4, Figure 4.2 |
| $\text{Dist}(\phi)$ | Distortion, (3.4) |
| $D_k$ | derivative of $\psi^k_v$ at the tip, (3.5) |
| $\delta_B$ | thickness of $B$, §5 |
| $\Delta_B$ | absolute thickness of $B$, §5 |
| $E^k$ | part of a dynamical partition, §4, Figure 2.1, 4.1 |
| $f_*$ | unimodal renormalization fixed point, §3 |
| $G_k$ | return map related to the partition by $E^k$, §4, Figure 2.1, 4.1 |
| $k_i(B)$ | depth of the $i^{th}$-predecessor of $B$, §4, Definition 7.2 |
| $\kappa_0(n)$ | minimal depth to safely push-up, §8, and §7.1 |
| $\kappa(n)$ | upper bound of the brute-force regime, §8, and Lemma 8.3 |
| $l(k)$ | maximal allowable depth, §4 |
| $\eta_\phi$ | Nonlinearity, (3.3) |
| $\mathcal{O}_F$ | invariant Cantor set of $F$, §3 |
| $\psi^k_{c,v}$ | coordinate changes related to the renormalization $R(R^kF)$, (3.1) |
| $\psi^n_\omega$ | coordinate change, §3 |
| $\Psi^n_k$ | coordinate change relating $R^{n-k}(R^kF)$ to $R^n$, (3.2) |
| $\mathcal{P}_n(k; q_0, q_1)$ | collection of $q_0, q_1$-controlled pieces, Definition 7.2 |
| $\mathcal{P}_n$ | pieces obtained by applying the three regimes, §8, and (8.7) |
| $q_0, q_1$ | boundary one-dimensional regime, §8, and Lemma 8.4 |
| $\sigma$ | scaling factor of the unimodal renormalization fixed point, §3 |
| $\sigma_B$ | scaling factor of $B$, §6 |
| $\mathcal{S}^n(\epsilon)$ | collection of pieces in $\mathcal{B}^n$ with $\epsilon$ precision, §6 |
| $t_k$ | tilt of the derivative of $\psi^k_v$ at the tip, (3.5) |
| $T$ | tangent line field to $\mathcal{O}_F$, §10 |
| $\tau_F$ | tip, §3 |
| $X$ | the differentiable part of $\mathcal{O}_F$, §10 |

## References

[BMT] C. Birkhoff, M. Martens, C. P. Tresser. On the Scaling Structure for Period Doubling. Asterisque (286) (2003), 167-186.

[CEK] P. Collet, J. P. Eckmann, H. Koch. Period doubling bifurcations for families of maps on $\mathbb{R}^n$. J. Stat. Physics 25 (1980), 1-15.

[CLM] A. de Carvalho, M. Lyubich, M. Martens. Renormalization in the Hénon family, I: universality but non-rigidity, J. Stat. Phys. 121 No. 5/6, (2005), 611-669.

[CMMT] V.V.M.S. Chandramouli, M. Martens, W. de Melo, C.P. Tresser, *Chaotic period doubling*, Erg. Th. and Dyn. Sys. 29, (2009), 381-418.

[Cv] P. Cvitanovic, Universality in Chaos (2nd edition). Adam Hilger, Bristol, (1989).

[CT] P. Coullet, C. Tresser. Itération d'endomorphismes et groupe de renormalisation. J. Phys. Colloque C 539, C5-25 (1978).

[Ep] H.Epstein, New proofs of the existence of the Feigenbaum functions. Comm. Math. Phys. 106 (1986), 395-426.

[F] M.J. Feigenbaum. Quantitative universality for a class of non-linear transformations. J. Stat. Phys., 19 (1978), 25-52.

[FMP] E. de Faria, W. de Melo, A. Pinto. Global hyperbolicity of renormalization for $C^r$ unimodal mappings, Ann. of Math. 164 No. 3, (2006), 731-824.

[GST] J.-M. Gambaudo, S. van Strien & C. Tresser. Hénon-like maps with strange attractors: there exist $C^\infty$ Kupka-Smale diffeomorphisms on $S^2$ with neither sinks nor sources. Nonlinearity 2 (1989), 287-304.

[J] P. Jones. Rectifiable sets and the traveling salesman problem. Invent. Math. 102, 1 (1990), 1-16.

[Lan] O.E. Lanford III, A computer assited proof of the Feigenbaum conjectures. Bull. Amer. Math. Soc. New Series 6, (1982), 427-434.

[L] M. Lyubich. Feigenbaum-Coullet-Tresser Universality and Milnor's Hairiness Conjecture. Ann. of Math. 149 (1999), 319 - 420.

[LM1] M. Lyubich, M. Martens. Renormalization in the Hénon family, II. The Heteroclinic Web, IMS Stony Brook preprint 08-2 and accepted for publication in Invent. Math.

[LM2] M. Lyubich, M. Martens. Renormalization of Hénon maps. Dynamics, Games and Science I, Springer Proceedings in Mathematics 1, M.M. Peixoto, A.A. Pinto, D.A. Rand (Eds.), (2011), 597-618.

[Ma1] M.Martens, *Distortion Results and Invariant Cantor sets for Unimodal maps*, Erg.Th and Dyn.Sys. 14, (1994), 331-349.

[Ma2] M.Martens, *The Periodic Points of Renormalization*, Ann. of Math. 147, (1998), 543-584.

[McM] C. McMullen. Renormalization and 3-manifolds which fiber over the circle. Annals of Math. Studies, 135, Princeton University Press, 1996.

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

[O] V.I. Oseledec. A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. Trans. Moscow. math. Soc. 19, (1969), 197-231.

[S] D. Sullivan. Bounds, quadratic differentials, and renormalization conjectures. AMS Centennial Publications. 2: Mathematics into Twenty-first Century (1992).

SUNY at Stony Brook