

Non-Uniformly Hyperbolic Horseshoes Arising from Bifurcations of Poincaré Heteroclinic Cycles

Jacob PALIS* and Jean-Christophe YOCCOZ

Abstract

The purpose of this paper is to advance the knowledge of the dynamics arising from the creation and subsequent bifurcation of Poincaré heteroclinic cycles. The problem is central to dynamics: it has to be addressed if, for instance, one aims at describing the typical orbit behaviour of a typical system, thus providing a global scenario for the ensemble of dynamical systems - see the Introduction and [P1, P2]. Here, we shall consider smooth, i.e. C^∞ , one-parameter families of dissipative, meaning non-conservative, surface diffeomorphisms. An heteroclinic cycle may appear when the parameter evolves and an orbit of tangency, say quadratic, is created between stable and unstable manifolds (lines) of periodic orbits that belong to a basic hyperbolic set. The key novelty is to allow this basic set, a horseshoe, to have Hausdorff dimension bigger than one. In the present paper we do assume such a dimension to be beyond one, but in a limited way, as explicitly indicated in the Introduction. [A mild non-degeneracy condition on the family of maps is assumed: at the orbit of tangency the invariant lines, stable and unstable, cross each other with positive relative speed]. We then prove that most diffeomorphisms, corresponding to parameter values near the bifurcating one, are non-uniformly hyperbolic in a neighborhood of the horseshoe and the orbit of tangency; such diffeomorphisms display no attractors nor repellers in such a neighborhood. A first precise formulation of our main theorem is at the Introduction and a more encompassing version at the end of the paper. These results were announced in [PY3].

*Partially support by CNPq and FAPERJ, Brazil.

Contents

1	Introduction	4
1.1	The Context	4
1.2	The Setting and a First Formulation of the Main Result	7
1.3	A Summary of the Next Sections of the Paper	9
2	Markov Partition and Folding Map	16
2.1	Markov Partition and Related Charts	16
2.2	The Parabolic Tongues L_u, L_s	16
2.3	The Folding Map G	18
3	Affine-like Maps	20
3.1	Definition and Implicit Representation	20
3.2	Cone Condition and Distortion	21
3.3	Simple Composition	22
3.4	Properties of the Markov Partition	23
3.5	Parabolic Composition	23
4	Structure of Parameter Space	28
4.1	Some Important Constants	28
4.2	One-Parameter Families	28
4.3	Parameter Intervals	29
4.4	The Selection Process	29
4.5	Strongly Regular Parameters	30
5	Classes of Affine-like Iterates and the Transversality Relation	31
5.1	Affine-like Iterates	31
5.2	The Classes $\mathcal{R}(I)$: General Overview	32
5.3	Definition of the Special Class of Affine-Like Iterates $\mathcal{R}(I)$	33
5.4	Definition of the Transversality Relation	36
5.5	The Class $\mathcal{R}(I_0)$	37
5.6	Definition of the Regularity Property	38
6	Some Properties of the Classes $\mathcal{R}(I)$	40
6.1	Transversality is Hereditary	40
6.2	Criticality and Decomposability	41
6.3	Concavity	41
6.4	Children are Born From Their Parent	43
6.5	Coherence and Parametric Concavity	45
6.6	Further Criteria for Transversality	48
6.7	A Structure Theorem for New Rectangles	49

7	Estimates for the Classes $\mathcal{R}(I)$	58
7.1	A Stretched Exponential Estimate for Widths	58
7.2	Uniform Cone Condition	59
7.3	Bounded Distortion	60
7.4	Estimates for the Special Rectangles P_s and Q_u	61
7.5	Proof of the Property (R4) of Section 5.3	64
7.6	Relative Speeds of Critical Rectangles	67
7.7	Variation of Width of Critical Rectangles	70
8	Number of Children and Dimension Estimates	78
8.1	Estimates on the Number of Children	78
8.2	A Dimension Estimate	83
8.3	Transfer to Parameter Space	86
9	Strong Regularity and Parameter Selection	94
9.1	Partitions of the Critical Locus	94
9.2	Size of the Critical Locus	95
9.3	Classes of Bicritical Rectangles	98
9.4	Number of Bicritical Rectangles	99
9.5	The Starting Interval	104
9.6	New Bicritical Rectangles	105
9.7	The Case $I = I_\alpha \neq I_\omega$	109
9.8	The Case $I = I_\alpha = I_\omega$: General Overview	112
9.9	Subclasses of Bicritical Elements	115
9.10	The Size of the Index set L	122
9.11	The Size of \widehat{B}	124
9.12	End of the Induction Step for $(\text{SR3})_s$	127
9.13	The Induction Step for $(\text{SR3})_u$	131
10	The Well-Behaved Part of the Dynamics for Strongly Regular Parameters	132
10.1	Prime Elements and Prime Decomposition	132
10.2	Number of Factors in a Prime Decomposition	133
10.3	A Weighted Estimate on the Number of Children	135
10.4	Stable Curves	139
10.5	Topology and Geometry of \mathcal{R}_+^∞ and $\widetilde{\mathcal{R}}_+^\infty$	141
10.6	Transverse Dilatation	143
10.7	Definition of a Transverse Operator	145
10.8	Spectral Properties of the Transfer Operator	147
10.9	The Gibbs Measure	150
10.10	Transverse Hausdorff Dimension of $\widetilde{\mathcal{R}}_+^\infty$	153
10.11	Invariant Measures	156

11 Some Further Geometric Properties of the Invariant Set	159
11.1 One-Dimensional Analysis of the Invariant Set	159
11.2 Parabolic Cores	160
11.3 Decomposition of $c(P) \cap \omega^*$	161
11.4 The Structure of $\omega^* \cap \Lambda$	163
11.5 Hausdorff Dimension of the Exceptional Set $\mathcal{E}(\omega^*)$	165
11.6 The Stable and Unstable Sets of Λ	169
11.7 Decomposition of \mathcal{E}^+	169
11.8 Size and Area of Parabolic Cores	171
11.9 Proof of Theorem 4	173

1 Introduction

1.1 The Context

Since Poincaré referred to a property valid “pour la plupart des coefficients” of a polynomial (analytic) dynamical system, the outstanding problem of describing at large the orbits of a “typical” system became the source of much creative work in dynamics.

A stumbling block in a possible program to solve this question is the understanding of dynamics arising from bifurcations of homoclinic or heteroclinic cycles. Such cycles were defined by Poincaré himself: they involve stable and unstable manifolds of invariant sets (typically periodic orbits) that successively intersect each other. In his classic book on Celestial Mechanics, he prophetically stated: “Rien n’est plus propre à nous donner une idée de la complication de tous les problèmes de Dynamique”.

In fact, the creation and unfolding of cycles, in particular homoclinic tangencies for surface diffeomorphisms and their unfolding, led Newhouse to show the non-denseness of hyperbolic dynamics, thus contradicting Smale’s remarkable conjecture of the early 60’s. That is, there are systems that cannot be approximated by one with a hyperbolic limit set. On the way, Newhouse showed that in this context always appear surface diffeomorphisms displaying infinitely many simultaneous sinks (periodic attractors) or sources (periodic repellers).

Abundance of other more intricate kind of attractors, the so called Hénon-like ones, was proved to be also present in the unfolding of such cycles. This was another striking fact. It resulted from the works of Benedicks-Carleson, Mora-Viana and Colli. Attractors here mean invariant sets that attract future orbits of points of a positive Lebesgue measure set in the phase space (space of events).

In view of all these intricacies inherent to homoclinic and heteroclinic bifurcations, a new global conjecture has been proposed in [P1] (see also [P2]) concerning a typical dynamical system: In particular, systems with finitely many attractors should be dense in the universe of dynamics, i.e. C^r flows, diffeomorphisms and maps, with $r \geq 1$. Also, their basins of attraction should cover the whole phase space, except for a Lebesgue zero measure set.

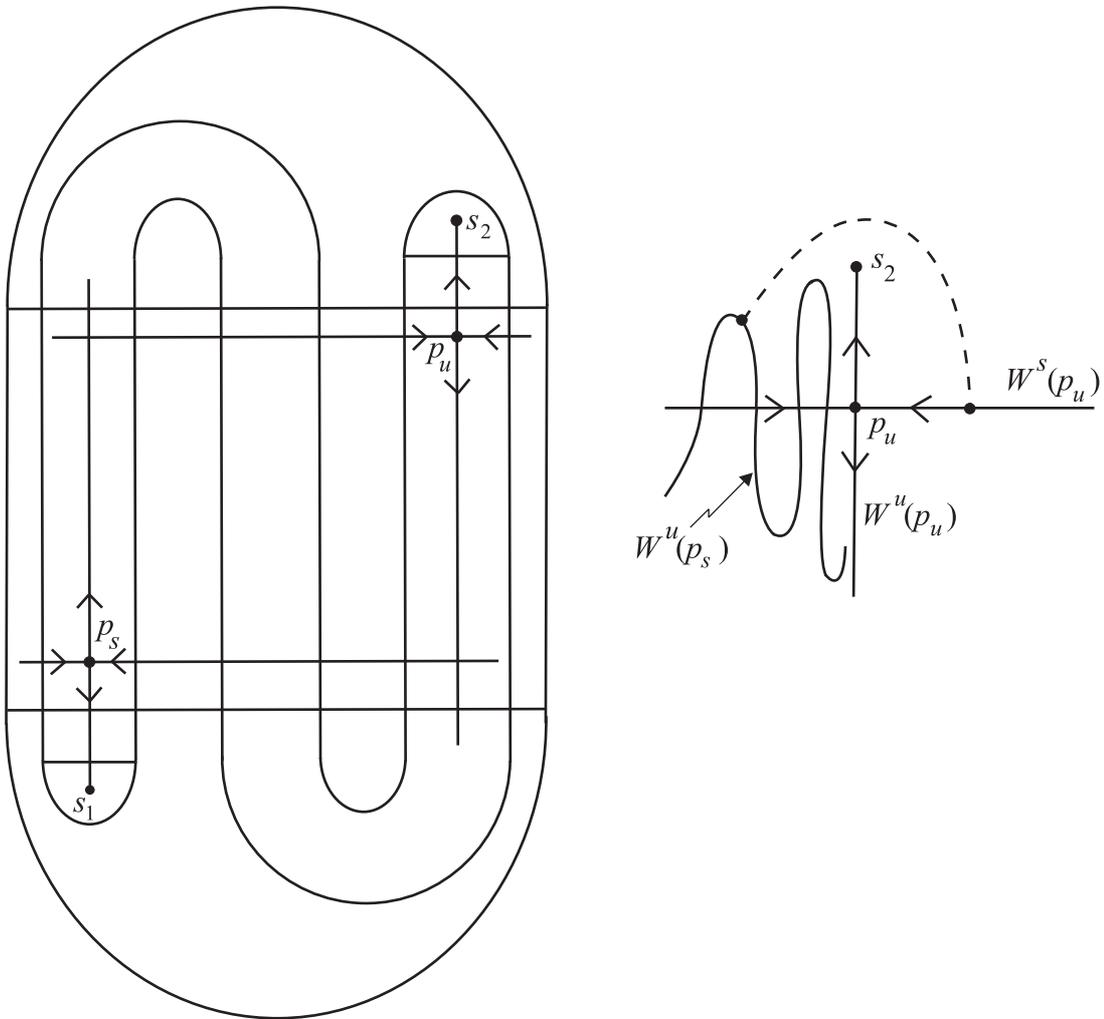


Figure 1

An example of the creation of a heteroclinic cycle associated to a (hyperbolic) horseshoe is indicated in Figure 1. Initially we have a classic Smale's horseshoe map (diffeomorphism) on the two-sphere S^2 with two saddle fixed points p_s, p_u with positive eigenvalues, a fixed point repeller outside the figure and two fixed points attractors s_1 and s_2 . The rectangle inside the figure is sent by the map to the snake-shaped piece, while the bigger top half-disk is sent to the small one around s_2 and the lower bigger half disk is sent to the small one around s_1 . At the right hand side of the figure, we show how to move a small neighborhood of a point in the stable line of p_u so to create a tangency with the unstable manifold of p_s . This is done through a one-parameter family of diffeomorphisms;

until we create such a tangency the corresponding map remains hyperbolic, i.e. having a hyperbolic limit set with no cycles among its basic sets.

Concerning Smale's conjecture on the denseness of hyperbolicity, it is to be remarked that the question is still open for rational maps of the Riemann sphere. On the other hand, it was known to be true in the early 60's for flows on orientable surfaces (Peixoto), preceded by the case of the two-disk (Andronov-Pontryagin). Much more recently, it was proved to be true for quadratic maps of the interval (Graczyk-Swiatek and Lyubich) and then, more generally, unimodal maps (Kozlovski) and, finally, multimodal maps (Kozlovski-Shen-Van Strien).

We point out that comprehensive accounts of the results mentioned above and many other related ones are in [BDV], [P1] and [P2].

The present paper represents a contribution to the understanding of the dynamics arising from bifurcating a cycle of a C^∞ surface diffeomorphism. The notion of "most" used here requires us to consider one parameter families of diffeomorphisms g_t strictly containing the initial bifurcating one, say g_0 at parameter value $t = 0$. We assume that g_t is hyperbolic for $t < 0$ and $|t|$ small. We suppose that the cycle is formed by a (hyperbolic) horseshoe K and an orbit of tangency $o(q)$ between stable and unstable manifolds of different periodic orbits of K . We assume the maximal invariant set in a small neighborhood of $K \cup o(q)$ to consist precisely of $K \cup o(q)$. A main novelty is that we allow the Hausdorff dimension of K to be larger than one, but not too far from it. We show that right after the bifurcation, i.e. for $t > 0$ small, most diffeomorphisms still display no attractors nor repellers in some neighborhood of $K \cup o(q)$. This means that the parameter values corresponding to diffeomorphisms displaying no attractors nor repellers should have total density at $t = 0$. The concept is again discussed in the next subsection.

Our results considerably extend those in [PT], [NP] obtained for the case when the Hausdorff dimension $HD(K)$ is smaller than one. They were announced in [PY3].

Of course, we expect the same to be true for all cases $0 < HD(K) < 2$. For that, it seems to us that our methods need to be considerably sharpened: we have to study deeper the dynamical recurrence of points near tangencies of higher order (cubic, quartic ...) between stable and unstable curves. We also hope that the ideas introduced in the present paper might be useful in broader contexts. In the horizon lies the famous question whether for the standard family of area preserving maps, one can find sets of positive Lebesgue probability in parameter space such that: the corresponding maps display non-zero Lyapunov exponents in sets of positive Lebesgue probability in phase space. Finally, we expect our results to be true in higher dimensions (see [MPV]).

We wish to thank W. de Melo e M. Viana for fruitful conversations.

1.2 The Setting and a First Formulation of the Main Result

Let f be a smooth, i.e. C^∞ diffeomorphism of a smooth surface M .

Recall that a *basic set* is a compact hyperbolic transitive locally maximal invariant set. A basic set is a *horseshoe* if it is infinite and is neither an attractor nor a repeller.

We assume that there exists a basic set K for f , points $p_s, p_u \in K$, $q \in M - K$ such that the following properties hold:

(H1) p_s and p_u are periodic points and belong to distinct periodic orbits;

(H2) $W^s(p_s)$ and $W^u(p_u)$ have a quadratic tangency at q ;

(H3) there exists a neighbourhood U of K , a neighbourhood V of the orbit $\mathcal{O}(q)$ of q , such that $K \cup \mathcal{O}(q)$ is the maximal invariant set in $U \cup V$.

We would like to understand, when U, V are appropriately small and g is C^∞ close to f , the maximal invariant set

$$(1.1) \quad \Lambda_g = \bigcap_{\mathbb{Z}} g^{-n}(U \cup V).$$

Observe that the smaller set

$$(1.2) \quad K_g = \bigcap_{\mathbb{Z}} g^{-n}(U)$$

is a horseshoe which is the hyperbolic continuation of K .

Let \mathcal{U} be an appropriately small neighbourhood of f in $\text{Diff}^\infty(M)$. We still denote by p_s, p_u the continuation of these hyperbolic periodic points in \mathcal{U} . The condition that $W^s(p_s), W^u(p_u)$ have a quadratic tangency near q defines a codimension 1 hypersurface \mathcal{U}_0 through f in \mathcal{U} . It divides \mathcal{U} into regions $\mathcal{U}_+, \mathcal{U}_-$ such that, for $g \in \mathcal{U}_-$, $W^s(p_s)$ and $W^u(p_u)$ do not intersect near q while, for $g \in \mathcal{U}_+$, $W^s(p_s)$ and $W^u(p_u)$ have two transverse intersection points near q (for obvious dynamical reasons, the intersection is actually infinite in this case; we are really considering here the intersection derived from the continuation of large *compact* curves contained in $W^s(p_s)$ and $W^u(p_u)$).

When $g \in \mathcal{U}_-$, we clearly have

$$(1.3) \quad \Lambda_g = K_g$$

When $g \in \mathcal{U}_0$, we have

$$(1.4) \quad \Lambda_g = K_g \cup \mathcal{O}(q_g),$$

where q_g is the tangency point close to q given by the definition of \mathcal{U}_0 .

The interesting case is therefore $g \in \mathcal{U}_+$.

It is actually not realistic to try to understand Λ_g for all $g \in \mathcal{U}_+$. One of the reasons is the so-called Newhouse's phenomenon [N]: there exists an open set $\mathcal{N} \subset \mathcal{U}_+$, with $\mathcal{U}_0 \subset \overline{\mathcal{N}}$, such that, residually in \mathcal{N} , Λ_g has infinitely many periodic sinks or sources and so its full dynamical description appears to be beyond reach. See also [BC], [MV], [C] for similar results involving Hénon-like attractors.

Still, we can and shall consider most $g \in \mathcal{U}_+$ in the following sense.

We will say that a subset $\mathcal{P} \subset \mathcal{U}_+$ contains most $g \in \mathcal{U}_+$ if, for any smooth 1-parameter family $(g_t)_{t \in (-t_0, t_0)}$ which is transverse to \mathcal{U}_0 at $t = 0$ (with $g_t \in \mathcal{U}_+$ for $t > 0$), we have

$$(1.5) \quad \lim_{t \rightarrow 0} \frac{1}{t} \text{Leb}(s \in (0, t], g_s \in \mathcal{P}) = 1.$$

Denote by $W^s(K)$ (resp. $W^u(K)$) the stable set (resp. unstable set) of K for f . This is a partial foliation with a $C^{1+\alpha}$ Cantor transverse structure; denote by d_s^0 (resp. d_u^0) the transverse Hausdorff dimension of $W^s(K)$ (resp. $W^u(K)$). The Hausdorff dimension of K is equal to $d_s^0 + d_u^0$. We then have, in some contrast to Newhouse's phenomenon:

Theorem. *[PT], [NP] Assume that $d_s^0 + d_u^0 < 1$. Then, for most $g \in \mathcal{U}_+$, Λ_g is a horseshoe.*

On the other hand, by [PY1], the same conclusion does not hold when $d_s^0 + d_u^0 > 1$. The paper [MY] gives substantially more geometric information in this case, specially concerning tangencies between stable and unstable manifolds (lines) in the hyperbolic continuation K_g of K . These results have been extended to higher dimensions, as announced in [MPV] and complete proofs to appear in the near future.

In the present work, we investigate the maximal invariant set Λ_g , for most $g \in \mathcal{U}_+$, provided that the dimensions d_s^0, d_u^0 satisfy (see figure 2)

$$(H4) \quad (d_s^0 + d_u^0)^2 + (\max(d_s^0, d_u^0))^2 < d_s^0 + d_u^0 + \max(d_s^0, d_u^0).$$

Our results can essentially be summarized as:

Main Theorem. *Assume that (H1), (H2), (H3), (H4) hold. Then, for most $g \in \mathcal{U}_+$, Λ_g is a non-uniformly hyperbolic horseshoe.*

The meaning of a non-uniformly hyperbolic horseshoe in the present context will be explained somewhat in the next section and more completely in the rest of the paper. We can, however, comment that, for most $g \in \mathcal{U}_+$, Λ_g will be a saddle-like object in the sense that both the stable set $W^s(\Lambda_g)$ and the unstable set $W^u(\Lambda_g)$ have Lebesgue measure zero and, so, it carries no attractors nor repellers. It will be (non-uniformly) hyperbolic in the sense that we will construct geometric

invariant measures, à la Sinai-Ruelle-Bowen [Si, Ru, BR], on $\Lambda_g \subset W^s(\Lambda_g)$ and $\Lambda_g \subset W^u(\Lambda_g)$ with non-zero Lyapunov exponents. Such properties of the invariant set Λ_g are made especially precise in Section 10 and 11, the last ones in the paper. They yield some rephrasing of the main result in these terms, which is presented at the end of Section 11.

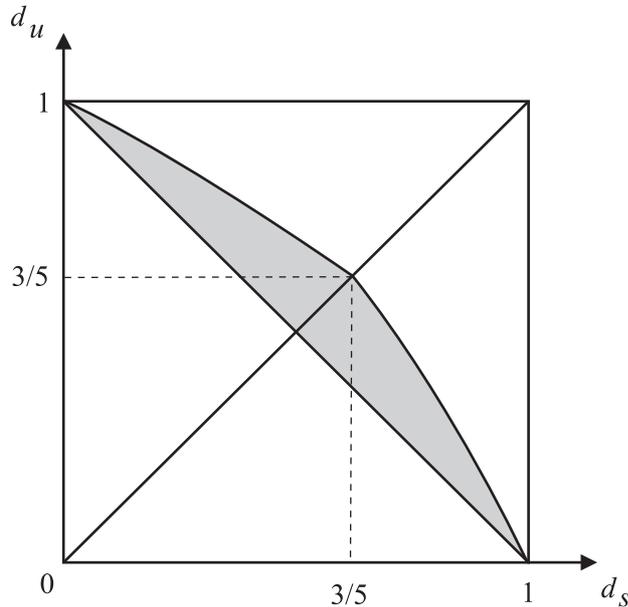


Figure 2

Remark. *In the case when $d_s^0 + d_u^0 < 1$, mentioned above and studied in [PT], [NP], it is not necessary to assume that p_s, p_u belong to distinct periodic orbits. It is probably not necessary in our case either, at least as far as the qualitative statements are concerned. But, this assumption seems to make the technicalities significantly easier in what is already a very long construction.*

1.3 A Summary of the Next Sections of the Paper

Sections 2–4 consist mainly of preparatory work.

In Section 2, we introduce a Markov partition by smooth disjoint rectangles $(R_a)_{a \in \mathcal{A}}$ for the horseshoe K . The dynamics in the neighbourhood U of K is given by the transition maps from one rectangle to another, which enjoy a nice hyperbolic behaviour. To understand the dynamics in the larger set $U \cup V$, we need to control the dynamics along a finite part of the orbit of q , stretching from the moment this orbit goes out of $R := \sqcup R_a$ until it comes back to R . The region of exit of R and the region of entry into R are two parabolic tongues L_u and L_s , and the transition map

$$G = g^{N_0} : L_u \rightarrow L_s$$

is a folding map which share many features with the Henon quadratic polynomial diffeomorphisms of the plane.

Section 3 is essentially a summary of our previous work [PY2], (which was written having the present paper in mind). The important concept of affine-like map is introduced. The basic idea, which goes back to the early stages of the hyperbolic theory, is to describe maps that present hyperbolic features in an implicit way exhibiting preference for coordinates with a macroscopic range. Concretely, if a two-dimensional diffeomorphism contracts the vertical coordinate y and expands the horizontal coordinate x , we use y_0 and x_1 as independent variables associated with a point (x_0, y_0) and its image (x_1, y_1) , writing x_0 and y_1 as functions of y_0 and x_1 .

Cone conditions are easy to formulate in this setting. A nice feature of this implicit representation of the dynamics is that it is time-symmetric: the map and its inverse satisfy symmetric formulas. Another even more important feature is that this formalism is well-adapted to the right concept of distortion (for 2-dimensional maps), yielding appropriate control on the partial derivatives of order two.

Composition of two affine-like maps which satisfy the same cone condition is also affine-like, and the distortion is only slightly bigger than the distortion of the two maps. Besides this “simple” composition, we study “parabolic” compositions of the form $F_1 \circ G \circ F_0$, where F_0, F_1 are affine-like and G is the folding map of Section 2. When the relative positions of the parabolic strip $G(Q_0)$ (where Q_0 is the image of F_0) and P_1 (the domain of F_1) are appropriate, the domain of $F_1 \circ G \circ F_0$ has two connected components and the restrictions F^\pm of $F_1 \circ G \circ F_0$ to each component is affine-like. A control of the distortion of F^+ and F^- is also obtained.

In Section 4, the general structure of parameter space is introduced. The parameter coordinate is normalized by the relative speed at the quadratic tangency of the tips of the stable and unstable manifolds. Then, with ε_0 very small, the starting interval $I_0 := [\varepsilon_0, 2\varepsilon_0]$ for the parameter selection process is introduced. A small parameter τ (with $\tau \ll 1$ but still $\varepsilon_0^\tau \ll 1$) determines a sequence of scales $(\varepsilon_k)_{k \geq 0}$ in parameter space through the formula $\varepsilon_{k+1} = \varepsilon_k^{1+\tau}$. At level k , we have disjoint parameter intervals of length ε_k (starting from level 0 with I_0). Each parameter interval of level k that has been selected is divided into $[\varepsilon_k^{-\tau}]$ disjoint candidates of length ε_{k+1} . These candidates will pass a test to decide whether they are selected at level $k+1$.

The test takes two forms. First, in Section 5, a property of the parameter interval called regularity (see below) will be introduced; candidates which do not possess this property are discarded. Such a property is sufficient to develop in Sections 5–8 some basic combinatorial and quantitative properties, but it is not well-adapted to an inductive scheme. Hence, in Section 9, a stronger property called strong regularity is introduced, and candidates have to satisfy this property in order to be selected.

Sections 5–7 constitute in some sense a single logical step: in Section 5, certain classes of restrictions of iterates of g_t are inductively defined, and the definition is only possible because of properties that are inductively proved in Sections 6 and 7.

In Section 5, the goal is to define, for each parameter interval I which is a candidate (i.e. its parent interval at the immediately upper level has been tested as regular), a class $\mathcal{R}(I)$ of I -persistent affine-like iterates. An I -persistent affine-like iterate is a triple (P, Q, n) where P is a vertical-like strip in some rectangle R_a depending on $t \in I$, Q is a horizontal-like strip in some rectangle $R_{a'}$, depending on $t \in I$, and the restriction of g_t^n to P is a diffeomorphism onto Q which is affine-like.

However, we do not want to have in $\mathcal{R}(I)$ all I -persistent affine-like iterates: we will argue about them by induction (on n , for instance) and in order to do this, we want to obtain them in some explicit constructive way. Therefore, a number of Axioms, (R1)–(R7), are introduced and together they completely determine the class $\mathcal{R}(I)$. The most important feature of these Axioms is the following: every element of $\mathcal{R}(I)$ consisting of more than one iteration of g_t can be obtained from simpler elements of $\mathcal{R}(I)$ by simple or parabolic composition; in this context, the notions of parent and simple or non-simple child introduced here, play a relevant role; simple composition is allowed in $\mathcal{R}(I)$ whenever it makes sense; and parabolic compositions of elements of $\mathcal{R}(I)$ is allowed if and only if a certain *transversality relation* is satisfied.

Thus, the definition of $\mathcal{R}(I)$ is reduced to the definition of this transversality relation, which is presented in Subsection 5.4. The intuitive notion behind the formal definition is the following: an element (P, Q, n) should be I -transverse to an element (P', Q', n') if the distance $\delta(Q, P')$ between the tip of the parabolic-like strip $G(Q)$ and P' satisfies

$$\delta(Q, P') \geq |Q|^{1-\eta} + |P'|^{1-\eta}$$

for all $t \in I$, where $|Q|$, $|P'|$ are the widths of the strips Q and P' . Here η is a small positive constant, fixed once and for all. However, a number of properties, presented in Section 6, are very helpful, and they require a definition of the transversality relation that may seem quite complicated. In Appendix C, we explain why this seemed complication is rather necessary.

For the starting interval I_0 , it follows from the formal definition that the transversality relation is never satisfied; therefore, parabolic composition is not allowed and the class $\mathcal{R}(I_0)$ is exactly the one associated with the symbolic dynamics given by the Markov partition. We conclude Section 5 with the definition of *regularity*. First, one says that a strip P (from an element $(P, Q, n) \in \mathcal{R}(I)$) is I -transverse if one can find finitely many Q_α , whose union contains the unstable set of Λ , such that Q_α and P are I -transverse for every α ; otherwise one says that P is I -critical. Then, given a constant $\beta > 1$, one says that the parameter interval is β -regular if any $(P, Q, n) \in \mathcal{R}(I)$ such that both P and Q are I -critical satisfies $|P| < |I|^\beta$, $|Q| < |I|^\beta$ for all $t \in I$. Intuitively, this means that no short return to the critical set is allowed.

In Section 6, we prove a number of properties of the transversality relation and the classes $\mathcal{R}(I)$.

Amongst the most important is the following: children are born from their parents. Let us explain what it means. Let $(P, Q, n) \in \mathcal{R}(I)$, and let $(\tilde{P}, \tilde{Q}, \tilde{n})$ the element of $\mathcal{R}(I)$ such that $P \subset \tilde{P}$, $P \neq \tilde{P}$ and \tilde{P} is the thinnest rectangle with this property; one says that P is a child of \tilde{P} and that \tilde{P} is the parent of P . There are two cases; either $n = \tilde{n} + 1$ and one says that P is a simple child; (P, Q, n) is obtained by simple composition of $(\tilde{P}, \tilde{Q}, \tilde{n})$ with an element of length 1; or $n > \tilde{n} + 1$ and one says that P is a non-simple child; one then proves that (P, Q, n) is obtained by parabolic composition of $(\tilde{P}, \tilde{Q}, \tilde{n})$ with some element (P_1, Q_1, n_1) .

The most important result in Section 6 is a structure theorem for new rectangles in Subsection 6.7. One considers an element (P, Q, n) which belongs to $\mathcal{R}(I)$ but not to $\mathcal{R}(\tilde{I})$, where \tilde{I} is the parameter interval containing I of the level immediately inferior (one says that \tilde{I} is the parent of I). Then there is a unique way to write (P, Q, n) as the result of a sequence of k parabolic compositions, possible in $\mathcal{R}(I)$ but not in $\mathcal{R}(\tilde{I})$, of elements $(P_0, Q_0, n_0), \dots, (P_k, Q_k, n_k)$. This fundamental result has several useful corollaries.

We have grouped in Section 7 a number of calculations and estimates related to the definitions of Section 5, and that are necessary for the inductive construction of the classes $\mathcal{R}(I)$. The first result relates the length n of an element (P, Q, n) to the width $|P|, |Q|$ of the strips. While there is no uniform exponential estimate as in the uniformly hyperbolic case, we are still able to prove a stretched-exponential uniform estimate.

We prove next that a uniform cone condition is satisfied by the affine-like iterates that we consider, and also that they have uniformly bounded distortion. After a technical estimate related to parabolic composition, we deal in Subsection 7.6 with the relative speed of the strips when the parameter varies; this is clearly of capital importance if we are to succeed. A point which is worth mentioning is that we are not able to obtain estimates for all pairs of strips (actually, it is easy to see that such estimates do not exist); we have to restrict ourselves to strips satisfying certain criticality conditions, that, fortunately, will be satisfied every time we need some information on these speeds. In the last Subsection 7.7, we investigate the oscillation of the widths of the strips with the parameter. While it is just not true that the relative oscillation is bounded (in the sense that the maximum over a parameter interval is no greater than a constant times the minimum), the result that we get will allow us to argue as if it was.

At the end of Section 7, the construction of the classes $\mathcal{R}(I)$ is complete, for every parameter interval I whose parent \tilde{I} is regular. But we still don't know whether a single interval I is regular.

In Section 8, we develop several quantitative estimates that will turn out to be crucial both in the parameter selection process of Section 9 and in the analysis of the dynamics for strongly regular parameters in Sections 10 and 11. We first investigate, for a given element $(\tilde{P}, \tilde{Q}, \tilde{n})$, the number of elements (P, Q, n) such that P is a non-simple child of \tilde{P} ; we show that, for every $\varepsilon > 0$, there are at most $\varepsilon^{-C\eta}$ such non-simple children with width $|P|$ larger than $\varepsilon|\tilde{P}|$. The constant η here is small

and related to the definition of the transversality relation. The meaning of this estimate is that the presence of non-simple children is not too significant from the point of view of Hausdorff (or box) dimension, as it is made clear in Subsection 8.2. In Subsection 8.3, we transfer this information to parameter space, combining it with the result on relative speed of strips in Subsection 7.6.

Section 9 is the longest one in the paper and deals with the parameter selection process. The concept of regularity is very useful to develop a number of properties of the classes $R(I)$, as we did in Sections 5–8. Unfortunately, we are not able to prove (and it is probably false) that, given a β -regular interval \tilde{I} , most candidates $I \subset \tilde{I}$ at the next level are β -regular. [It is a consequence of the structure theorem of Section 6 that all candidates are $\bar{\beta}$ -regular, where $\bar{\beta} = \beta(1 + \tau)^{-1}$ is very close to β ; this allow us to obtain all qualitative consequences of regularity for all candidates; but obviously we cannot repeat this at many successive levels of parameter intervals, because we need to keep $\beta > 1$.] The problem with the concept of regularity is that it is dealing with only one scale $|\tilde{I}|^\beta$; it could happen a priori that for a regular parameter interval \tilde{I} we have many \tilde{I} -bicritical $(P, Q, n) \in \mathcal{R}(I)$ with $|P|$ or $|Q|$ only slightly below the threshold $|\tilde{I}|^\beta$ (and therefore above the next threshold $|I|^\beta$ for candidates $I \subset \tilde{I}$); for each such (P, Q, n) , we have to eliminate candidates I such that (P, Q, n) is I -bicritical, and no candidate will survive this selection process if there are too many (P, Q, n) .

The solution to this difficulty is to introduce the condition of strong regularity, which implies regularity and gives a quantitative control at all scales. Actually, the strong regularity condition involves two parts. In the first, one controls the size of the critical locus (in several slightly different ways) by a series of eight inequalities which all amount to say that the "dimension" of the critical locus is not much larger than $d_s^0 + d_u^0 - 1$. In this case, the parameter selection process is based on the result mentioned above in the last part of Section 8. The second part of the strong regularity condition, by far the most subtle one, is a quantitative estimate for the number of bicritical elements at all scales. Because of the inductive nature of the argument, which relies in an essential way on the structure theorem of Section 6, we need to control the number of elements $(P, Q, n) \in \mathcal{R}(I)$ such that P is I_α -critical, Q is I_ω -critical and $|P| \geq x$ for some $t \in I$. Here, I_α and I_ω are parameter intervals containing I , and the control will depend on I_α , I_ω and x . The formulas in Subsection 9.4 present a phase transition with respect to the width parameter x . Discussing this phase transition leads naturally to the hypothesis (H4) on the transverse dimensions d_s^0 , d_u^0 : a small calculation shows that (H4) is exactly what one needs to obtain β -regularity with $\beta > 1$.

Having stated the strong regularity condition, the goal in the rest of Section 9 is to prove that, given a strongly regular parameter interval \tilde{I} , most candidates $I \subset \tilde{I}$ at the next level are also strongly regular (the proportion of failed candidates turns out to be not larger than $C|I|^\tau$). This requires the control of two things. First, is to bound the number of "new" bicritical elements $(P, Q, n) \in \mathcal{R}(I)$ which did not belong to $\mathcal{R}(\tilde{I})$; this is based on the structure theorem of Section 6 and leads to a long but straightforward calculation. Second, is to estimate which proportion of bicritical elements for \tilde{I} are still bicritical for I ; this is only necessary when I_α or I_ω is equal to

I ; when only one of the two intervals I_α, I_ω is equal to I , the idea is simply to estimate what is the mean proportion (over all candidates), and to discard candidates for which the proportion is much above the mean. To compute the mean proportion, we rely again on the result at the end of Section 8. The case where $I = I_\alpha = I_\omega$ is the most important and the most difficult. When x is “large”, the same argument than when $I = I_\alpha \neq I_\omega$ still applies; but when x is “small”, the phase transition of the estimate means that the argument is not sufficient any more. A more complicated strategy is required, which is explained in Subsection 9.8 and carried out in 9.9 through 9.13.

It is worth mentioning that up to the end of Section 9, we never consider the dynamics for a single parameter, only for parameter intervals. In the last two sections, we study the dynamics for a strongly regular parameter value, i.e. the intersection of a decreasing sequence (I_m) of strongly regular parameter intervals.

In Section 10, we study the dynamics on the set of stable curves. A *stable curve* ω is the decreasing intersection of a sequence of vertical-like strips P_k , where $(P_k, Q_k, n_k) \in \mathcal{R} = \bigcup_m \mathcal{R}(I_m)$. The set of stable curves is denoted by \mathcal{R}_+^∞ , their union by $\tilde{\mathcal{R}}_+^\infty$. In order to define a map on $\tilde{\mathcal{R}}_+^\infty$ (which is not invariant under g), we introduce the concept of prime element in \mathcal{R} , i.e. one which cannot be written as the simple composition of shorter elements. Let then ω be a stable curve which is not contained in infinitely many prime elements P_k , and let (P, Q, n) be such that P is the thinnest prime element containing ω . The image $g^n(\omega)$ is contained in a stable curve ω' and we set $T^+(\omega) = \omega', \tilde{T}^+/\omega = g^n/\omega$. This defines a map T^+ from a subset \mathcal{D}_+ of \mathcal{R}_+^∞ onto \mathcal{R}_+^∞ which lifts to a map \tilde{T}^+ from the union $\tilde{\mathcal{D}}_+$ of curves in \mathcal{D}_+ to $\tilde{\mathcal{R}}_+^\infty$.

The map T^+ is Bernoulli in the following sense: its domain \mathcal{D}_+ splits into countably many pieces $\mathcal{R}_+^\infty(P)$ indexed by prime elements, and each piece is sent homeomorphically by T^+ onto the intersection of \mathcal{R}_+^∞ with some rectangle R_a of the Markov partition.

The map T^+ is uniformly expanding (with countably many branches) and we introduce a one parameter family of weighted transfer operators in the spirit of classical uniformly hyperbolic maps. One has only to be careful because the presence of countably many branches is the source of some problems, which are dealt with in Subsection 10.3 using the estimates of Section 8 on the number of children.

As expected, the transfer operators L_d , considered in the appropriate function space, turn out to have a positive eigenfunction h_d associated with a dominant eigenvalue $\lambda_d > 0$. There is a unique value d_s such that $\lambda_{d_s} = 1$. This value turns out to be, unsurprisingly, the transverse Hausdorff dimension of the partial foliation $\tilde{\mathcal{R}}_+^\infty$ (which is proved in Subsection 10.5 to be transversally Lipschitzian). The transfer operator also allows us to identify, as usual, a measure μ_d with prescribed Jacobian and an invariant measure $\nu_d = h_d \mu_d$. For $d = d_s$, the μ_d -measure (or ν_d -measure) of the set of stable curves contained in any vertical-like strip P is proportional to $|P|^{d_s}$.

The set $\tilde{\mathcal{R}}_+^\infty - \tilde{\mathcal{D}}_+$ where \tilde{T}^+ is not defined, has transverse dimension smaller than d_s , hence is negligible in a geometrical sense. One can lift the T^+ -invariant measure $\nu = \nu_{d_s}$ to a \tilde{T}^+ -invariant measure $\tilde{\nu}$ which is ergodic and then spread it to a g -invariant measure on Λ .

In Section 11, the last in the paper, we pursue the study of the dynamics of g_t on $\Lambda = \Lambda_{g_t}$ for a strongly regular parameter t , looking now beyond the well-behaved set $\tilde{\mathcal{R}}_+^\infty$ which was studied in Section 10. In the first part (Subsections 11.1–11.5), we study the intersection of the invariant set Λ with an unstable curve ω^* (defined as a stable curve, exchanging P 's and Q 's). The main part of this intersection is a countable disjoint union of dynamical copies of the set \mathcal{R}_+^∞ studied in Section 10. There are also at most countably many critical points, corresponding to quadratic tangencies between stable curves and images under G of unstable curves. And, finally, there is an exceptional set (formed by points which come very close to the critical locus infinitely many times); but this exceptional set is small; its Hausdorff dimension is explicitly controlled by a value much smaller than the dimension d_s of $\omega^* \cap \Lambda$.

In the second part of Section 11, we prove that the invariant set Λ is a saddle-like object in the metric sense: both its stable set $W^s(\Lambda)$ and its unstable set $W^u(\Lambda)$ have Lebesgue measure 0. So, no attractors nor repellers are present on Λ . One actually expects more: certainly the Hausdorff dimension of $W^s(\Lambda)$ should be strictly less than 2, probably it is close to $1 + d_s$, and perhaps even equal to $1 + d_s$. However, we stick to the simpler, but still very meaningful result: it implies that Λ_g carries no attractor nor a repeller for most g .

One has a nice combinatorial decomposition of the restricted stable set $W^s(\Lambda, R)$, but to compute Lebesgue measure (or Hausdorff dimension), one has to transport the pieces of this decomposition by affine-like iterates of g of high order. This is easy to do as far as Lebesgue measure is concerned, because bounded distortion of affine-like maps mean also bounded distortion of measure (bounded relative oscillation of Jacobians). This is much more delicate with respect to Hausdorff dimension: the geometry of the pieces after iteration can get very distorted.

In Appendix A, we recall all formulas related to the implicit representation of affine-like maps; many of them can already be found in [PY2], but we have also to consider the derivatives with respect to parameter, a setting which was not considered in [PY2].

In Appendix B, we perform some calculations related to proposition 40 in Subsection 10.5, which generates the transversally Lipschitz regularity of the partial foliation $\tilde{\mathcal{R}}_+^\infty$.

In Appendix C, we give some justification for what seems to be a convoluted definition of the transversality relation in Subsection 5.4.

2 Markov Partition and Folding Map

2.1 Markov Partition and Related Charts

We will choose once and for all a finite system of smooth charts

$$I_a^s \times I_a^u \xrightarrow{\approx} R_a \subset M, \quad a \in \mathcal{a}$$

indexed by a finite alphabet \mathcal{a} . Each chart depends smoothly on $g \in \mathcal{U}$; the intervals I_a^s, I_a^u are compact; the rectangles R_a are disjoint.

Let $R = \bigcup_a R_a$. We choose the charts in order to have:

(MP1) for each $g \in \mathcal{U}$, K_g is the maximal invariant set in $\text{int } R$; for each $g \in \mathcal{U}$, $a \in \mathcal{a}$, one has

$$(2.1) \quad g(\partial I_a^s \times I_a^u) \cap R = \emptyset,$$

$$(2.2) \quad g^{-1}(I_a^s \times \partial I_a^u) \cap R = \emptyset;$$

(MP2) for each $g \in \mathcal{U}$, the family $(R_a \cap K_g)_{a \in \mathcal{a}}$ induces a Markov partition for the horseshoe K_g .
Let

$$(2.3) \quad \mathcal{B} = \{(a, a') \in \mathcal{a}^2, f(R_a) \cap R_{a'} \neq \emptyset\}.$$

The Markov partition provides a coding which is a topological conjugacy between the horseshoe K_g and the subshift of finite type of $\mathcal{a}^{\mathbb{Z}}$ defined by \mathcal{B} .

2.2 The Parabolic Tongues L_u, L_s

Denote by $a_s, a_u \in \mathcal{a}$ the letters such that $p_s \in R_{a_s}, p_u \in R_{a_u}$. We choose the corresponding charts in order to have:

(MP3) for each $g \in \mathcal{U}$, the equation of the local stable manifold $W_{\text{loc}}^s(p_s)$ is $\{x_{a_s} = 0\}$, the equation of the local unstable manifold $W_{\text{loc}}^u(p_u)$ is $\{y_{a_u} = 0\}$.

We have written x_a (resp. y_a) for the coordinate in I_a^s (resp. I_a^u). We also choose the rectangles R_a in order to have, for some integer $N_0 \geq 2$:

(MP4) for each $g \in \mathcal{U}_0$, there are points q_s, q_u in the orbit of q such that

- for $n \geq 0$, $g^n(q_s)$ and $g^n(p_s)$ belong to the interior of the same rectangle;
- for $n \leq 0$, $g^n(q_u)$ and $g^n(p_u)$ belong to the interior of the same rectangle;

- $q_s = g^{N_0}(q_u)$ and $g^i(q_u)$ does not belong to R for $0 < i < N_0$.

Consider small pieces of $W^s(p_s)$, $W^u(p_u)$ which are tangent at q_u for $g \in \mathcal{U}_0$. When $g \in \mathcal{U}_+$, these pieces will meet in two points and bound a compact lenticular region $L_u \subset \text{int } R_{a_u}$. Taking the image under g^{N_0} , we get another lenticular region $L_s \subset \text{int } R_{a_s}$. These regions are called *parabolic tongues*. See figure 3.

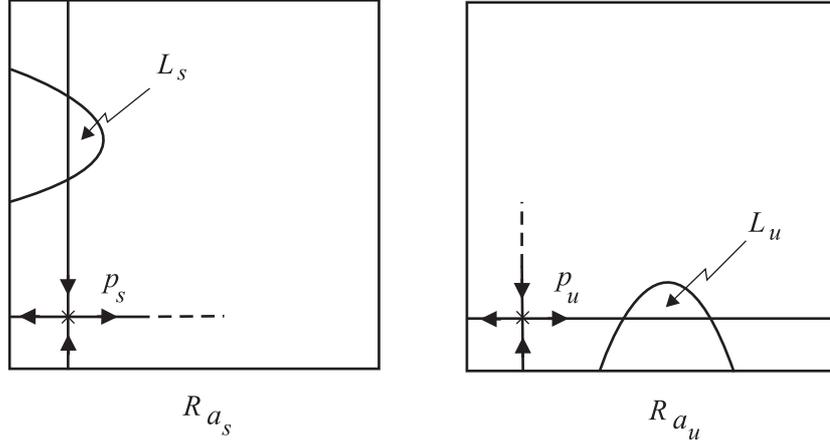


Figure 3

Define then, for $g \in \mathcal{U}_+$

$$(2.4) \quad \widehat{R} = R \bigcup_{0 < i < N_0} g^i(L_u)$$

The maximal invariant set we are interested in is

$$(2.5) \quad \Lambda_g = \bigcap_{\mathbb{Z}} g^{-n}(\widehat{R}).$$

We also define

$$(2.6) \quad W^s(\Lambda_g, \widehat{R}) = \bigcap_{n \geq 0} g^{-n}(\widehat{R}),$$

$$(2.7) \quad W^u(\Lambda_g, \widehat{R}) = \bigcap_{n \leq 0} g^{-n}(\widehat{R})$$

The dynamics in \widehat{R} are generated by

- the transition maps related to the Markov partition:

$$g : R_a \cap g^{-1}(R_{a'}) \rightarrow g(R_a) \cap R_{a'}, \text{ for } (a, a') \in \mathcal{B};$$

- the folding map $G := g^{N_0}$ from L_u onto L_s .

2.3 The Folding Map G

For simplicity, we write $(x_s, y_s), (x_u, y_u)$ for the coordinates in $R_{a_s} \supset L_s, R_{a_u} \supset L_u$.

The folding map G is *parabolic* in the sense of [PY2]; let us recall this definition.

Consider the graph Γ_G of the restriction G of g^{N_0} to the component of $R_{a_u} \cap g^{-N_0}(R_{a_s})$ which contains L_u (for $g \in \mathcal{U}_+$; we then follow this component in the rest of \mathcal{U}). Using the corresponding charts, we can view Γ_G as a surface in $I_{a_u}^s \times I_{a_u}^u \times I_{a_s}^s \times I_{a_s}^u$. Denote by π the projection from $I_{a_u}^s \times I_{a_u}^u \times I_{a_s}^s \times I_{a_s}^u$ onto $I_{a_u}^u \times I_{a_s}^s$. For \mathcal{U} small enough, from the quadratic tangency at q and (MP3) we deduce that:

(P1) the restriction of π to Γ_G is a fold map (in the sense of singularity theory).

Denote by $\Gamma_0 \subset I_{a_u}^u \times I_{a_s}^s$ the smooth curve which is the image of the critical locus of this fold map. It divides $I_{a_u}^u \times I_{a_s}^s$ into two regions Γ_+, Γ_- such that $\Gamma_+ \cup \Gamma_0$ is the image of the fold map. We can reformulate (P1) as:

(P'1) (i) for $(y^0, x^0) \in \Gamma_0$, the image $G(\{y_u = y^0\})$ meets $\{x_s = x^0\}$ in a single point, interior to both curves, at which the curves have a quadratic tangency;

(ii) for $(y^0, x^0) \in \Gamma_-$, the curves $G(\{y_u = y^0\})$ and $\{x_s = x^0\}$ do not intersect;

(iii) for $(y^0, x^0) \in \Gamma_+$, the curves $G(\{y_u = y^0\})$ and $\{x_s = x^0\}$ intersect transversally in two points.

As G is a diffeomorphism, the tangents to Γ_0 are never vertical or horizontal. Therefore, we can and will choose a smooth function θ on $I_{a_u}^u \times I_{a_s}^s$ such that

(P2) $\theta \equiv 0$ on Γ_0 , $\theta > 0$ on Γ_+ , $\theta < 0$ on Γ_- ;

(P3) the partial derivatives θ_y, θ_x of θ do not vanish on $I_{a_u}^u \times I_{a_s}^s$.

Remark. *The choice of θ is far from unique. One could for instance choose θ of the form*

$$(2.8) \quad \theta(y_u, x_s) = \varepsilon_u y_u + \varepsilon_s \chi(x_s),$$

with $\varepsilon_s, \varepsilon_u \in \{-1, +1\}$ and χ monotone increasing. We prefer not to specify a particular choice in order to keep a time-symmetric setting between positive and negative iterations.

From θ , we define a smooth function w on Γ_G by

(P4) $w^2 = \theta \circ \pi$

(there are two choices for w ; the other is $-w$).

Then, from (P3) we obtain smooth maps Y_u, X_s implicitly defined by

$$(P5) \quad \begin{aligned} w^2 &= \theta(Y_u(w, x_s), x_s) \\ &= \theta(y_u, X_s(w, y_u)) \end{aligned}$$

On the graph Γ_G , we can use either (x_u, y_u) or (x_s, y_s) or (w, y_u) or (x_s, w) as coordinates; therefore we can factorize G as $G_+ \circ G_0 \circ G_-$:

$$(2.9) \quad (x_u, y_u) \xrightarrow{G_-} (w, y_u) \xrightarrow{G_0} (x_s, w) \xrightarrow{G_+} (x_s, y_s)$$

with

$$(P6) \quad \begin{aligned} G_0(w, y_u) &= (X_s(w, y_u), w), \\ G_0^{-1}(x_s, w) &= (w, Y_u(w, x_s)), \\ G_+(x_s, w) &= (x_s, Y_s(w, x_s)), \\ G_-^{-1}(w, y_u) &= (X_u(w, y_u), y_u). \end{aligned}$$

The last two formulas define smooth maps Y_s, X_u and the partial derivatives $Y_{s,w}, X_{u,w}$ do not vanish as G_+, G_- are diffeomorphisms. Observe that the map G_0 is very similar to a quadratic Hénon-like map.

3 Affine-like Maps

This section is essentially a summary of [PY2].

3.1 Definition and Implicit Representation

Let $I_0^s, I_0^u, I_1^s, I_1^u$ be non trivial compact intervals, x_0, y_0, x_1, y_1 the corresponding coordinates. Consider a smooth diffeomorphisms F whose domain is a *vertical strip*

$$P = \{\varphi^-(y_0) \leq x_0 \leq \varphi^+(y_0)\} \subset I_0^s \times I_0^u$$

and whose image is a *horizontal strip*

$$Q = \{\psi^-(x_1) \leq y_1 \leq \psi^+(x_1)\} \subset I_1^s \times I_1^u.$$

We say that F is *affine-like* if

(AL1) the restriction to the graph of F of the projection onto $I_0^u \times I_1^s$ is a diffeomorphism onto $I_0^u \times I_1^s$.

This allow us to define smooth maps A, B on $I_0^u \times I_1^s$ such that

$$(3.1) \quad F(x_0, y_0) = (x_1, y_1) \iff \begin{cases} x_0 = A(y_0, x_1) \\ y_1 = B(y_0, x_1). \end{cases}$$

The pair (A, B) is the *implicit representation* (or definition) of the affine-like map F . See figure 4. In the formulas below, we shall most of the time omit the arguments of the functions considered, which should be obvious from the context. We will write $A_x, A_y, A_{xx}, B_x, B_y \dots$ for partial derivatives.

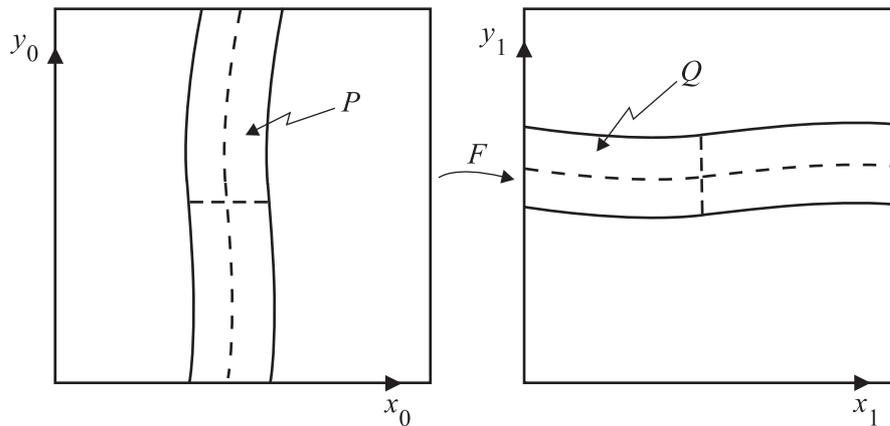


Figure 4

On the graph of F , we have

$$(3.2) \quad \begin{aligned} dx_0 &= A_y dy_0 + A_x dx_1, \\ dy_1 &= B_y dy_0 + B_x dx_1, \end{aligned}$$

which leads to

$$(3.3) \quad DF = A_x^{-1} \begin{pmatrix} 1 & -A_y \\ B_x & A_x B_y - A_y B_x \end{pmatrix}$$

$$(3.4) \quad DF^{-1} = B_y^{-1} \begin{pmatrix} A_x B_y - A_y B_x & A_y \\ -B_x & 1 \end{pmatrix}$$

$$(3.5) \quad \det DF = A_x^{-1} B_y.$$

The main advantage of the implicit representation is the symmetry between positive and negative iteration.

3.2 Cone Condition and Distortion

Let $\lambda, u, v > 0$ satisfy

$$(3.6) \quad 1 < uv \leq \lambda^2.$$

Let (X_0, Y_0) be a tangent vector at some point in the domain of F , and let (X_1, Y_1) be its image under TF . The usual cone condition with parameters (λ, u, v) is:

(AL2) (i) if $|Y_0| \leq u|X_0|$, then $|Y_1| \leq v^{-1}|X_1|$ and $|X_1| \geq \lambda|X_0|$;

(ii) if $|X_1| \leq v|Y_1|$, then $|X_0| \leq u^{-1}|Y_0|$ and $|Y_0| \geq \lambda|Y_1|$.

This is readily seen to be equivalent to

$$(AL'2) \quad \begin{aligned} \lambda|A_x| + u|A_y| &\leq 1, \\ \lambda|B_y| + v|B_x| &\leq 1, \end{aligned}$$

everywhere on $I_0^u \times I_1^s$.

We will also need to control partial derivatives of second order of A, B . By (3.5), the partial derivatives A_x, B_y do not vanish on $I_0^u \times I_1^s$. It turns out that the right way to look at partial derivatives of second order is to consider the six functions

$$\begin{aligned} \partial_x \log|A_x|, \partial_y \log|A_x|, A_{yy}, \\ \partial_y \log|B_y|, \partial_x \log|B_y|, B_{xx}. \end{aligned}$$

We define the *distortion* of an affine-like map F , and denote by $D(F)$, the maximal absolute value attained by any one of these six functions on $I_0^u \times I_1^s$.

We also define the *width* of the domain P of F by

$$(3.7) \quad |P| := \max|A_x|,$$

and the width of the image Q by

$$(3.8) \quad |Q| := \max|B_y|.$$

3.3 Simple Composition

The composition of two affine-like maps is not always affine-like. However, the composition of two affine-like maps which also satisfy the same cone condition (AL2) will again be affine-like and satisfy the same cone condition (actually a better one).

More precisely, let $I_0^s, I_0^u, I_1^s, I_1^u, I_2^s, I_2^u$ be compact intervals. Let $F : P \rightarrow Q$ and $F' : P' \rightarrow Q'$ be affine-like maps with domains $P \subset I_0^s \times I_0^u, P' \subset I_1^s \times I_1^u$ and images $Q \subset I_1^s \times I_1^u, Q' \subset I_2^s \times I_2^u$. We assume that both F and F' satisfy (AL2) (or (AL'2)) with parameters λ, u, v . The composition $F'' = F' \circ F$ has domain $P'' = P \cap F^{-1}(P')$ and image $Q'' = Q' \cap F'(Q)$. It satisfies (AL1) and (AL2) with parameters λ^2, u, v (cf. [PY2]). See figure 5.

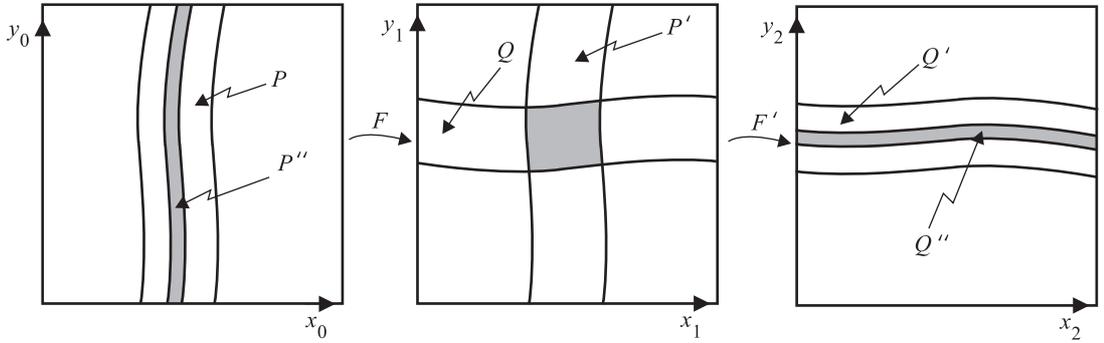


Figure 5

Let $(A, B), (A', B'), (A'', B'')$ be the implicit representations of F, F', F'' respectively. Define

$$(3.9) \quad \Delta := 1 - A'_y B_x > 1 - u^{-1} v^{-1} > 0.$$

The partial derivatives of first order of A'', B'' are given by

$$(3.10) \quad \begin{aligned} A''_x &= A_x A'_x \Delta^{-1}, \\ B''_y &= B_y B'_y \Delta^{-1}, \end{aligned}$$

$$(3.11) \quad \begin{aligned} A_y'' &= A_y + A_y' A_x B_y \Delta^{-1}, \\ B_x'' &= B_x' + B_x A_x' B_y' \Delta^{-1}. \end{aligned}$$

From (3.10), we get

$$(3.12) \quad \begin{aligned} C^{-1} &\leq \frac{|P''|}{|P| |P'|} \leq C, \\ C^{-1} &\leq \frac{|Q''|}{|Q| |Q'|} \leq C, \end{aligned}$$

where the constants are uniform once u, v are fixed and the distortions are uniformly bounded.

The formulas for the partial derivatives of second order are derived in [PY2] and recalled in Appendix A. They lead to the following estimate for the distortion:

$$(3.13) \quad D(F'') \leq \max \left\{ D(F) + C|Q|(D(F) + D(F')), D(F') + C|P'|(D(F) + D(F')) \right\},$$

where C depends only on u, v .

3.4 Properties of the Markov Partition

We choose charts for the Markov partition discussed in Subsection 2.1 in order to have the following property, for some λ, u, v satisfying (3.6):

(MP5) for any $(a, a') \in \mathcal{B}$, any $g \in \mathcal{U}$, the transition map $g_{a,a'}$ from $P_{aa'} = R_a \cap g^{-1}(R_{a'})$ onto $Q_{aa'} = R_{a'} \cap g(R_a)$ is affine-like and also satisfies the cone condition (AL2).

These values of (λ, u, v) will be fixed in what follows.

To any finite word $\underline{a} = (a_0, \dots, a_n)$ with transitions in \mathcal{B} , we have a composition

$$g_{\underline{a}} = g_{a_{n-1}a_n} \circ \dots \circ g_{a_0a_1}$$

which satisfies also (AL1) and (AL2).

Moreover, as the widths decrease exponentially with the number of iterations, it follows from (3.13) that there exists $D_0 > 0$ such that all $g_{\underline{a}}$ satisfy

(MP6) $D(g_{\underline{a}}) \leq D_0$.

3.5 Parabolic Composition

Let G be the folding map of Subsection 2.3, satisfying properties (P1)–(P6).

Let also $I_0^s, I_0^u, I_1^s, I_1^u$ be compact intervals ; let F_0 be an affine-like map from a vertical strip $P_0 \subset I_0^s \times I_0^u$ to a horizontal strip $Q_0 \subset I_{a_u}^s \times I_{a_u}^u$; let F_1 be an affine-like map from a vertical strip $P_1 \subset I_{a_s}^s \times I_{a_s}^u$ to a horizontal strip $Q_1 \subset I_1^s \times I_1^u$.

We recall from [PY2] how, under appropriate hypotheses, the composition $F_1 \circ G \circ F_0$ defines two affine-like maps F^\pm with domains $P^\pm \subset P_0$ and image $Q^\pm \subset Q_1$. See figure 6.

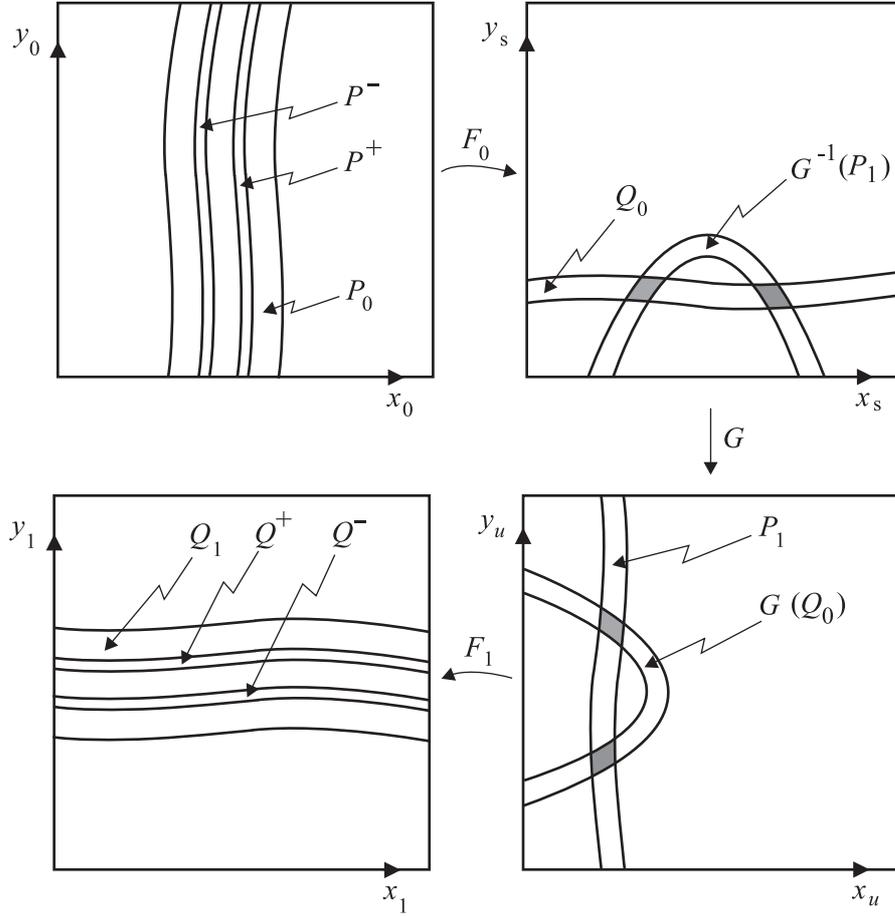


Figure 6

Let $(A_0, B_0), (A_1, B_1)$ be implicit representations of F_0, F_1 , respectively. We assume that

$$\begin{aligned}
 \text{(PC1)} \quad & |A_{1,y}| < b, & |A_{1,yy}| < b, \\
 & |B_{0,x}| < b, & |B_{0,xx}| < b,
 \end{aligned}$$

with $b \ll 1$. In the system

$$\begin{aligned}
 (3.14) \quad & x_u = X_u(w, y_u), \\
 & y_u = B_0(y_0, x_u),
 \end{aligned}$$

we can, as $|B_{0,x}| \ll 1$, eliminate y_u and solve for x_u to define

$$(3.15) \quad x_u = X(w, y_0)$$

Similarly, in the system

$$(3.16) \quad \begin{aligned} y_s &= Y_s(w, x_s), \\ x_s &= A_1(y_s, x_1), \end{aligned}$$

we eliminate x_s , solve for y_s to get

$$(3.17) \quad y_s = Y(w, x_1).$$

The next step is to define

$$(3.18) \quad C(w, y_0, x_1) := w^2 - \theta\left(B_0(y_0, X(w, y_0)), A_1(Y(w, x_1), x_1)\right).$$

This quantity has the following geometrical interpretation. Fix values y_0^*, x_1^* for y_0, x_1 . The image $G_- \circ F_0(\{y_0 = y_0^*\})$ is the graph

$$(3.19) \quad \gamma_0 = \left\{ y_u = B_0(y_0^*, X(w, y_0^*)) \right\};$$

symmetrically, $G_+^{-1} \circ F_1^{-1}(\{x_1 = x_1^*\})$ is the graph

$$(3.20) \quad \gamma_1 = \left\{ x_s = A_1(Y(w, x_1^*), x_1^*) \right\}.$$

Then, $C(w, y_0^*, x_1^*)$ gives the relative position of the two curves γ_0 and $G_0^{-1}(\gamma_1)$ (or equivalently $G_0(\gamma_0)$ and γ_1). More precisely, it is positive for all w if the two curves do not intersect; it vanishes at the intersection points and is negative between the intersection points.

It follows from (PC1) just above that

$$(3.21) \quad |C_w - 2w| \ll 1,$$

$$(3.22) \quad |C_{ww} - 2| \ll 1.$$

Therefore, for fixed values of y_0 and x_1 , C has a unique minimum as a function of w ; we denote by $\bar{C}(y_0, x_1)$ the corresponding minimum value. We have $\bar{C}(y_0^*, x_1^*) > 0$ (resp. $= 0$, resp. < 0) if and only if the curves γ_0 and $G_0^{-1}(\gamma_1)$ do not intersect (resp. are tangent, resp. have two transverse intersection points).

In order to consider parabolic compositions, we shall require that $\bar{C}(y_0, x_1) < 0$ everywhere on $I_0^u \times I_1^s$. Setting

$$(3.23) \quad \delta = \min_{y_0, x_1} -\bar{C}(y_0, x_1)$$

we actually want to have

$$(PC2) \quad \delta > b^{-1}(|P_1| + |Q_0|).$$

The geometric interpretation of this requirement is clear: the displacement of one of the rectangles and the image of the other should be much bigger than the sum of their widths. In other words, the distance between the tip of the parabolic strip $G_0^{-1}(P_1)$ and the horizontal strip Q_0 should be much bigger than the widths of these strips.

Assume now that (PC1) and (PC2) are satisfied; the equation $C(w, y_0, x_1) = 0$ defines two smooth functions

$$(3.24) \quad w = W^\pm(y_0, x_1)$$

with $w^+ > w^-$. One then defines

$$(3.25) \quad A^\pm(y_0, x_1) := A_0\left(y_0, X(W^\pm(y_0, x_1), y_0)\right),$$

$$(3.26) \quad B^\pm(y_0, x_1) := B_1\left(Y(W^\pm(y_0, x_1), x_1), x_1\right).$$

As shown in [PY2], the pair (A^+, B^+) (resp. (A^-, B^-)) implicitly defines an affine-like map F^+ (resp. F^-).

Denote by P^+ (resp. P^-) the domain of F^+ and by Q^+ (resp. Q^-) the domain of F^- . Then P^+ and P^- are the two components of $P_0 \cap (G \circ F_0)^{-1}(P_1)$, Q^+ and Q^- are the two components of $Q_1 \cap (F_1 \circ G)(Q_0)$; F^+ (resp. F^-) is the restriction of $F_1 \circ G \circ F_0$ to P^+ (resp. P^-).

The formulas for the partial derivatives of A^\pm, B^\pm are derived in [PY2] and recalled in Appendix A. They provide the following estimate for the widths:

$$(3.27) \quad C^{-1} \leq \frac{|P^\pm|}{|P_0| |P_1| \delta^{-\frac{1}{2}}} \leq C,$$

$$(3.28) \quad C^{-1} \leq \frac{|Q^\pm|}{|Q_0| |Q_1| \delta^{-\frac{1}{2}}} \leq C,$$

where the constants are uniform once b is fixed and the distortions are uniformly bounded.

From [PY2, Theorem 3.7], we also have the following estimate for the distortion of F^\pm : assuming that b is small enough (in terms of the partial derivatives of first order of X_u, Y_s, θ), we have

$$(3.29) \quad D(F^\pm) \leq \max\left\{D(F_0) + C|Q_0|\delta^{-1}, D(F_1) + C|P_1|\delta^{-1}\right\},$$

provided that $D(F_0) + D(F_1) \leq \delta^{-\frac{1}{2}}$. The constant C in (3.29) depends only on the partial derivatives of first order of X_s, Y_u, θ .

We also recall from [PY2, formula 3.50] the estimate:

$$(3.30) \quad |A_y^\pm - A_{0,y}| \leq C|P_0| |Q_0| \delta^{-\frac{1}{2}},$$

$$(3.31) \quad |B_x^\pm - B_{1,x}| \leq C|P_1| |Q_1| \delta^{-\frac{1}{2}},$$

where the right-hand terms must be small by (PC2).

As a concluding remark for this section, let us observe that, while conditions (PC1), (PC2) are *necessary* in order to consider parabolic composition, they are not *sufficient*: in Section 5, the requirement for parabolic composition will be much more restrictive than (PC2).

4 Structure of Parameter Space

4.1 Some Important Constants

Throughout the rest of the paper, we will use four main constants $\varepsilon_0, \eta, \tau, \beta$ which satisfy

$$(4.1) \quad 0 < \varepsilon_0 \ll \eta \ll \tau \ll \beta - 1 < 1.$$

We roughly explain the meaning of each constant:

- ε_0 is the maximal width of the parabolic tongues L_u, L_s . It is also the size of the parameter interval we start with.

- η is involved in the transversality relation (defined in Section 5) which allows parabolic composition: instead of the condition (PC2) of Subsection 3.5, roughly speaking we will ask that

$$(4.2) \quad \delta \geq (|P_1| + |Q_0|)^{1-\eta}.$$

- τ relates the successive scales of the parameter intervals we will consider through the formula $\varepsilon_{k+1} = \varepsilon_k^{1+\tau}$.

- β will actually be given in Section 9 by an explicit formula in terms of d_s^0, d_u^0 ; the condition (H4) involving d_s^0, d_u^0 in Section 1 is required because we need $\beta > 1$. It appears in the definition of regularity in Section 5, which controls the recurrence of the "critical locus".

4.2 One-Parameter Families

From now on, we fix a one-parameter family $(g_t)_{t \in (-t_0, t_0)}$ in \mathcal{U} . We assume that the family is transverse to \mathcal{U}_0 at $t = 0$, with $g_t \in \mathcal{U}_+$ for $t > 0$ and $g_t \in \mathcal{U}_-$ for $t < 0$.

Observe that g_0 satisfies exactly the same assumptions as f , provided \mathcal{U} is small enough. Therefore, we may and shall, assume that $g_0 = f$.

We will first reparametrize the family in order to make some computations simpler. Consider the folding map $G_t = g_t^{N_0}$ of Subsection 2.3. If t_0 is small enough, G_t is a fold map for all values of $t \in (-t_0, t_0)$. Moreover, we can in properties (P2), (P3) of Subsection 2.3 choose a function θ which depends smoothly on t .

From (MP3), Subsection 2.2, the values $y_u = 0, x_s = 0$ of the arguments of θ correspond to $W_{\text{loc}}^u(p_u)$ and $W_{\text{loc}}^s(p_s)$ respectively. Therefore, the transversality of the family to \mathcal{U}_0 is equivalent to

$$(4.3) \quad \frac{\partial}{\partial t} \theta_t(0, 0) |_{t=0} > 0.$$

Taking t_0 small enough, we can therefore reparametrize our family in order to have

$$(4.4) \quad \theta_t(0, 0) \equiv t, \quad t \in (-t_0, t_0).$$

4.3 Parameter Intervals

The starting parameter interval will be

$$(4.5) \quad I_0 := [\varepsilon_0, 2\varepsilon_0],$$

where, as explained above, ε_0 will be taken very small. This is the only parameter interval at level 0.

At level k , we will deal with parameter intervals of length ε_k , where the sequence of scales ε_k is defined inductively by

$$(4.6) \quad \varepsilon_{k+1} = \varepsilon_k^{1+\tau}.$$

The constant τ is small, but ε_0 is much smaller and in particular we will have $\varepsilon_0^{\tau^2} \ll 1$. Every parameter interval of level k is divided into $[\varepsilon_k^{-\tau}]$ parameter intervals of level $k+1$.

The remaining part, if any, is discarded; it is of length $< \varepsilon_{k+1}$; the total length discarded in this way is smaller than $\varepsilon_1 \ll \varepsilon_0$.

Let \tilde{I} be a parameter interval of level k and I be a parameter interval of level $k+1$ contained in \tilde{I} . We say that \tilde{I} is the *parent* of I and that I is a *child* of \tilde{I} .

4.4 The Selection Process

In Section 5, we will define what it means for a parameter interval to be *regular*. The starting interval I_0 will be regular.

Given a regular parameter interval \tilde{I} of level k , we divide it into its children: these parameter intervals of level $k+1$ are the *candidates*. We then test each candidate for regularity and discard those which are not regular. We then proceed to level $k+1$ with each surviving candidate.

The *regular* parameters are those which are the intersection of a decreasing sequence of regular parameter intervals. For such parameters, we are able to carry out some analysis of the maximal invariant set Λ_{g_t} .

4.5 Strongly Regular Parameters

The regularity property is, in some sense, the minimal requirement that is needed to keep control on the geometry and dynamics of the maximal invariant set. However, this requirement is of an essentially qualitative character and this leads in particular to the following difficulty: we are not able to estimate which proportion of the children of a regular parameter interval are also regular.

To circumvent this problem, we define in Section 9 a stronger property for parameter intervals, called *strong regularity*. It implies regularity, and is better adapted to the inductive selection process. It also gives additional geometric information on the maximal invariant set.

When \tilde{I} is a strongly regular parameter interval of level k , we will show in Section 9 that most candidates of level $k + 1$ contained in \tilde{I} are also strongly regular. The proportion of discarded candidates is less than α_k , with

$$(4.7) \quad \sum_{k \geq 0} \alpha_k \ll 1,$$

the \ll sign means that the sum gets arbitrarily small as ε_0 goes to zero. Then we can conclude that most parameters are strongly regular in the sense that they are equal to the intersection of decreasing sequences of strongly regular parameter intervals.

The non-uniformly hyperbolic horseshoes that are the subject of our study are exactly the maximal invariant set Λ_g for strongly regular $g \in \mathcal{U}_+$.

5 Classes of Affine-like Iterates and the Transversality Relation

5.1 Affine-like Iterates

Let I be a parameter interval of some level.

Definition. An I -persistent affine-like iterate is a triple (P, Q, n) such that

- P is a vertical strip in some R_a , depending smoothly on $t \in I$;
- Q is a horizontal strip in some $R_{a'}$, depending smoothly on $t \in I$;
- n is a nonnegative integer;
- for each $t \in I$, the restriction of g_t^n to P_t is an affine-like map onto Q_t , i.e. property (AL1) of Subsection 3.1 holds;
- for each $t \in I$, each $m \in [0, n]$, we have $g_t^m(P_t) \subset \widehat{R}$.

Examples.

1. For $n = 0$, the I -persistent affine-like iterates are the $(R_a, R_a, 0)$, $a \in \mathcal{A}$.
2. For $n = 1$, the I -persistent affine-like iterates are the $(P_{aa'}, Q_{aa'}, 1)$, $(a, a') \in \mathcal{B}$.
3. More generally, for any finite word $\underline{a} = (a_0, \dots, a_n)$ with transitions in \mathcal{B} , the map $g_{\underline{a}}$ of Subsection 3.4 defined an I -persistent affine-like iterate $(P_{\underline{a}}, Q_{\underline{a}}, n)$.

Notation. If P is a vertical strip $\{\varphi_-(y) \leq x \leq \varphi_+(y)\}$ we denote by ∂P the vertical part of the boundary, i.e. the two graphs $\{x = \varphi^\pm(y)\}$. Similarly for horizontal strips.

If (P, Q, n) is an I -persistent affine-like iterate and I' is a parameter interval contained in I , (P, Q, n) also defines by restriction an I' -persistent affine-like iterate. A slightly less trivial property is given by

Proposition 1. *Let $(P, Q, n), (P', Q', n')$ be I -persistent affine-like iterates. We have*

- a) *if $n = n'$, then either $P = P'$ and $Q = Q'$ for all $t \in I$ or $P \cap P' = \emptyset, Q \cap Q' = \emptyset$ for all $t \in I$.*
- b) *if $n < n'$, then either $P \supset P', \partial P \cap P' = \emptyset$ for all $t \in I$ or $P \cap P' = \emptyset$, for all $t \in I$.*

Remark. *Throughout the paper, except in Section 9 (where we break the symmetry assuming $d_s^0 \geq d_u^0$), we will keep a time-symmetric setting. Thus every property stated for the domains P 's is also valid for the images Q 's. This apply for instance to part b) of the proposition.*

Proof. By the definition of an I -persistent affine-like iterate, for all $t \in I$, P is a connected component of $R \cap g_t^{-n}(R)$ and also of $\bigcap_{0 \leq m \leq n} g_t^{-m}(\widehat{R})$.

a) If $n = n'$ and $P \cap P' \neq \emptyset$ for some $t_0 \in I$, we must have $P = P'$ at t_0 and hence, $P \cap P' \neq \emptyset$ for t close to t_0 . It follows that $P = P'$ for all $t \in I$, and also $Q = Q'$ for all $t \in I$.

b) Assume that $n < n'$ and $P \cap P' \neq \emptyset$ for some $t_0 \in I$, then $P' \subset P$ at t_0 (since $\bigcap_{0 \leq m \leq n'} g_{t_0}^{-m}(\widehat{R})$ is contained in $\bigcap_{0 \leq m \leq n} g_{t_0}^{-m}(\widehat{R})$), hence $P' \cap P \neq \emptyset$ for t close to t_0 and $P' \subset P$ for all $t \in I$.

Let $t \in I$, $z \in \partial P$; then, $g_t^n(z)$ belongs to the vertical boundary of R and $g_t^{n+1}(z) \notin \widehat{R}$; therefore, $z \notin P'$. This proves that $\partial P \cap P'$ is empty for all $t \in I$. \square

5.2 The Classes $\mathcal{R}(I)$: General Overview

It would be nice to work with the class of all I -persistent affine-like iterates, but with this approach one faces two problems:

- I -persistent affine-like iterates do not satisfy a uniform cone condition, and they do not have uniformly bounded distortion;
- even if we force such uniformity in the definition, a major problem is that we lack some control on the way in which long I -persistent affine-like iterates are constructed from shorter ones by simple or parabolic composition.

To overcome these problems, we will define, for every parameter interval (whose parent is regular, see Subsection 4.4 and the end of the present section on the definition of regularity) a subset $\mathcal{R}(I)$ of the set of all I -persistent affine-like iterates; all elements of $\mathcal{R}(I)$ with $n > 1$ can be, from the very definition, obtained from shorter ones by simple or parabolic composition; the elements of $\mathcal{R}(I)$ will turn out to satisfy a uniform cone condition and have uniformly bounded distortion.

The main ingredient in the definition of $\mathcal{R}(I)$ is a *transversality relation* which is an appropriate strengthening of condition (PC2) in Subsection 3.5. Simple composition is allowed whenever it makes sense, but parabolic composition is only allowed when this transversality relation holds.

The definition of the transversality relation, given later in this section, is quite involved; this is because we want some combinatorial properties proved in Section 6 to be satisfied. Such properties make our later work much easier.

All the process of this definition is based on a double induction:

- an induction on the level of the parameter interval, starting with $I_0 = [\varepsilon_0, 2\varepsilon_0]$ at level 0;
- for a given parameter interval, an induction on the length n of the I -persistent affine-like iterates which are considered.

In order for the definition of $\mathcal{R}(I)$ to be consistent, a number of properties (uniform cone condition, uniform bound on distortion and many others) must hold; unfortunately, the proof of any of these properties is inductive, using several *other* properties for shorter iterates. Our presentation, is therefore, organized through the following scheme: we will first define $\mathcal{R}(I)$ and list all the required properties conditioning the definition. Then we will proceed to the proofs of these properties, and of other related ones. The whole process will occupy the rest of the present section and Sections 6 and 7.

5.3 Definition of the Special Class of Affine-Like Iterates $\mathcal{R}(I)$

Let I be a parameter interval of some level.

We claim that there exists a (unique) class of I -persistent affine-like iterates which satisfy the properties (R1)–(R7) below.

(R1) For any word $\underline{a} = (a_0, \dots, a_n)$ with transitions in \mathcal{B} , the element $(P_{\underline{a}}, Q_{\underline{a}}, n)$ (see example 3 above) belongs to $\mathcal{R}(I)$.

For the starting interval $I_0 = [\varepsilon_0, 2\varepsilon_0]$, it will turn out that one obtains in this way all elements of $\mathcal{R}(I_0)$.

Recall from (MP5), (MP6) in Subsection 3.4 that all $(P_{\underline{a}}, Q_{\underline{a}}, n)$ with $n > 0$ satisfy for all $t \in I_0$ a uniform cone condition (AL2) with parameters λ, u, v (satisfying $1 < uv \leq \lambda^2$), and have distortion bounded by D_0 . Let $u_0 = u^{\frac{1}{2}}, v_0 = v^{\frac{1}{2}}$.

(R2) All $(P, Q, n) \in \mathcal{R}(I)$ satisfy for all $t \in I$ the cone condition (AL2) with parameters λ, u_0, v_0 and have distortion bounded by $2D_0$ for all $t \in I$.

Let $(P, Q, n), (P', Q', n')$ be elements of $\mathcal{R}(I)$ such that $Q \subset R_a, P' \subset R_a$ for some $a \in \mathcal{A}$. As both iterates satisfy the cone condition (AL2) with parameters λ, u_0, v_0 , we know from Subsection 3.3 that the simple composition defined by

$$(5.1) \quad P'' = P \cap g^{-n}(P'), \quad Q'' = Q' \cap g^{n'}(Q), \quad n'' = n + n',$$

is an (I -persistent) affine-like iterate.

The next condition states that it should also belong to $\mathcal{R}(I)$.

(R3) The class $\mathcal{R}(I)$ is stable under simple composition.

We now turn to parabolic composition.

We first define two special elements which belong to $\mathcal{R}(I)$ according to (R1): define (P_s, Q_s, n_s) (resp. (P_u, Q_u, n_u)) to be the element (P_a, Q_a, n) with maximal length n such that $L_s \subset P_s$ for all $t \in I_0$ (resp. $L_u \subset Q_u$ for all $t \in I_0$). We have that $p_s \in P_s$ and $p_u \in Q_u$. See figure 7.

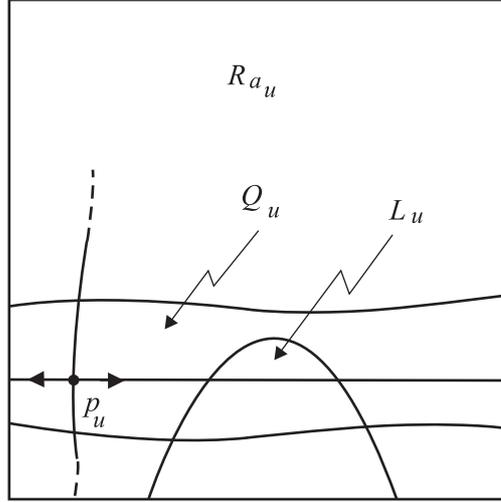


Figure 7

We obviously have, for all $t \in I_0$

$$(5.2) \quad \begin{aligned} C^{-1}\varepsilon_0 &\leq |P_s| \leq C\varepsilon_0, \\ C^{-1}\varepsilon_0 &\leq |Q_u| \leq C\varepsilon_0. \end{aligned}$$

The next condition guarantees that property (PC1) in Subsection 3.5 is satisfied.

(R4) Let (A, B) be the implicit representation of an affine-like iterate $(P, Q, n) \in \mathcal{R}(I)$.

a) If $P \subset P_s$, then for all $t \in I$ we have

$$|A_y| \leq C\varepsilon_0, \quad |A_{yy}| \leq C\varepsilon_0.$$

b) If $Q \subset Q_u$, then for all $t \in I$ we have

$$|B_x| \leq C\varepsilon_0, \quad |B_{xx}| \leq C\varepsilon_0.$$

Here and in the sequel, the letter C denotes various constants which depend only on our initial diffeomorphism f , but *not* on $\eta, \tau, \varepsilon_0$.

Let $(P_0, Q_0, n_0), (P_1, Q_1, n_1)$ be elements in $\mathcal{R}(I)$ with $Q_0 \subset Q_u, P_1 \subset P_s$. In these circumstances, we will define in Subsection 5.4 a *transversality relation* denoted by $Q_0 \pitchfork_I P_1$ which may or may not hold. When it holds, it implies condition (PC2) of Subsection 3.5 for all $t \in I$ (see (R7) below).

(R5) If $(P_0, Q_0, n_0), (P_1, Q_1, n_1)$ as above satisfy $Q_0 \pitchfork_I P_1$, then both I -persistent affine-like iterates obtained from the parabolic composition $g_t^{n_1} \circ G_t \circ g_t^{n_0}$ belong to $\mathcal{R}(I)$.

Writing (P^+, Q^+, n) and (P^-, Q^-, n) for these two iterates, we have $n = n_0 + n_1 + N_0$. The domains P^+ and P^- are the two components of $g^{-n_0}(Q_0 \cap G_t^{-1}(P_1))$; the images Q^+ and Q^- are the two components of $g^{n_1}(P_1 \cap G_t(Q_0))$. See figure 6, Subsection 3.5.

When $(P_0, Q_0, n_0), (P_1, Q_1, n_1)$ satisfy $Q_0 \subset Q_u, P_1 \subset P_s, Q_0 \pitchfork_I P_1$, we say that their parabolic composition is allowed in $\mathcal{R}(I)$.

(R6) Any $(P, Q, n) \in \mathcal{R}(I)$ with $n > 1$ can be obtained from shorter elements by simple composition or (allowed) parabolic composition.

Typically, an element of $\mathcal{R}(I)$ can be obtained in many ways by composition of shorter ones. We say that an element of $\mathcal{R}(I)$ is *prime* if it cannot be obtained by *simple* composition of shorter ones. Prime elements play a key role in the description of the dynamics for regular parameters in Section 10.

It is pretty clear from conditions (R1), (R3), (R5), (R6) alone that there is at most one class $\mathcal{R}(I)$ satisfying these conditions. The existence of $\mathcal{R}(I)$, i.e. the proof of the consistency of conditions (R1)–(R6), is much more delicate. There is actually a seventh property (R7) formulated in the next subsection and related to the condition (PC2) for parabolic composition.

Parent-Child Terminology and Notations for Compositions.

Let $(P, Q, n), (\tilde{P}, \tilde{Q}, \tilde{n}) \in \mathcal{R}(I)$ with $P \subset \tilde{P}, n > \tilde{n}$. If there is no $(\hat{P}, \hat{Q}, \hat{n}) \in \mathcal{R}(I)$ with $P \subset \hat{P} \subset \tilde{P}$ and $n > \hat{n} > \tilde{n}$, we say that P is a *child* of \tilde{P} and that \tilde{P} is the *parent* of P ; if moreover $n = \tilde{n} + 1$, we say that P is a *simple* child; otherwise it is a non-simple child.

Let $(P_0, Q_0, n_0), (P_1, Q_1, n_1) \in \mathcal{R}(I)$. If Q_0, P_1 are contained in a same rectangle R_a , the simple composition $(P, Q, n) \in \mathcal{R}(I)$ of these elements will be written as

$$(5.3) \quad (P, Q, n) = (P_0, Q_0, n_0) * (P_1, Q_1, n_1),$$

If $Q_0 \subset Q_u, P_1 \subset P_s$ and $Q_0 \pitchfork_I P_1$, any of the two elements $(\hat{P}, \hat{Q}, \hat{n})$ obtained by the corresponding allowed parabolic composition will be written as

$$(5.4) \quad (\hat{P}, \hat{Q}, \hat{n}) \in (P_0, Q_0, n_0) \square (P_1, Q_1, n_1).$$

5.4 Definition of the Transversality Relation

Let I be a parameter interval of some level, and let $(P_0, Q_0, n_0), (P_1, Q_1, n_1)$ be elements of $\mathcal{R}(I)$ which satisfy $Q_0 \subset Q_u, P_1 \subset P_s$.

From (R4) the condition (PC1) of Subsection 3.5 is satisfied provided ε_0 small enough. Denote by (x_0, y_0) (resp. (x_1, y_1)) the coordinates in the rectangle containing P_0 (resp. Q_1). A function $\overline{C}(y_0, x_1)$ was defined in Subsection 3.5, together with

$$(5.5) \quad \delta(Q_0, P_1) = \min_{y_0} \min_{x_1} -\overline{C}(y_0, x_1).$$

In Subsection 3.5, we were asking for δ to be much larger than $|P_0|$ and $|Q_1|$. From the formulas of Appendix A, we have

$$(5.6) \quad C^{-1}|P_1| \leq |\overline{C}_x| \leq C|P_1|,$$

$$(5.7) \quad C^{-1}|Q_0| \leq |\overline{C}_y| \leq C|Q_0|.$$

We will introduce

$$(5.8) \quad \delta_L(Q_0, P_1) := \max_{y_0} \min_{x_1} -\overline{C}(y_0, x_1),$$

$$(5.9) \quad \delta_R(Q_0, P_1) := \min_{y_0} \max_{x_1} -\overline{C}(y_0, x_1),$$

$$(5.10) \quad \delta_{LR}(Q_0, P_1) := \max_{y_0} \max_{x_1} -\overline{C}(y_0, x_1).$$

All together, $\delta, \delta_L, \delta_R, \delta_{LR}$ are the values of $-\overline{C}$ at the four corners of the rectangle of definition of \overline{C} . We have from (5.6), (5.7) that

$$(5.11) \quad C^{-1}|Q_0| \leq \delta_L(Q_0, P_1) - \delta(Q_0, P_1) \leq C|Q_0|,$$

$$(5.12) \quad C^{-1}|P_1| \leq \delta_R(Q_0, P_1) - \delta(Q_0, P_1) \leq C|P_1|,$$

$$(5.13) \quad C^{-1}|Q_0| \leq \delta_{LR}(Q_0, P_1) - \delta_R(Q_0, P_1) \leq C|Q_0|,$$

$$(5.14) \quad C^{-1}|P_1| \leq \delta_{LR}(Q_0, P_1) - \delta_L(Q_0, P_1) \leq C|P_1|.$$

Preliminary Definition. We write $Q_0 \overline{\cap}_I P_1$ if the following holds

(T1) for all $t \in I$,

$$\delta_{LR}(Q_0, P_1) \geq 2|I|,$$

(T2) for some $t_0 \in I$,

$$\delta_R(Q_0, P_1) \geq 2|Q_0|^{1-\eta},$$

(T3) for some $t_1 \in I$,

$$\delta_L(Q_0, P_1) \geq 2|P_1|^{1-\eta}.$$

Definition.

We say that Q_0, P_1 are I -transverse and write $Q_0 \pitchfork_I P_1$ if there exist a parameter interval $\tilde{I} \supset I$, elements $(\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0), (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) \in \mathcal{R}(\tilde{I})$ with $\tilde{P}_1 \supset P_1, \tilde{Q}_0 \supset Q_0$ such that $\tilde{Q}_0 \overline{\pitchfork}_{\tilde{I}} \tilde{P}_1$.

Remark.

1. Taking $\tilde{I} = I, \tilde{P}_0 = P_0, \tilde{Q}_1 = Q_1$, it is obvious that if $Q_0 \overline{\pitchfork}_I P_1$, then $Q_0 \pitchfork_I P_1$.
2. In view of our inductive procedure, all $(\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0), (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ which have to be considered have been constructed before $(P_0, Q_0, n_0), (P_1, Q_1, n_1)$.
3. As mentioned before, the definition of the transversality relation is quite involved. Some justification for the choice of quantifiers in (T1), (T2), (T3) can be found in Appendix C.

At first sight, it appears that properties (T2), (T3) above are not quite sufficient to guarantee condition (PC2) of parabolic composition (Subsection 3.5), because they involve only one value of the parameter. The next property takes care of this problem.

(R7) If $(P_0, Q_0, n_0), (P_1, Q_1, n_1) \in \mathcal{R}(I)$ satisfy $Q_0 \subset Q_u, P_1 \subset P_s$ and $Q_0 \pitchfork_I P_1$ holds, then, for all $t \in I$, we have

$$\delta(Q_0, P_1) \geq C^{-1} \left(|P_1|^{1-\eta} + |Q_0|^{1-\eta} \right).$$

5.5 The Class $\mathcal{R}(I_0)$

Recall that in Subsection 4.2 we had the normalization

$$(4.4) \quad \theta(0, 0, t) \equiv t, \quad |t| \leq t_0.$$

As θ is monotonous in both variables, it follows from the definition of C (formula (3.18) of Subsection 3.5) that for all $t \in (-t_0, t_0)$:

$$(5.15) \quad -\overline{C}(y_0, x_1, t) \leq t$$

Therefore, for the starting interval $I_0 = [\varepsilon_0, 2\varepsilon_0]$, condition (T1) above is never satisfied.

Parabolic composition is never allowed in $\mathcal{R}(I_0)$. Thus (see Subsection 5.3) the class $\mathcal{R}(I_0)$ consists only of the elements (P_a, Q_a, n) given by (R1) which are associated with the horseshoe K .

Condition (R2) is satisfied because of the choice of u_0, v_0, D_0 . Condition (R3) is obviously satisfied, as are conditions (R5) and (R6). Condition (R4) will be checked in Section 7.

5.6 Definition of the Regularity Property

We introduce first some terminology and some concepts related to the transversality relation.

5.6.1 Let I be a parameter interval of some level, and let $(P_0, Q_0, n_0), (P_1, Q_1, n_1)$ be elements of $\mathcal{R}(I)$ such that $Q_0 \subset Q_u$ and $P_1 \subset P_s$. When Q_0 and P_1 are not I -transverse, two cases may happen:

- if $G_t(Q_0) \cap P_1 = \emptyset$ for all $t \in I$, we say that Q_0 and P_1 are *I-separated*;
- otherwise, we say that Q_0 and P_1 are *I-critically related*.

5.6.2 Let $(P, Q, n) \in \mathcal{R}(I)$. An *I-decomposition* of P is a finite family $(P_\alpha, Q_\alpha, n_\alpha)$ of elements of $\mathcal{R}(I)$ such that the P'_α 's are disjoint, contained in P and satisfy

$$(5.16) \quad W^s(\Lambda, \widehat{R}) \cap P = \bigsqcup_{\alpha} (W^s(\Lambda, \widehat{R}) \cap P_{\alpha}),$$

where $W^s(\Lambda, \widehat{R})$ was defined in Subsection 2.2. We say that P is *I-decomposable* if it admits a non-trivial *I-decomposition*. Then, there is a coarsest non-trivial *I-decomposition*, namely by the children of P , which is called the *canonical I-decomposition*.

Remark. We will see in Section 8 that any P has only finitely many children.

5.6.3 Let $(P, Q, n) \in \mathcal{R}(I)$. We say that Q is *I-transverse* if either $Q \cap Q_u = \emptyset$ or $Q \subset Q_u$ and there exists an *I-decomposition* $(P_\alpha, Q_\alpha, n_\alpha)$ of P_s such that, for any α , Q and P_α are either *I-transverse* or *I-separated*.

We say that Q is *I-critical* when it is not *I-transverse*. This is always the case if $Q \supset Q_u$.

We also define in a symmetric way an *I-decomposition* for Q , and *I-transversality* or *I-criticality* for P .

5.6.4 We say that $(P, Q, n) \in \mathcal{R}(I)$ is *I-bicritical* if both P and Q are *I-critical*. The corresponding iterate should be seen as describing some recurrence to the “critical locus”.

Definition. Let $\beta > 1$. We say that the parameter interval I is *β -regular* (or just regular when the value of β is fixed) if any *I-bicritical* element $(P, Q, n) \in \mathcal{R}(I)$ satisfies, for all $t \in I$:

$$(5.17) \quad |P| < |I|^\beta, \quad |Q| < |I|^\beta.$$

Remark.

1. When $d_s^0 \geq d_u^0$, in Section 9 we will take for β a number satisfying

$$(5.18) \quad 1 < \beta < \beta_{max} := \frac{(1 - d_u^0)(d_s^0 + d_u^0)}{d_s^0(2d_s^0 + d_u^0 - 1)}.$$

Condition (H4) in Section 1.2 is actually equivalent to $\beta_{max} > 1$. When $d_u^0 \geq d_s^0$, we exchange d_s^0 and d_u^0 in the definition of β_{max} .

2. We will see in Section 6 that if \tilde{I} is a β -regular parameter interval, and $I \subset \tilde{I}$ is a candidate at the next level, then I is at least $\bar{\beta}$ -regular with

$$(5.19) \quad \bar{\beta} = \beta(1 + \tau)^{-1}.$$

This will allow us to apply to candidates, with only slightly worse constants, all results that have been proven for regular parameter intervals. But, obviously, we cannot let the regularity exponent deteriorate too much (it must stay > 1), which explains why candidates have to pass the β -regularity test.

6 Some Properties of the Classes $\mathcal{R}(I)$

We recall that all parameter intervals that we consider in the sequel are assumed to be regular or candidates (meaning in this case that their parent is regular).

6.1 Transversality is Hereditary

The following obvious but fundamental property was forced into the definition of the transversality relation.

Proposition 2. *Let $\tilde{I} \supset I$ be parameter intervals. Let $(P_0, Q_0, n_0), (P_1, Q_1, n_1) \in \mathcal{R}(I)$ and $(\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0), (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) \in \mathcal{R}(\tilde{I})$. Assume that $Q_0 \subset \tilde{Q}_0 \subset Q_u$ and $P_1 \subset \tilde{P}_1 \subset P_s$. If \tilde{Q}_0 and \tilde{P}_1 are \tilde{I} -transverse, then Q_0 and P_1 are I -transverse.*

Corollary 1. *Let $\tilde{I} \supset I$ be parameter intervals, and let $(P_0, Q_0, n_0), (P_1, Q_1, n_1) \in \mathcal{R}(\tilde{I}) \cap \mathcal{R}(I)$ be such that $Q_0 \subset Q_u, P_1 \subset P_s$. If their parabolic composition is allowed in $\mathcal{R}(\tilde{I})$, it is also allowed in $\mathcal{R}(I)$.*

Proof. This is the case $\tilde{Q}_0 = Q_0, \tilde{P}_1 = P_1$ of the proposition. □

Corollary 2. *Let $\tilde{I} \supset I$ be parameter intervals. Then $\mathcal{R}(\tilde{I}) \subset \mathcal{R}(I)$.*

Remark. *This is a slight abus de langage of no consequence: properly speaking, we mean that the restriction to I of any $(P, Q, n) \in \mathcal{R}(\tilde{I})$ belongs to $\mathcal{R}(I)$.*

Proof. As simple composition does not depend on the parameter interval, Corollary 2 is an immediate consequence by induction on length of property (R6). □

Corollary 3. *Let $\tilde{I} \supset I$ be parameter intervals, and let $(P, Q, n) \in \mathcal{R}(\tilde{I})$. If Q is \tilde{I} -transverse, then it is also I -transverse.*

Proof. If $Q \cap Q_u = \emptyset$, this is obvious. Assume therefore that $Q \subset Q_u$. Then there exists an \tilde{I} -decomposition $(P_\alpha, Q_\alpha, n_\alpha)_\alpha$ of P_s by elements of $\mathcal{R}(\tilde{I})$ such that for all α , Q and P_α are either \tilde{I} -transverse or \tilde{I} -separated.

First observe that $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{R}(I)$ and therefore this is also an I -decomposition of P_s . By Corollary 1, if Q and P_α are \tilde{I} -transverse, they are also I -transverse. On the other hand, it is obvious from the definition that if Q and P_α are \tilde{I} -separated they are also I -separated. The result follows. □

6.2 Criticality and Decomposability

Proposition 3. *Let I be a parameter interval, and let $(P, Q, n) \in \mathcal{R}(I)$. If Q is I -transverse, then P is I -decomposable.*

Proof. Let us first assume that $Q \cap Q_u = \emptyset$. Let $a \in \mathcal{a}$ be such that $Q \subset R_a$. We have

$$(6.1) \quad R_a \cap W^s(\Lambda, \widehat{R}) \subset \bigcup_{(a, a') \in \mathcal{B}} (P_{a, a'} \cap W^s(\Lambda, \widehat{R})) \cup L_u;$$

for each $a' \in \mathcal{a}$ such that $(a, a') \in \mathcal{B}$, we have the simple child of P :

$$(6.2) \quad (P(a'), Q(a'), n+1) = (P, Q, n) * (P_{aa'}, Q_{aa'}, 1),$$

and together they form by (6.1) an I -decomposition of P (the canonical one).

Let us now assume that $Q \subset Q_u$. As Q is I -transverse, there is an I -decomposition $(P_\alpha, Q_\alpha, n_\alpha)_\alpha$ of P_s such that, for each α , Q and P_α are not I -critically related. For each α such that Q and P_α are I -transverse, let $(P_\alpha^\pm, Q_\alpha^\pm, n_\alpha + n + N_0)$ be the two elements produced by the allowed parabolic composition. Together with the simple children defined by (6.2), they form an I -decomposition of P . \square

Corollary 4. *Let I be a β -regular parameter interval and let $(P, Q, n) \in \mathcal{R}(I)$. If P is I -critical and $|P| > |I|^\beta$ or $|Q| > |I|^\beta$ for some $t \in I$, then P is I -decomposable.*

Proof. Indeed, by the very definition of regularity, Q cannot be I -critical. \square

The decomposability of "fat" critical rectangles is crucial to our analysis. As mentioned before, Corollary 4 will apply to candidate intervals with $\bar{\beta} = \beta(1 + \tau)^{-1}$ instead of β .

6.3 Concavity

The following result is a partial converse to Proposition 2. We call *concavity* the corresponding property of $\mathcal{R}(I)$.

Proposition 4. *Let I be a parameter interval. Let $(P_0, Q_0, n_0), (P'_0, Q'_0, n'_0), (P_1, Q_1, n_1), (P'_1, Q'_1, n'_1)$ be elements of $\mathcal{R}(I)$ such that*

$$Q_0 \subset Q'_0 \subset Q_u, \quad P_1 \subset P'_1 \subset P_s.$$

If both $Q_0 \pitchfork_I P'_1$ and $Q'_0 \pitchfork_I P_1$ hold, then $Q'_0 \pitchfork_I P'_1$ also holds.

Proof. From the definition (3.18) of the function C and the monotonicity of θ with respect to each variable, it follows that we have, for each $t \in I$:

$$(6.3) \quad \delta(Q_0, P_1) \geq \max\left(\delta(Q'_0, P_1), \delta(Q_0, P'_1)\right),$$

$$(6.4) \quad \delta_L(Q'_0, P_1) \geq \delta_L(Q_0, P_1) \geq \delta_L(Q_0, P'_1),$$

$$(6.5) \quad \delta_R(Q_0, P'_1) \geq \delta_R(Q_0, P_1) \geq \delta_R(Q'_0, P_1),$$

$$(6.6) \quad \min\left(\delta_{LR}(Q'_0, P_1), \delta_{LR}(Q_0, P'_1)\right) \geq \delta_{LR}(Q_0, P_1).$$

By definition of the transversality relation, there exist parameter intervals \tilde{I}_1, \tilde{I}_2 containing I , elements $(\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0), (\tilde{P}'_1, \tilde{Q}'_1, \tilde{n}'_1)$ in $\mathcal{R}(\tilde{I}_1)$, elements $(\tilde{P}'_0, \tilde{Q}'_0, \tilde{n}'_0), (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ in $\mathcal{R}(\tilde{I}_2)$ such that $Q_0 \subset \tilde{Q}_0, Q'_1 \subset \tilde{Q}'_1, Q'_0 \subset \tilde{Q}'_0, P_1 \subset \tilde{P}_1$ and

$$(6.7) \quad \tilde{Q}_0 \bar{\cap}_{\tilde{I}_1} \tilde{P}'_1,$$

$$(6.8) \quad \tilde{Q}'_0 \bar{\cap}_{\tilde{I}_2} \tilde{P}_1.$$

If we have either $Q'_0 \subset \tilde{Q}_0$ or $P'_1 \subset \tilde{P}_1$, we can already conclude that $Q'_0 \cap_I P'_1$. Assume therefore that $\tilde{Q}_0 \subset Q'_0, \tilde{P}_1 \subset P'_1$. Assume also for instance that $\tilde{I}_1 \subset \tilde{I}_2$. We have $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) \in \mathcal{R}(\tilde{I}_2)$ and $\tilde{P}_1 \subset P'_1$. By the coherence property (Proposition 6 in Subsection 6.5), the element (P'_1, Q'_1, n'_1) belongs to $\mathcal{R}(\tilde{I}_2)$. We will show that

$$(6.9) \quad \tilde{Q}'_0 \bar{\cap}_{\tilde{I}_2} P'_1$$

which implies that Q'_0 and P'_1 are I -transverse. We check properties (T1)–(T3) of Subsection 5.5. First, by (6.6) and (6.8), we have, for all $t \in \tilde{I}_2$

$$(6.10) \quad \delta_{LR}(\tilde{Q}'_0, P'_1) \geq \delta_{LR}(\tilde{Q}'_0, \tilde{P}_1) \geq 2|\tilde{I}_2|.$$

Next, by (6.8) and (6.5), there exists $t_0 \in \tilde{I}_2$ such that

$$(6.11) \quad \delta_R(\tilde{Q}'_0, P'_1) \geq \delta_R(\tilde{Q}'_0, \tilde{P}_1) \geq 2|\tilde{Q}'_0|^{1-\eta}.$$

Finally, by (6.7), there exists $t_1 \in \tilde{I}_1 \subset \tilde{I}_2$ such that

$$(6.12) \quad \delta_L(\tilde{Q}_0, \tilde{P}'_1) \geq 2|\tilde{P}'_1|^{1-\eta}.$$

For this value t_1 , we then have, by (6.4):

$$(6.13) \quad \delta_L(\tilde{Q}'_0, P'_1) \geq \delta_L(\tilde{Q}_0, \tilde{P}'_1) \geq 2|\tilde{P}'_1|^{1-\eta} \geq 2|P'_1|^{1-\eta}.$$

We have proved (6.9) and this concludes the proof of the proposition. \square

Remark. *The concavity property is very helpful in the sequel. The proof of the proposition shows why the definition of the transversality relation had to be complicated.*

Corollary 5. *Let I be a parameter interval, (P_0, Q_0, n_0) , (P_1, Q_1, n_1) , (P'_1, Q'_1, n'_1) be elements of $\mathcal{R}(I)$ such that $Q_0 \subset Q_u$, $P_1 \subset P'_1 \subset P_s$. If P'_1 is I -transverse and $Q_0 \pitchfork_I P_1$ holds, then $Q_0 \pitchfork_I P'_1$ also holds.*

Proof. There exists an I -decomposition $(P_\alpha, Q_\alpha, n_\alpha)$ of Q_u such that, for any α , Q_α and P'_1 are either I -transverse or I -separated. There exists α such that Q_α and Q_0 do intersect. As $Q_0 \pitchfork_I P_1$ holds, Q_α and P'_1 must be I -transverse. If $Q_\alpha \supset Q_0$, it follows from Proposition 2 that Q_0 and P'_1 are I -transverse. If $Q_\alpha \subset Q_0$, the same conclusion follows from Proposition 4. \square

6.4 Children are Born From Their Parent

Let I be a parameter interval, $(\tilde{P}, \tilde{Q}, \tilde{n})$ be an element of $\mathcal{R}(I)$ and $a \in \mathcal{a}$ the letter such that $\tilde{Q} \subset \mathcal{R}_a$. The simple children of \tilde{P} are given by formula (6.2) and parametrized by the elements $a' \in \mathcal{a}$ such that $(a, a') \in \mathcal{B}$. All children are simple unless $\tilde{Q} \subset Q_u$.

Proposition 5. *Assume $\tilde{Q} \subset Q_u$ and let (P, Q, n) be a non-simple child of \tilde{P} . Then there exists $(P_1, Q_1, n_1) \in \mathcal{R}(I)$ such that $\tilde{Q} \pitchfork_I P_1$ and*

$$(P, Q, n) \in (\tilde{P}, \tilde{Q}, \tilde{n}) \square (P_1, Q_1, n_1).$$

Moreover, the parent \tilde{P}_1 of P_1 is I -critical.

Proof.

1. By (R6) in Subsection 5.3, (P, Q, n) is obtained from shorter elements by simple or parabolic composition. Let us first assume that we can write

$$(6.14) \quad (P, Q, n) = (P_0, Q_0, n_0) * (P_1, Q_1, n_1),$$

with $n_0, n_1 > 0$. As P is a non-simple child, we must have $n_1 > 1$. Let $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ be the element of $\mathcal{R}(I)$ such that \tilde{P}_1 is the parent of P_1 . If P_1 was a simple child, one would be able to write

$$(6.15) \quad (P_1, Q_1, n_1) = (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) * (P_{aa'}, Q_{aa'}, 1)$$

for some $(a, a') \in \mathcal{B}$, and then

$$(6.16) \quad (P, Q, n) = (\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0) * (P_{aa'}, Q_{aa'}, 1)$$

where

$$(6.17) \quad (\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0) = (P_0, Q_0, n_0) * (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1).$$

Thus P would be a simple child, in contradiction with the assumption of the proposition. Therefore, P_1 is a non-simple child; by induction on the length, we can write

$$(6.18) \quad (P_1, Q_1, n_1) \in (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) \square (P_2, Q_2, n_2)$$

for some $(P_2, Q_2, n_2) \in \mathcal{R}(I)$. We must have

$$(6.19) \quad \tilde{Q}_1 \pitchfork_I P_2.$$

We define

$$(6.20) \quad (\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0) := (P_0, Q_0, n_0) * (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1).$$

We then have $\tilde{Q}_0 \subset \tilde{Q}_1$ and hence, from (6.19) and Proposition 1

$$(6.21) \quad \tilde{Q}_0 \pitchfork_I P_2.$$

Thus, the parabolic composition of $(\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0)$ and (P_2, Q_2, n_2) is allowed; we obviously have:

$$(6.22) \quad (P, Q, n) \in (\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0) \square (P_2, Q_2, n_2).$$

2. We have shown so far that it is always possible to write

$$(6.23) \quad (P, Q, n) \in (P_0, Q_0, n_0) \square (P_1, Q_1, n_1)$$

for some allowed parabolic composition in $\mathcal{R}(I)$. We take n_0 maximal in (6.23), and assume by contradiction that $n_0 < \tilde{n}$. Let $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ be the element of $\mathcal{R}(I)$ such that \tilde{P}_1 is the parent of P_1 ; let (P'_0, Q'_0, n'_0) be the element of $\mathcal{R}(I)$ such that P'_0 is the child of P_0 containing P . As $n_0 < \tilde{n}$, we have $n'_0 < n$. As $g_t^{n_0}(P) \subset L_u$ by (6.23), P'_0 must be a non-simple child. Then, from the induction hypothesis, we can write

$$(6.24) \quad (P'_0, Q'_0, n'_0) \in (P_0, Q_0, n_0) \square (P'_1, Q'_1, n'_1)$$

for some $(P'_1, Q'_1, n'_1) \in \mathcal{R}(I)$ with $P'_1 \supset P_1$. We have thus $P'_1 \supset \tilde{P}_1$. As Q_0 is I -transverse to P'_1 , it is also I -transverse to \tilde{P}_1 (Proposition 2), and we can define $(\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0) \in \mathcal{R}(I)$ by

$$(6.25) \quad (\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0) \in (P_0, Q_0, n_0) \square (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1),$$

and $\tilde{P}_0 \supset P$. If P_1 was a simple child of \tilde{P}_1 , P would be a simple child of \tilde{P}_0 . Therefore, by the induction hypothesis, we can write

$$(6.26) \quad (P_1, Q_1, n_1) \in (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) \square (P_2, Q_2, n_2).$$

for some (P_2, Q_2, n_2) with $\tilde{Q}_1 \pitchfork_I P_2$. But then $\tilde{Q}_0 \pitchfork_I P_2$ (Proposition 2) and we have

$$(6.27) \quad (P, Q, n) \in (\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0) \square (P_2, Q_2, n_2)$$

which contradicts the maximality of n_0 . We have proven the first part of the proposition.

3. Assume by contradiction that the parent \tilde{P}_1 of P_1 is I -transverse. As $Q_0 \pitchfork_I P_1$, we have $\tilde{P}_1 \subset P_s$. There exists, therefore, an I -decomposition $(P_\alpha, Q_\alpha, n_\alpha)$ of Q_u such that Q_α and \tilde{P}_1 are never critically related. By definition of an I -decomposition, there exists α such that $g^{\tilde{n}}(P) \cap Q_\alpha \neq \emptyset$; then Q_α and \tilde{P}_1 are not I -separated and they must be I -transverse.

As $\tilde{Q} \cap Q_\alpha \neq \emptyset$, we have either $\tilde{Q} \subset Q_\alpha$, and $\tilde{Q} \pitchfork_I \tilde{P}_1$ by Proposition 2, or $\tilde{Q} \supset Q_\alpha$, and $\tilde{Q} \pitchfork_I \tilde{P}_1$ again by concavity (Proposition 4), because both $\tilde{Q}_\alpha \pitchfork_I \tilde{P}_1$ and $\tilde{Q} \pitchfork_I P_1$ hold. Thus, we can define $(\hat{P}, \hat{Q}, \hat{n})$ by

$$(6.28) \quad (\hat{P}, \hat{Q}, \hat{n}) \in (\tilde{P}, \tilde{Q}, \tilde{n}) \square (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1),$$

and $P \subset \hat{P}$. We then have $P \subset \hat{P} \subset \tilde{P}$, $\hat{P} \neq P$, $\hat{P} \neq \tilde{P}$; thus, P would not be a child of \tilde{P} . \square

6.5 Coherence and Parametric Concavity

The coherence property, asserted in the next proposition, means that larger rectangles are constructed before thinner ones.

Proposition 6. *Let $I \subset \tilde{I}$ be parameter intervals, and let $(P, Q, n) \in \mathcal{R}(I)$, $(\tilde{P}, \tilde{Q}, \tilde{n}) \in \mathcal{R}(\tilde{I})$. If $\tilde{P} \subset P$, then $(P, Q, n) \in \mathcal{R}(\tilde{I})$.*

The next property, called parametric concavity, is another partial converse to Proposition 2; it is formally very similar to Proposition 4.

Proposition 7. *Let $I \subset I'$ be parameter intervals and let (P_0, Q_0, n_0) , (P'_0, Q'_0, n'_0) , (P_1, Q_1, n_1) be elements of $\mathcal{R}(I')$ such that $Q_0 \subset Q'_0 \subset Q_u$ and $P_1 \subset P_s$. If both $Q_0 \pitchfork_{I'} P_1$ and $Q'_0 \pitchfork_I P_1$ hold, then $Q'_0 \pitchfork_{I'} P_1$ also holds.*

Obviously there is a similar statement exchanging P 's and Q 's. In the proof of Proposition 6, we will use the following result, which is of independent interest.

Proposition 8. *Let I be a parameter interval, and let (P, Q, n) , (P', Q', n') , (P'', Q'', n'') be elements $\mathcal{R}(I)$ with $P \subset P' \subset P''$, $P' \neq P''$. We have:*

a) *If $(P, Q, n) = (P'', Q'', n'') * (P_1, Q_1, n_1)$ for some $(P_1, Q_1, n_1) \in \mathcal{R}(I)$, then*

$$(P', Q', n') = (P'', Q'', n'') * (P'_1, Q'_1, n'_1)$$

for some $(P'_1, Q'_1, n'_1) \in \mathcal{R}(I)$.

b) If $(P, Q, n) \in (P'', Q'', n'') \square (P_1, Q_1, n_1)$ for some $(P_1, Q_1, n_1) \in \mathcal{R}(I)$, then

$$(P', Q', n') \in (P'', Q'', n'') \square (P'_1, Q'_1, n'_1)$$

for some $(P'_1, Q'_1, n'_1) \in \mathcal{R}(I)$.

Proof of Proposition 8. We only consider case b), case a) being similar but easier. It is sufficient to consider the case where P' is the parent of P . Let $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ be the element of $\mathcal{R}(I)$ such that \tilde{P}_1 is the parent of P_1 . We claim that Q'' and \tilde{P}_1 are I -transverse.

Indeed, let $(\bar{P}, \bar{Q}, \bar{n}) \in \mathcal{R}(I)$ such that \bar{P} is the child of P'' containing P' . As $g^{n''}(P) \subset L_u$, \bar{P} is a non-simple child; by Proposition 5, there exists $(\bar{P}_1, \bar{Q}_1, \bar{n}_1)$ such that

$$(6.29) \quad (\bar{P}, \bar{Q}, \bar{n}) \in (P'', Q'', n'') \square (\bar{P}_1, \bar{Q}_1, \bar{n}_1).$$

We then have $Q'' \cap_I \bar{P}_1$ and $\bar{P}_1 \supset \tilde{P}_1$; the claim then follows from Proposition 2.

We can therefore define $(\tilde{P}, \tilde{Q}, \tilde{n}) \in \mathcal{R}(I)$ by

$$(6.30) \quad (\tilde{P}, \tilde{Q}, \tilde{n}) \in (P'', Q'', n'') \square (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$$

and $P \subset \tilde{P}$. Let us show that $\tilde{P} = P'$. Otherwise, we have $P' \subset \tilde{P}$, and $P' \neq \tilde{P}$. Let $(\hat{P}, \hat{Q}, \hat{n})$ be the element of $\mathcal{R}(I)$ such that \hat{P} is the child of \tilde{P} containing P' .

If P_1 was a simple child of \tilde{P}_1 , we would have $\tilde{n}_1 = n_1 - 1$ and $\tilde{n} = n - 1$, forcing $\tilde{P} = P'$. Hence, P_1 is a non-simple child; this implies that $g^{\tilde{n}}(P) \subset L_u$ and that \hat{P} is also a non-simple child. By Proposition 5, there exist (P_2, Q_2, n_2) , $(\hat{P}_2, \hat{Q}_2, \hat{n}_2)$ in $\mathcal{R}(I)$ such that $P_2 \subset \hat{P}_2$ and

$$(6.31) \quad (P_1, Q_1, n_1) \in (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) \square (P_2, Q_2, n_2),$$

$$(6.32) \quad (\hat{P}, \hat{Q}, \hat{n}) \in (\tilde{P}, \tilde{Q}, \tilde{n}) \square (\hat{P}_2, \hat{Q}_2, \hat{n}_2).$$

Then both $\tilde{Q}_1 \cap_I P_2$ and $\tilde{Q} \cap_I \hat{P}_2$ hold. By concavity (Proposition 4), we also have $\tilde{Q}_1 \cap_I \hat{P}_2$. We then define $(\hat{P}_1, \hat{Q}_1, \hat{n}_1)$ by:

$$(6.33) \quad (\hat{P}_1, \hat{Q}_1, \hat{n}_1) \in (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) \square (\hat{P}_2, \hat{Q}_2, \hat{n}_2)$$

and $P_1 \subset \hat{P}_1$. Then $P_1 \subset \hat{P}_1 \subset \tilde{P}_1$ and \hat{P}_1 is distinct from P_1 and \tilde{P}_1 , which contradicts that \tilde{P}_1 is the parent of P_1 . \square

Proof of Proposition 6. It is sufficient to consider the case where P is the parent of \tilde{P} in $\mathcal{R}(I)$. Let (P_0, Q_0, n_0) be the element of $\mathcal{R}(\tilde{I})$ such that P_0 is the parent of \tilde{P} in $\mathcal{R}(\tilde{I})$. We want to show that $P_0 = P$. This is clear if \tilde{P} is a simple child of P_0 . We assume therefore that \tilde{P} is a non-simple child of P_0 . We know that $P \subset P_0$. By Proposition 5, there exists $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) \in \mathcal{R}(\tilde{I})$ such that

$$(6.34) \quad (\tilde{P}, \tilde{Q}, \tilde{n}) \in (P_0, Q_0, n_0) \square (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$$

with $Q_0 \pitchfork_{\tilde{I}} \tilde{P}_1$. By Proposition 8, there exists $(P_1, Q_1, n_1) \in \mathcal{R}(I)$ such that

$$(6.35) \quad (P, Q, n) \in (P_0, Q_0, n_0) \square (P_1, Q_1, n_1)$$

with $Q_0 \pitchfork_I P_1$. By induction on the length, as $\tilde{P}_1 \subset P_1$ we must have $(P_1, Q_1, n_1) \in \mathcal{R}(\tilde{I})$. By parameter concavity (Proposition 7), as both $Q_0 \pitchfork_{\tilde{I}} \tilde{P}_1$ and $Q_0 \pitchfork_I P_1$ hold, $Q_0 \pitchfork_{\tilde{I}} P_1$ must also hold, which implies $(P, Q, n) \in \mathcal{R}(\tilde{I})$. \square

Proof of Proposition 7. By definition of the transversality relation, there exist parameter intervals $\tilde{I} \supset I$, $\tilde{I}' \supset I'$ and elements $(\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0), (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) \in \mathcal{R}(\tilde{I}')$, $(\tilde{P}'_0, \tilde{Q}'_0, \tilde{n}'_0), (\tilde{P}'_1, \tilde{Q}'_1, \tilde{n}'_1) \in \mathcal{R}(\tilde{I})$ such that $Q_0 \subset \tilde{Q}_0$, $Q'_0 \subset \tilde{Q}'_0$, $P_1 \subset \tilde{P}_1$, $P_1 \subset \tilde{P}'_1$ and

$$(6.36) \quad \tilde{Q}_0 \pitchfork_{\tilde{I}'} \tilde{P}_1,$$

$$(6.37) \quad \tilde{Q}'_0 \pitchfork_{\tilde{I}} \tilde{P}'_1$$

both hold. If either $Q'_0 \subset \tilde{Q}_0$ or $I' \subset \tilde{I}$, we conclude immediately that $Q'_0 \pitchfork_{I'} P_1$ holds. Assume therefore that $\tilde{Q}_0 \subset Q'_0$ and $\tilde{I} \subset I'$. Let P_1^* be the largest of $\tilde{P}_1, \tilde{P}'_1$.

If $\tilde{P}_1 \subset \tilde{P}'_1$, $(\tilde{P}'_1, \tilde{Q}'_1, \tilde{n}'_1) \in \mathcal{R}(\tilde{I}')$ by coherence (Proposition 6), hence (P_1^*, Q_1^*, n_1^*) always belong to $\mathcal{R}(\tilde{I}')$. We will show that

$$(6.38) \quad \tilde{Q}'_0 \pitchfork_{\tilde{I}'} P_1^*$$

holds, which implies $Q'_0 \pitchfork_{I'} P_1$.

We check properties (T1)–(T3) of Subsection 5.4. For all $t \in \tilde{I}'$, we have by (6.36) and (6.6)

$$(6.39) \quad \delta_{LR}(\tilde{Q}'_0, P_1^*) \geq \delta_{LR}(\tilde{Q}_0, \tilde{P}_1) \geq 2|\tilde{I}'|.$$

By (6.37), there exists $t_0 \in \tilde{I} \subset I' \subset \tilde{I}'$ such that

$$(6.40) \quad \delta_R(\tilde{Q}'_0, \tilde{P}'_1) \geq 2|\tilde{Q}'_0|^{1-\eta}.$$

Then, by (6.5), for the same t_0 , we have

$$(6.41) \quad \delta_R(\tilde{Q}'_0, P_1^*) \geq \delta_R(\tilde{Q}'_0, \tilde{P}'_1) \geq 2|\tilde{Q}'_0|^{1-\eta}.$$

When $P_1^* = \tilde{P}'_1$, it follows directly from (6.37) that we have

$$(6.42) \quad \delta_L(\tilde{Q}'_0, P_1^*) \geq 2|P_1^*|^{1-\eta}$$

for some $t_1 \in \tilde{I} \subset \tilde{I}'$.

When $P_1^* = \tilde{P}_1$, we use (6.36) and (6.4) to find $t_1 \in \tilde{I}'$ such that

$$(6.43) \quad \delta_L(\tilde{Q}'_0, P_1^*) \geq \delta_L(\tilde{Q}_0, \tilde{P}_1) \geq 2|P_1^*|^{1-\eta}.$$

We have thus proved (6.38). □

Remark.

1. The coherence property means in particular that the parent-child relation does not depend on the parameter interval (once both parent and child are defined).
2. We have presented together Propositions 6, 7 and 8 because the proofs are interconnected. But, actually, coherence (Proposition 6) was already used in the proof of Proposition 4. Logically speaking, all properties in Sections 5–7 should be proved together (as was already mentioned in Section 5).

6.6 Further Criteria for Transversality

In this subsection, we give other sufficient conditions for transversality that can be seen as partial converses to Proposition 2.

Proposition 9. *Let I be a parameter interval and let (P, Q, n) , (P_0, Q_0, n_0) , (P_1, Q_1, n_1) be elements of $\mathcal{R}(I)$ such that $Q \subset Q_u$ and $P_0 \subset P_1 \subset P_s$. Assume that $Q \pitchfork_I P_0$ holds and that $2|P_1|^{1-\eta} \leq |I|$ for some $t_1 \in I$. Then Q and P_1 are also I -transverse.*

Proof. By definition of the transversality relation, there exist $\tilde{I} \supset I$, $(\tilde{P}, \tilde{Q}, \tilde{n})$, $(\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0) \in \mathcal{R}(\tilde{I})$ such that $Q \subset \tilde{Q}$, $P_0 \subset \tilde{P}_0$ and $\tilde{Q} \pitchfork_{\tilde{I}} \tilde{P}_0$.

If $P_1 \subset \tilde{P}_0$ this already implies that $Q \pitchfork_I P_1$. Let us assume that $\tilde{P}_0 \subset P_1$. We will show that $\tilde{Q} \pitchfork_{\tilde{I}} P_1$ holds. By coherence (Proposition 4), we have $(P_1, Q_1, n_1) \in \mathcal{R}(\tilde{I})$. Let us check (T1)–(T3).

By (T1) for \tilde{Q} , \tilde{P}_0 and (6.6), we have, for all $t \in \tilde{I}$:

$$(6.44) \quad \delta_{LR}(\tilde{Q}, P_1) \geq \delta_{LR}(\tilde{Q}, \tilde{P}_0) \geq 2|\tilde{I}|.$$

By (T2) for \tilde{Q} , \tilde{P}_0 and (6.5), there exists $t_0 \in \tilde{I}$ such that

$$(6.45) \quad \delta_R(\tilde{Q}, P_1) \geq \delta_R(\tilde{Q}, \tilde{P}_0) \geq 2|\tilde{Q}|^{1-\eta}.$$

Finally, we have, for all $t \in I$, by (5.14)

$$(6.46) \quad \begin{aligned} \delta_L(\tilde{Q}, P_1) &\geq \delta_{LR}(\tilde{Q}, P_1) - C|P_1| \\ &\geq 2|I| - C|P_1|. \end{aligned}$$

But, for $t = t_1$, we have, if ε_0 is small enough

$$(6.47) \quad 2|I| - C|P_1| \geq 4|P_1|^{1-\eta} - C|P_1| \geq 2|P_1|^{1-\eta},$$

□

Proposition 10. *Let I be a parameter interval and let (P, Q, n) , (P_0, Q_0, n_0) , (P_1, Q_1, n_1) be elements of $\mathcal{R}(I)$ such that $Q \subset Q_u$, $P_0 \subset P_1 \subset P_s$. Assume that $Q \pitchfork_I P_0$ holds and that $|P_1| \leq \frac{1}{2} |Q|$ for all $t \in I$. Then Q and P_1 are also I -transverse.*

Proof. The argument is the same as in Proposition 9, with a slight difference to check (T3) for \tilde{Q} , P_1 . We use (5.13), (5.14) to obtain, for the value t_0 of the parameter given by (T2) for \tilde{Q} , \tilde{P}_0 :

$$(6.48) \quad \begin{aligned} \delta_L(\tilde{Q}, P_1) &\geq \delta_R(\tilde{Q}, P_1) - C|P_1|, \\ &\geq 2|Q|^{1-\eta} - C|P_1|, \\ &\geq 2(2|P_1|)^{1-\eta} - C|P_1|, \\ &\geq 2|P_1|^{1-\eta}, \end{aligned}$$

if ε_0 is small enough. □

Proposition 11. *Let $I \subset I'$ be parameter intervals and let (P_0, Q_0, n_0) , (P_1, Q_1, n_1) be elements of $\mathcal{R}(I')$ such that $Q_0 \subset Q_u$, $P_1 \subset P_s$. Assume that $Q_0 \pitchfork_{I'} P_1$ holds and that we have $2|I'| < |P_1|^{1-\eta}$, for all $t \in I$. Then Q_0 and P_1 are also I' -transverse.*

Proof. By definition of the transversality relation, there exists $\tilde{I} \supset I$, $(\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0)$, $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) \in \mathcal{R}(\tilde{I})$ such that $\tilde{Q}_0 \supset Q_0$, $\tilde{P}_1 \supset P_1$ and $\tilde{Q}_0 \pitchfork_{\tilde{I}} \tilde{P}_1$ holds. If $\tilde{I} \supset I'$, this already implies that $Q_0 \pitchfork_{I'} P_1$ holds. Assume thus that $\tilde{I} \subset I'$. We show that $\tilde{Q}_0 \pitchfork_{I'} \tilde{P}_1$.

Conditions (T2), (T3) in I' follow from the same conditions in \tilde{I} . For the value t_1 given by (T3), we have

$$(6.49) \quad \delta_{LR}(\tilde{Q}_0, \tilde{P}_1) \geq \delta_L(\tilde{Q}_0, \tilde{P}_1) \geq 2|P_1|^{1-\eta} \geq 4|I'|.$$

By Corollary 11 in Subsection 7.6, this implies

$$(6.50) \quad \delta_{LR}(\tilde{Q}_0, \tilde{P}_1) \geq 2|I'|, \quad \forall t \in I'.$$

which is (T1). □

6.7 A Structure Theorem for New Rectangles

6.7.1 Associativity of Parabolic Composition. Let I be a parameter interval, and let (P_0, Q_0, n_0) , (P_1, Q_1, n_1) , (P_2, Q_2, n_2) be elements in $\mathcal{R}(I)$ such that $Q_0 \subset Q_u$, $Q_1 \subset Q_u$, $P_1 \subset P_s$, $P_2 \subset P_s$. We assume that both $Q_0 \pitchfork_I P_1$ and $Q_1 \pitchfork_I P_2$ hold.

Parabolic composition of (P_0, Q_0, n_0) , (P_1, Q_1, n_1) produces two elements $(P_{01}^+, Q_{01}^+, n_{01}^+)$, $(P_{01}^-, Q_{01}^-, n_{01}^-)$.

As Q_{01}^+ and Q_{01}^- are contained in Q_1 , it follows from Prop. 2 that both $Q_{01}^+ \pitchfork_I P_2$ and $Q_{01}^- \pitchfork_I P_2$ hold.

In the same way, parabolic composition of (P_1, Q_1, n_1) , (P_2, Q_2, n_2) produces two elements $(P_{12}^+, Q_{12}^+, n_{12}^+)$, $(P_{12}^-, Q_{12}^-, n_{12}^-)$ such that both $Q_0 \pitchfork_I P_{12}^+$ and $Q_0 \pitchfork_I P_{12}^-$ hold.

It is clear that the four elements of $\mathcal{R}(I)$ obtained by parabolic composition of $(P_{01}^+, Q_{01}^+, n_{01}^+)$ or $(P_{01}^-, Q_{01}^-, n_{01}^-)$ with (P_2, Q_2, n_2) are the same as the four elements obtained by the parabolic composition of (P_0, Q_0, n_0) with $(P_{12}^+, Q_{12}^+, n_{12}^+)$ or $(P_{12}^-, Q_{12}^-, n_{12}^-)$. Their domains are the components of $P_0 \cap (G_t \circ g_t^{n_0})^{-1} P_1 \cap (G_t \circ g_t^{n_1} \circ G_t \circ g_t^{n_0})^{-1} P_2$. If (P, Q, n) is any of these four elements, we will write

$$(6.51) \quad (P, Q, n) \in (P_0, Q_0, n_0) \square (P_1, Q_1, n_1) \square (P_2, Q_2, n_2).$$

The same considerations extend immediately, by induction on k , to the case of elements (P_0, Q_0, n_0) , (P_k, Q_k, n_k) such that $P_i \subset P_s$ for $0 < i < k$, $Q_i \subset Q_u$ for $0 \leq i < k$, and $Q_i \pitchfork_I P_{i+1}$ holds for $0 \leq i < k$. Then the successive parabolic compositions of $(P_0, Q_0, n_0), \dots, (P_k, Q_k, n_k)$ produce 2^k elements and we will write for any such element (P, Q, n) :

$$(6.52) \quad (P, Q, n) \in (P_0, Q_0, n_0) \square \dots \square (P_k, Q_k, n_k).$$

6.7.2 Statement of the Structure Theorem. We have seen in Subsection 5.5 that parabolic composition is never allowed in the class $\mathcal{R}(I_0)$ associated to the starting interval $I_0 = [\varepsilon_0, 2\varepsilon_0]$. This class consists exactly of the affine-like iterates associated to the Markov partition of the initial horseshoe K_{g_t} .

On the other hand, for elements (P, Q, n) belonging to some class $\mathcal{R}(I)$ but which are not (restrictions of) an element of $\mathcal{R}(I_0)$, parabolic composition must occur. The following theorem gives some rather precise information on this process.

Theorem 1. *Let I , be a parameter interval of level > 0 , \tilde{I} be the parent interval, and let (P, Q, n) be an element of $\mathcal{R}(I)$ which is not (the restriction of) an element of $\mathcal{R}(\tilde{I})$. Then there exists $k > 0$, elements $(P_0, Q_0, n_0), \dots, (P_k, Q_k, n_k)$ of $\mathcal{R}(\tilde{I})$ such that $Q_i \subset Q_u$ for $0 \leq i < k$, $P_i \subset P_s$ for $0 < i \leq k$, $Q_i \pitchfork_I P_{i+1}$ holds for $0 \leq i < k$, $Q_i \pitchfork_{\tilde{I}} P_{i+1}$ does not hold for $0 \leq i < k$ and*

$$(P, Q, n) \in (P_0, Q_0, n_0) \square \dots \square (P_k, Q_k, n_k).$$

Moreover, these elements are uniquely determined by these conditions, P_i is \tilde{I} -critical for $0 < i \leq k$ and Q_i is \tilde{I} -critical for $0 \leq i < k$.

6.7.3 We will first introduce a concept, relative to an element (P, Q, n) as in the theorem above, that leads to the determination of the (P_i, Q_i, n_i) .

Let m, p be integers such that $0 \leq m \leq p \leq n$. We say that $[m, p]$ is an \tilde{I} -interval if there exists $(\tilde{P}, \tilde{Q}, \tilde{n}) \in \mathcal{R}(\tilde{I})$ such that

$$g_t^m(P) \subset \tilde{P} \text{ for all } t \in I \text{ and } \tilde{n} = p - m.$$

Lemma 1. *The union of two \tilde{I} -intervals with non empty intersection is an \tilde{I} -interval.*

Proof. Let $[m, p], [m', p']$ be these \tilde{I} -intervals, and let $(\tilde{P}, \tilde{Q}, \tilde{n}), (\tilde{P}', \tilde{Q}', \tilde{n}')$ be the corresponding elements of $\mathcal{R}(\tilde{I})$. Without loss of generality, we may assume that $m < m' \leq p < p'$. Replacing if necessary \tilde{P} by a larger rectangle, we also assume that the element $(\hat{P}, \hat{Q}, \hat{n})$ of $\mathcal{R}(\tilde{I})$ such that \hat{P} is the parent of \tilde{P} satisfies $m + \hat{n} < m'$. There are now two cases:

a) $p = m'$.

Let R_a be the rectangle containing \tilde{Q} . Then $R_a \supset \tilde{Q} = g^{\tilde{n}}(\tilde{P}) \supset g^p(P) = g^{m'}(P)$; thus \tilde{P}' is also contained in R_a and the simple composition

$$(6.53) \quad (\tilde{P}'', \tilde{Q}'', \tilde{n}'') := (\tilde{P}, \tilde{Q}, \tilde{n}) * (\tilde{P}', \tilde{Q}', \tilde{n}')$$

is defined. We have $m + \tilde{n}'' = p'$ and $g_t^m(P) \subset P''$.

b) $p > m'$.

Then, \tilde{P} is not a simple child of \hat{P} , because otherwise we would have $\hat{n} = \tilde{n} - 1 \geq m' - m$. By Proposition 5, there exists $(\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0)$ in $\mathcal{R}(\tilde{I})$ such that

$$(6.54) \quad (\tilde{P}, \tilde{Q}, \tilde{n}) \in (\hat{P}, \hat{Q}, \hat{n}) \square (\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0).$$

The element $(\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0)$ of $\mathcal{R}(\tilde{I})$ is associated to the \tilde{I} -interval $[\hat{m}, p]$, where $\hat{m} = m + \hat{n} + N_0$. We have $m + \hat{n} < m'$ and $g_t^{m+\hat{n}}(P) \subset L_u$, hence also $m + \hat{n} + N_0 = \hat{m} \leq m'$.

To conclude the proof, we argue by induction on the total length $p' - m$ of the interval considered. The case $p' - m = 0$ is trivial. In the other case, we have the \tilde{I} -intervals $[\hat{m}, p]$ and $[m', p']$ with $m < \hat{m} \leq m'$ and hence by induction $[\hat{m}, p']$ is an \tilde{I} -interval. Let $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ be the corresponding element of $\mathcal{R}(\tilde{I})$; we have $g_t^{\hat{m}}(P) \subset \tilde{P}_1 \subset \tilde{P}_0$. From (6.54), $\tilde{Q} \cap_{\tilde{I}} \tilde{P}_0$ holds, hence also does $\tilde{Q} \cap_{\tilde{I}} \tilde{P}_1$ by Proposition 2. Then, the parabolic composition of $(\hat{P}, \hat{Q}, \hat{n})$ and $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ is allowed and defines an element of $\mathcal{R}(\tilde{I})$ which guarantees that $[m, p']$ is an \tilde{I} -interval. \square

6.7.4 We will now show that the (P_i, Q_i, n_i) in the theorem are uniquely determined by their properties. Indeed, define $m_0 = 0, p_0 = n_0$ and for $i > 0$:

$$(6.55) \quad m_i = p_{i-1} + N_0, \quad p_i = m_i + n_i.$$

Lemma 2. *The maximal \tilde{I} -intervals are exactly the $[m_i, p_i]$, $0 \leq i \leq k$, with associated elements (P_i, Q_i, n_i) .*

Proof. First, the $[m_i, p_i]$ are indeed \tilde{I} -intervals with associated elements (P_i, Q_i, n_i) . To complete the proof, it is sufficient to show that no \tilde{I} -interval $[m, p]$ can intersect a gap (p_i, m_{i+1}) . Assume by contradiction that there exists such a $[m, p]$ with associated element $(\tilde{P}, \tilde{Q}, \tilde{n})$ and minimal $\tilde{n} = p - m$. As $g^\ell(P) \subset g^{\ell-p_i}(L_u)$ does not intersect R for $p_i < \ell < m_{i+1}$, we must have $m \leq p_i$ and $m_{i+1} \leq p$. By property (R6) of $\mathcal{R}(\tilde{I})$ (Subsection 5.3) and the minimality of \tilde{n} , there exists $(\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0), (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ in $\mathcal{R}(\tilde{I})$ such that

$$(6.56) \quad (\tilde{P}, \tilde{Q}, \tilde{n}) \in (\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0) \square (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$$

with $\tilde{n}_0 = p_i - m \leq n_i$, $\tilde{n}_1 = p - m_{i+1} \leq n_{i+1}$. But then, from $\tilde{Q}_0 \supset Q_i$, $\tilde{P}_1 \supset P_{i+1}$ and $\tilde{Q}_0 \cap_{\tilde{I}} \tilde{P}_1$, we deduce from Proposition 2 that $Q_i \cap_{\tilde{I}} P_{i+1}$ holds, a contradiction. \square

6.7.5 Lemma 2 allow us to *define* k as being the number of maximal \tilde{I} -intervals minus one, and to define the $(P_i, Q_i, n_i) \in \mathcal{R}(\tilde{I})$ as the elements of $\mathcal{R}(\tilde{I})$ associated to the successive maximal \tilde{I} -intervals. Observe that the maximal \tilde{I} -intervals $[m_i, p_i]$, ($0 \leq i \leq k$) must indeed satisfy $m_0 = 0$, $m_{i+1} = p_i + N_0$ for $0 \leq i < k$: every $\ell \in [0, n]$ not contained in an \tilde{I} -interval is such that $g_i^{\ell-N}(P) \subset L_u$ for some $0 < N < N_0$ and then no \tilde{I} -interval intersects with $(\ell - N, \ell - N + N_0)$, while $\{[\ell - N, \ell - N + N_0]\}$ are \tilde{I} -intervals. We observe also that $Q_i \cap_{\tilde{I}} P_{i+1}$ does not hold because otherwise $[m_i, p_{i+1}]$ would be an \tilde{I} -interval.

6.7.6 Let $0 \leq i < k$. Let us assume by induction over i that P_j is \tilde{I} -critical for $0 < j \leq i$, Q_j is \tilde{I} -critical for $0 \leq j < i$, $Q_j \cap_I P_{j+1}$ holds for $0 \leq j < i$ and that we have an element of $\mathcal{R}(I)$:

$$(6.57) \quad (P^{(i)}, Q^{(i)}, p_i) \in (P_0, Q_0, n_0) \square \cdots \square (P_i, Q_i, n_i)$$

such that $P \subset P^{(i)}$. The assumption is vacuously true for $i = 0$. We will prove it at step $i + 1$. For $i = k$, it gives the properties stated in the theorem for the (P_i, Q_i, n_i) .

6.7.7 We first prove that Q_i is \tilde{I} -critical. Assume by contradiction that Q_i is \tilde{I} -transverse. Then P_i is \tilde{I} -decomposable. Let $(\hat{P}_i, \hat{Q}_i, \hat{n}_i)$ be an element of $\mathcal{R}(\tilde{I})$ such that \hat{P}_i is a child of P_i intersecting $g^{m_i}(P \cap \Lambda)$. On the other hand, let $(\hat{P}^{(i)}, \hat{Q}^{(i)}, \hat{n}^{(i)})$ be the element of $\mathcal{R}(I)$ such that $\hat{P}^{(i)}$ is the child of $P^{(i)}$ containing P . We apply Proposition 5 twice. We find $(\tilde{P}_{i+1}, \tilde{Q}_{i+1}, \tilde{n}_{i+1})$ in $\mathcal{R}(I)$, $(\tilde{P}'_{i+1}, \tilde{Q}'_{i+1}, \tilde{n}'_{i+1})$ in $\mathcal{R}(\tilde{I})$ such that

$$(6.58) \quad (\hat{P}_i, \hat{Q}_i, \hat{n}_i) \in (P_i, Q_i, n_i) \square (\tilde{P}'_{i+1}, \tilde{Q}'_{i+1}, \tilde{n}'_{i+1}),$$

$$(6.59) \quad (\hat{P}^{(i)}, \hat{Q}^{(i)}, \hat{n}^{(i)}) \in (P^{(i)}, Q^{(i)}, p_i) \square (\tilde{P}_{i+1}, \tilde{Q}_{i+1}, \tilde{n}_{i+1}).$$

If we had $\tilde{n}'_{i+1} > \tilde{n}_{i+1}$, from $Q_i \cap_{\tilde{I}} \tilde{P}'_{i+1}$ and $Q^{(i)} \cap_I \tilde{P}_{i+1}$, we would deduce first by Proposition 2 that $Q_i \cap_I \tilde{P}'_{i+1}$, then by concavity (Proposition 4) that $Q_i \cap_I \tilde{P}_{i+1}$; parabolic composition would yield an

element $(\bar{P}_i, \bar{Q}_i, \bar{n}_i) \in \mathcal{R}(I)$ with $P_i \supsetneq \bar{P}_i \supsetneq \hat{P}_i$, in contradiction with coherence (Proposition 6) and the definition of \hat{P}_i .

Therefore, we must have $\tilde{n}'_{i+1} \leq \tilde{n}_{i+1}$; as $\hat{n}^{(i)} \leq n$, we have also $m_i + \hat{n}_i \leq n$, and $[m_i, m_i + \hat{n}_i]$ is an \hat{I} -interval strictly larger than $[m_i, p_i]$; this contradiction shows that Q_i is \tilde{I} -critical.

6.7.8 The proof that P_{i+1} is \tilde{I} -critical is rather similar. We assume by contradiction that it is \tilde{I} -transverse. Then Q_{i+1} is \tilde{I} -decomposable and we can find $(P_{i+1}^*, Q_{i+1}^*, n_{i+1}^*) \in \mathcal{R}(\tilde{I})$ such that Q_{i+1}^* is a child of Q_{i+1} intersecting $g^{p_{i+1}}(P \cap \Lambda)$. By Proposition 5, there exists $(P_*^{(i)}, Q_*^{(i)}, n_*^{(i)}) \in \mathcal{R}(\tilde{I})$ such that

$$(6.60) \quad (P_{i+1}^*, Q_{i+1}^*, n_{i+1}^*) \in (P_*^{(i)}, Q_*^{(i)}, n_*^{(i)}) \square (P_{i+1}, Q_{i+1}, n_{i+1}).$$

If we had $n_*^{(i)} > p_i$, we would derive a contradiction as follows: we should have $Q^{(i)} \pitchfork_I P_{i+1}$ either (if $\tilde{n}_{i+1} \leq n_{i+1}$) from $Q^{(i)} \pitchfork_I \tilde{P}_{i+1}$ by Proposition 2 or (if $\tilde{n}_{i+1} \geq n_{i+1}$) from $Q^{(i)} \pitchfork_I \tilde{P}_{i+1}$ and $Q_*^{(i)} \pitchfork_I P_{i+1}$ by Proposition 4; then, parabolic composition of $(P^{(i)}, Q^{(i)}, p_i)$ and $(P_{i+1}, Q_{i+1}, n_{i+1})$ produces an element $(P^{(i+1)}, Q^{(i+1)}, p_{i+1})$ in $\mathcal{R}(I)$ with $Q_{i+1} \supsetneq Q^{(i+1)} \supsetneq Q_{i+1}^*$, which is not compatible with coherence (Proposition 6).

Thus, we must have $n_*^{(i)} \leq p_i$; then we also have $n_{i+1}^* \leq p_{i+1}$ and $[p_{i+1} - n_{i+1}^*, p_{i+1}]$ is an \tilde{I} -interval strictly larger than $[m_{i+1}, p_{i+1}]$. This contradiction shows that P_{i+1} is \tilde{I} -critical.

6.7.9 We now prove that $Q^{(i)}$ and P_{i+1} are I -transverse. If $\tilde{n}_{i+1} \leq n_{i+1}$, we have $Q^{(i)} \pitchfork_I \tilde{P}_{i+1}$ by (6.59) and thus also $Q^{(i)} \pitchfork_I P_{i+1}$ by Proposition 2. Let us assume that $\tilde{n}_{i+1} > n_{i+1}$. In this case, we claim that Q_{i+1} is \tilde{I} -critical. Indeed, if it was \tilde{I} -transverse, P_{i+1} would be \tilde{I} -decomposable and we would find an element $(\hat{P}_{i+1}, \hat{Q}_{i+1}, \hat{n}_{i+1})$ of $\mathcal{R}(\tilde{I})$ such that \hat{P}_{i+1} is a child of P_{i+1} intersecting $g^{m_{i+1}}(P \cap \Lambda)$. By coherence (Proposition 6), we should have $\hat{P}_{i+1} \supset \tilde{P}_{i+1}$ and $[m_{i+1}, m_{i+1} + \hat{n}_{i+1}]$ would be an \tilde{I} -interval larger than $[m_{i+1}, p_{i+1}]$, a contradiction.

As $(P_{i+1}, Q_{i+1}, n_{i+1})$ is \tilde{I} -bicritical, and the parent interval \tilde{I} is always assumed to be β -regular, we have, for all $t \in \tilde{I}$

$$(6.61) \quad |P_{i+1}| < |\tilde{I}|^\beta$$

and thus also (with ε_0 small enough)

$$(6.62) \quad 2|P_{i+1}|^{1-\eta} < |I|.$$

It now follows from Proposition 9 and $Q^{(i)} \pitchfork_I \tilde{P}_{i+1}$ that $Q^{(i)}$ and P_{i+1} are I -transverse.

6.7.10 When $i = 0$, $Q^{(0)} = Q_0$ and we have already shown that Q_i and P_{i+1} are I -transverse.

When $i > 0$, (P_i, Q_i, n_i) is \tilde{I} -bicritical and, therefore, we have for all $t \in \tilde{I}$:

$$(6.63) \quad |Q_i| < |\tilde{I}|^\beta,$$

and thus also

$$(6.64) \quad 2|Q_i|^{1-\eta} < |I|.$$

It follows from Proposition 9 and $Q^{(i)} \pitchfork_I P_{i+1}$ that Q_i and P_{i+1} are I -transverse.

To conclude the induction step of 6.7.6, we simply observe that the parabolic composition of $(P^{(i)}, Q^{(i)}, n^{(i)})$ and $(P_{i+1}, Q_{i+1}, n_{i+1})$ is allowed in $\mathcal{R}(I)$; it produces an element $(P^{(i+1)}, Q^{(i+1)}, p_{i+1}) \in \mathcal{R}(I)$ such that $P^{(i+1)}$ intersects P and therefore contains P .

The proof of the theorem is now complete.

6.7.11 In the next two Corollaries, the setting and notations are those of Theorem 1.

Corollary 6. *For all $t \in I$, we have*

$$|P| \leq C^k |P_0| |P_1| \cdots |P_k| |I|^{-\frac{k}{2}}$$

Proof. We have $P = P^{(k)}$, with $P^{(i)}$ defined in (6.57); we prove that, for all $t \in I$

$$(6.65) \quad |P^{(i)}| \leq C^i |P_0| \cdots |P_i| |I|^{-\frac{i}{2}}.$$

As $P^{(0)} = P_0$, this is true for $i = 0$. The induction step is a consequence of the key estimate (3.27) for parabolic composition if we know that, for all $t \in I$:

$$(6.66) \quad \delta(Q^{(i)}, P_{i+1}) \geq |I|.$$

As Q_i and P_{i+1} are I -transverse, there exists $I^* \supset I$, (P_i^*, Q_i^*, n_i^*) , $(P_{i+1}^*, Q_{i+1}^*, n_{i+1}^*) \in \mathcal{R}(I^*)$ such that $Q_i \subset Q_i^*$, $P_{i+1} \subset P_{i+1}^*$ and $Q_i^* \bar{\pitchfork}_{I^*} P_{i+1}^*$ holds. From (T1) in Subsection 5.4, we have, for all $t \in I^*$

$$(6.67) \quad \delta_{LR}(Q_i^*, P_{i+1}^*) \geq 2|I^*|.$$

From (6.3), (5.11)–(5.14) and (R7), we get, for all $t \in I$:

$$(6.68) \quad \begin{aligned} \delta(Q^{(i)}, P_{i+1}) &\geq \delta(Q_i^*, P_{i+1}^*) \\ &\geq \delta_{LR}(Q_i^*, P_{i+1}^*) - c(|P_{i+1}^*| + |Q_i^*|) \\ &\geq \frac{1}{2} \delta_{LR}(Q_i^*, P_{i+1}^*) \\ &\geq |I^*| \geq |I|, \end{aligned}$$

which concludes the proof. □

Corollary 7. *For all $0 < i < k$, and $t \in \tilde{I}$, we have*

$$|P_i| < |\tilde{I}|^\beta.$$

For all $t \in I$, we also have

$$|P_k| < c|\tilde{I}|^{(1-\eta)^{-1}}.$$

Proof. The first assertion is an immediate consequence of the regularity of the parameter interval \tilde{I} , because (P_i, Q_i, n_i) is \tilde{I} -bicritical for $0 < i < k$.

For the second assertion, we first observe that in the proof of Corollary 6, we must have $I^* = I$ because $Q_{k-1} \cap_{\tilde{I}} P_k$ does not hold; for the same reason, there exists $t^* \in \tilde{I}$ such that

$$(6.69) \quad \delta_{LR}(Q_{k-1}^*, P_k^*) < 2|\tilde{I}|.$$

But then, from (6.6) and Corollary 11 in Subsection 7.6, we have, for all $t \in \tilde{I}$:

$$(6.70) \quad \begin{aligned} \delta(Q_{k-1}, P_k) &< \delta_{LR}(Q_{k-1}, P_k) \\ &< \delta_{LR}(Q_{k-1}^*, P_k^*) \\ &< c|\tilde{I}|. \end{aligned}$$

From (R7), we have, for all $t \in I$

$$(6.71) \quad \delta(Q_{k-1}, P_k) > c^{-1}|P_k|^{1-\eta}$$

and the second assertion of the Corollary follows. \square

6.7.12

Corollary 8. *Any (P, Q, n) in $\mathcal{R}(I)$ but not in $\mathcal{R}(\tilde{I})$ satisfies, for all $t \in I$:*

$$|P| \leq |\tilde{I}|^{\frac{1}{2}}, \quad |Q| \leq |\tilde{I}|^{\frac{1}{2}}.$$

Proof. From (6.70), (6.71) and (3.27), we actually have, for all $t \in I$

$$(6.72) \quad |P| \leq |P_0| |\tilde{I}|^{\frac{1}{2}},$$

and we must have $|P_0| \ll 1$ as $Q_0 \subset Q_u$. \square

6.7.13

Corollary 9. *Any candidate interval I is $\bar{\beta}$ -regular, with $\bar{\beta} = \beta(1 + \tau)^{-1}$.*

Proof. Let (P, Q, n) be an I -bicritical element of $\mathcal{R}(I)$. If (P, Q, n) belongs to $\mathcal{R}(\tilde{I})$, then it is \tilde{I} -bicritical and we have, for all $t \in \tilde{I}$:

$$(6.73) \quad \max(|P|, |Q|) < |\tilde{I}|^{\bar{\beta}} = |I|^{\bar{\beta}}.$$

Assume now that (P, Q, n) does not belong to $\mathcal{R}(\tilde{I})$. We apply Theorem 1. The element $(P_0, Q_0, n_0) \in \mathcal{R}(\tilde{I})$ is \tilde{I} -bicritical: Q_0 is \tilde{I} -critical by Theorem 1, and P_0 is \tilde{I} -critical because P is I -critical. Applying Corollaries 6 and 7 gives, for all $t \in I$

$$(6.74) \quad \begin{aligned} |P| &\leq C|P_0| |P_k| |I|^{-\frac{1}{2}} \\ &\leq C|\tilde{I}|^{\beta+1-\frac{1}{2}(1+\tau)}, \end{aligned}$$

with $\beta + 1 - \frac{1}{2}(1 + \tau) > \beta + \frac{1}{3}$, so I -bicritical elements are much thinner in this case. \square

Remark. *The same phenomenon will be important again in Section 9: "fat" bicritical elements were created much earlier.*

6.7.14 The last result in this section is a complement to Proposition 5 in Subsection 6.4.

Let I be a parameter interval, and let $(\tilde{P}, \tilde{Q}, \tilde{n}), (P, Q, n)$ be elements of $\mathcal{R}(I)$ such that P is a non-simple child of \tilde{P} . From Proposition 5, we know that there exists $(P_1, Q_1, n_1) \in \mathcal{R}(I)$ such that $\tilde{Q} \pitchfork_I P_1$ holds and

$$(6.75) \quad (P, Q, n) \in (\tilde{P}, \tilde{Q}, \tilde{n}) \square (P_1, Q_1, n_1).$$

Proposition 12. *Assume moreover that \tilde{P} is I -critical. Then, for all $t \in I$ we have*

$$\delta(\tilde{Q}, P_1) \geq |\tilde{P}|^{\frac{1}{\beta}}.$$

Proof. Consider first the case where \tilde{Q} is I -critical. Then $(\tilde{P}, \tilde{Q}, \tilde{n})$ is I -bicritical. From Corollary 9, we know that I is $\bar{\beta}$ -regular and therefore we have, for all $t \in I$:

$$(6.76) \quad |\tilde{P}| < |I|^{\bar{\beta}}.$$

On the other hand, as $\tilde{Q} \pitchfork_I P_1$ holds, we can find $(\tilde{P}^*, \tilde{Q}^*, \tilde{n}^*), (P_1^*, Q_1^*, n_1^*) \in \mathcal{R}(I)$ with $\tilde{Q} \subset \tilde{Q}^*$, $P_1 \subset P_1^*$ such that, for all $t \in I$

$$(6.77) \quad \delta_{LR}(\tilde{Q}^*, P_1^*) \geq 2|I|.$$

From (R7) (cf. Subsection 5.4) and (5.11)–(5.14), it now follows that

$$(6.78) \quad \begin{aligned} \delta(\tilde{Q}, P_1) &\geq \delta(\tilde{Q}^*, P_1^*) \\ &\geq \frac{1}{2} \delta_{LR}(\tilde{Q}^*, P_1^*) \\ &\geq |I| \\ &\geq |\tilde{P}|^{\frac{1}{\beta}}. \end{aligned}$$

From now on we assume that \tilde{Q} is I -transverse. Let $I^* \supset I$ be the largest parameter interval such that $(\tilde{P}, \tilde{Q}, \tilde{n}) \in \mathcal{R}(I^*)$ and \tilde{Q} is I^* -transverse. As the transversality relation never holds for the

starting interval $[\varepsilon_0, 2\varepsilon_0] = I_0$, I^* is not equal to I_0 . Let \tilde{I}^* be the parent of I^* . We claim that for all $t \in I^*$, we have:

$$(6.79) \quad |\tilde{P}| \leq |\tilde{I}^*|^\beta = |I^*|^{\bar{\beta}}.$$

Indeed, from the definition of I^* , we have that either $(\tilde{P}, \tilde{Q}, \tilde{n}) \notin \mathcal{R}(\tilde{I}^*)$, or $(\tilde{P}, \tilde{Q}, \tilde{n}) \in \mathcal{R}(\tilde{I}^*)$ and \tilde{Q} is \tilde{I}^* -critical. In the second case, $(\tilde{P}, \tilde{Q}, \tilde{n})$ is \tilde{I}^* -bicritical and \tilde{I}^* is β -regular, which gives (6.79). In the first case, we obtain from (6.74) in the proof of Corollary 9 (where only the I -criticality of P was used) that, for all $t \in I^*$

$$(6.80) \quad |\tilde{P}| \leq C|\tilde{I}^*|^{\beta+\frac{1}{3}} < |I^*|^{\bar{\beta}}.$$

The claim is proved.

As \tilde{Q} is I^* -transverse, there exists an \tilde{I} -decomposition $(P_\alpha, Q_\alpha, n_\alpha)$ of P_s in $\mathcal{R}(I^*)$ such that, for any α , \tilde{Q} and P_α are either I^* -separated or I^* -transverse. There exists α_0 such that P_1 and P_{α_0} intersect; \tilde{Q} and P_{α_0} must be I^* -transverse. This implies, as for (6.77) above, that we have, for all $t \in I^*$:

$$(6.81) \quad \delta(\tilde{Q}, P_{\alpha_0}) \geq \frac{3}{2} |I^*|.$$

If $P_1 \subset P_{\alpha_0}$, we have

$$(6.82) \quad \delta(\tilde{Q}, P_1) \geq \delta(\tilde{Q}, P_{\alpha_0}).$$

If $P_{\alpha_0} \subset P_1$, we have from (5.11)–(5.14) and (R7), for all $t \in I$:

$$(6.83) \quad \begin{aligned} \delta(\tilde{Q}, P_1) &\geq \frac{2}{3} \delta_R(\tilde{Q}, P_1) \\ &\geq \frac{2}{3} \delta_R(\tilde{Q}, P_{\alpha_0}) \\ &\geq \frac{2}{3} \delta(\tilde{Q}, P_{\alpha_0}). \end{aligned}$$

In all cases, combining this with (6.81) and (6.79) gives the required estimate. \square

7 Estimates for the Classes $\mathcal{R}(I)$

7.1 A Stretched Exponential Estimate for Widths

The next proposition is a substitute for the uniform exponential estimates for widths that are characteristic of the uniformly hyperbolic dynamics. We denote by γ the constant

$$(7.1) \quad \gamma := \frac{\log \frac{3}{2}}{\log 2} \in (0, 1).$$

Proposition 13. *Let I be a parameter interval and let (P, Q, n) be an element of $\mathcal{R}(I)$. For all $t \in I$, we have*

$$|P| \leq C \exp(-n^\gamma)$$

with the stronger estimate

$$|P| \leq C \exp(-2n^\gamma)$$

when the parent of P is I -critical.

Proof. If I is the starting interval $I_0 = [\varepsilon_0, 2\varepsilon_0]$, a stronger exponential bound actually holds. The proof is by induction on n , and the estimates are therefore valid when $n = O(\log \varepsilon_0^{-1})$ (in which case $(P, Q, n) \in \mathcal{R}(I_0)$). Let $(\tilde{P}, \tilde{Q}, \tilde{n})$ be the element of $\mathcal{R}(I)$ such that \tilde{P} is the parent of P . In the case where P is a simple child of \tilde{P} , the bound for P easily follows from the bound for \tilde{P} . Let us therefore assume that P is a non-simple child of \tilde{P} , in which case there exists by Proposition 5 an element $(P_1, Q_1, n_1) \in \mathcal{R}(I)$ such that $\tilde{Q} \pitchfork_I P_1$ holds, the parent of P_1 is I -critical and

$$(7.2) \quad (P, Q, n) \in (\tilde{P}, \tilde{Q}, \tilde{n}) \square (P_1, Q_1, n_1).$$

By the induction hypothesis, we have

$$(7.3) \quad |\tilde{P}| \leq C \exp(-\tilde{n}^\gamma),$$

$$(7.4) \quad |P_1| \leq C \exp(-2n_1^\gamma).$$

From (R7) and (3.27), we have, for all $t \in I$:

$$(7.5) \quad |P| \leq C |\tilde{P}| |P_1|^{\frac{1+\eta}{2}} \ll |\tilde{P}| |P_1|^{\frac{1}{2}},$$

and this gives

$$(7.6) \quad |P| \leq \exp(-\tilde{n}^\gamma - n_1^\gamma).$$

This proves the required estimate in the general case, because, when \tilde{n} and n_1 are large, we have

$$(7.7) \quad \tilde{n}^\gamma + n_1^\gamma \geq (\tilde{n} + n_1 + N_0)^\gamma = n^\gamma.$$

Assume now that \tilde{P} is I -critical. Instead of (7.3), we have

$$(7.8) \quad |\tilde{P}| \leq C \exp(-2\tilde{n}^\gamma),$$

and then from (7.5) we obtain

$$(7.9) \quad |P| \leq \exp(-2\tilde{n}^\gamma - n_1^\gamma).$$

This will be useful when $\tilde{n} \geq n_1$. When $\tilde{n} \leq n_1$, we prefer to rely on Proposition 12 which gives

$$(7.10) \quad \delta(\tilde{Q}, P_1) \geq |\tilde{P}|^{1/\beta} \gg |\tilde{P}|.$$

When we use this in (3.27) and combine with (7.4) and (7.8), we get

$$(7.11) \quad |P| \leq \exp(-\tilde{n}^\gamma - 2n_1^\gamma).$$

To get the required estimate from (7.9), (7.11), we have only to observe that the function $u \mapsto u^\gamma + 2(1-u)^\gamma$ is concave on $[0, \frac{1}{2}]$ and equal to 2 for $u = 0$ and $u = \frac{1}{2}$. \square

7.2 Uniform Cone Condition

In this subsection, we will check that all elements $(P, Q, n) \in \mathcal{R}(I)$ satisfy the cone condition (AL2) of Subsection 3.2 for the parameters λ, u_0, v_0 of Subsection 5.3: we have $u_0 = u^{1/2}$, $v_0 = v^{1/2}$, and all $(P, Q, n) \in \mathcal{R}([\varepsilon_0, 2\varepsilon_0])$ satisfy (AL2) with parameters λ, u, v .

Let (A, B) be the implicit representation of the affine-like iterate (P, Q, n) ; we have to prove that

$$(AL2) \quad \begin{aligned} \lambda|A_x| + u_0|A_y| &\leq 1, \\ \lambda|B_y| + v_0|B_x| &\leq 1. \end{aligned}$$

Let $u_1 = u^{\frac{3}{4}}$, $v_1 = v^{\frac{3}{4}}$. We will prove that, for all $t \in I$, we have

$$(7.12) \quad |A_y| < u_1^{-1}, \quad |B_x| < v_1^{-1}.$$

This is sufficient to obtain (AL2): we already know that if $(P, Q, n) \in \mathcal{R}(I_0)$ then (AL2) is satisfied; on the other hand, if $(P, Q, n) \notin \mathcal{R}(I_0)$, then, for all $t \in I$, we have from Corollary 8

$$(7.13) \quad |A_x| < C\varepsilon_0^{\frac{1}{2}}, \quad |B_y| < C\varepsilon_0^{\frac{1}{2}};$$

with ε_0 small enough, (7.13) and (7.12) give (AL2).

Let us now proceed with the proof of (7.12). When $(P, Q, n) \in \mathcal{R}(I_0)$, we have the stronger estimate:

$$(7.14) \quad |A_y| < u^{-1}, \quad |B_x| < v^{-1}.$$

Let $(\tilde{P}, \tilde{Q}, \tilde{n})$ be the element of $\mathcal{R}(I)$ such that \tilde{P} is the parent of P . Denote by (\tilde{A}, \tilde{B}) the implicit representation of the corresponding affine-like iterate.

If P is a simple child, we use formula (3.11) of Subsection 3.3 to obtain

$$(7.15) \quad |A_y - \tilde{A}_y| \leq C |\tilde{P}| |\tilde{Q}|.$$

If P is a non-simple child, it is obtained (Proposition 5) by the parabolic composition of $(\tilde{P}, \tilde{Q}, \tilde{n})$ with some (P_1, Q_1, n_1) ; we use formula (3.30) of Subsection 3.5 to obtain

$$(7.16) \quad |A_y - \tilde{A}_y| \leq C |\tilde{P}| |\tilde{Q}| (\delta(\tilde{Q}, P_1))^{-\frac{1}{2}}.$$

From (R7), $\delta(\tilde{Q}, P_1)$ is much larger than $|\tilde{Q}|$. In all cases, we have

$$(7.17) \quad |A_y - \tilde{A}_y| \leq C |\tilde{P}| \leq C \exp(-\tilde{n}^\gamma),$$

where we have used Proposition 13 in the last inequality. We only need (7.17) when \tilde{n} is large (because we already have (7.14) otherwise), and the series $\sum \exp(-n^\gamma)$ is convergent. Therefore (7.12) is a consequence of (7.14) and (7.17).

The proof of (AL2), i.e., the first part of condition (R2) in Subsection 5.3, is now complete.

7.3 Bounded Distortion

We now check the second half of property (R2) in Subsection 5.3. We have to prove that, for all $(P, Q, n) \in \mathcal{R}(I)$, we have the following estimate on distortion:

$$(7.18) \quad D(g_t^n/P) \leq 2D_0.$$

Here, the constant D_0 corresponds to the stronger estimate we obtain from (MP6) when $(P, Q, n) \in \mathcal{R}(I_0)$:

$$(7.19) \quad D(g_t^n/P) \leq D_0.$$

For $m > 0$, define

$$(7.20) \quad D(m) = \sup_{\substack{(P, Q, n) \in \mathcal{R}(I) \\ n \leq m}} \sup_{t \in I} D(g_t^n/P).$$

To obtain, by induction on m , a bound for the non-decreasing sequence $D(m)$, we combine (7.19), (which gives $D(m) \leq D_0$ for $m = O(\log \varepsilon_0^{-1})$), Proposition 13 and the bounds (3.13) in Subsection 3.3 (for simple composition) and (3.29) in Subsection 3.5 (for parabolic composition). We set

$$(7.21) \quad D_s(m) = \max_{\substack{n > 0, n' > 0 \\ n + n' \leq m}} D_s(n, n')$$

with

$$(7.22) \quad D_s(n, n') = D(n) + c \exp(-n^\gamma)(D(n) + D(n')).$$

We also set

$$(7.23) \quad D_p(m) = \max_{\substack{n \gg 0, n' \gg 0 \\ n+n'+N_0 \leq m}} D_p(n, n')$$

with

$$(7.24) \quad D_p(n, n') = D(n) + c \exp(-\eta n^\gamma).$$

Then, we have, as long as $D(m)$ is not too large (cf. the condition for (3.29) to hold)

$$(7.25) \quad D(m) \leq \max(D_s(m), D_p(m)).$$

The reason for formula (7.24) to hold is that the term $c|P_1|\delta^{-1}$ of (3.29) is smaller than $|P_1|^\eta$ by (R7); then one uses Proposition 13.

It is now clear that (7.18) follows from (7.19), (7.21)–(7.25).

7.4 Estimates for the Special Rectangles P_s and Q_u

In the next subsection, we will check the estimates contained in condition (R4) of Subsection 5.3 concerning the class $\mathcal{R}(I)$.

These estimates, which are related to parabolic composition, are valid for an element (P, Q, n) of $\mathcal{R}(I)$ which satisfies $Q \subset Q_u$ (or $P \subset P_s$).

In the present section, we will be concerned with the affine-like iterates which are directly associated with the elements (P_s, Q_s, n_s) and (P_u, Q_u, n_u) .

We will make the computations for (P_s, Q_s, n_s) the other case is obviously symmetric. We will assume that the periodic point p_s is *fixed*: the general case is completely similar, but the notations are more awkward.

In this subsection, we just write (x, y) for the coordinates in the rectangle R_{a_s} containing p_s ; we denote by (A, B) the implicit representation of the affine-like iterate

$$(7.26) \quad G_t : (R_{a_s}) \cap g_t^{-1}(R_{a_s}) \longrightarrow g_t(R_{a_s}) \cap R_{a_s}.$$

For $n \geq 0$, we denote by $(A^{(n)}, B^{(n)})$ the implicit representation of the n^{th} iterate of this restriction.

As the equation of $W_{\text{loc}}^s(p_s)$ is $\{x = 0\}$ (cf. (MP3) in Subsection 2.2), we have

$$(7.27) \quad A(y, 0, t) \equiv 0,$$

from which we deduce

$$(7.28) \quad \begin{aligned} |A_y(y, x, t)| &\leq c|x|, \\ |A_t(y, x, t)| &\leq c|x|, \\ |A_{yy}(y, x, t)| &\leq c|x|, \\ |A_{yt}(y, x, t)| &\leq c|x|. \end{aligned}$$

Denote by $\mu = \mu(t)$ the unstable eigenvalue of D_{g_t} at p_s . For all t, x, y, n , we have

$$(7.29) \quad c^{-1}\mu^{-n} \leq |A_x^{(n)}(x, y, t)| \leq c\mu^{-n}.$$

Let $(x_i, y_i)_{0 \leq i \leq n}$ be an orbit of g_t in R_{a_s} . For all $0 \leq \ell \leq m \leq n$, we have:

$$(7.30) \quad c^{-1}\mu^{m-\ell}|x_\ell| \leq |x_m| \leq c\mu^{m-\ell}|x_\ell|.$$

Proposition 14. *The following estimates hold:*

$$(7.31) \quad |A_y^{(n)}(y_0, x_n, t)| \leq c|x_0| \leq c\mu^{-n}|x_n|,$$

$$(7.32) \quad |A_t^{(n)}(y_0, x_n, t)| \leq cn|x_0| \leq cn\mu^{-n}|x_n|,$$

$$(7.33) \quad |A_{yy}^{(n)}(y_0, x_n, t)| \leq c|x_0| \leq c\mu^{-n}|x_n|.$$

Proof of 7.31: From formula (3.11) in Subsection 3.3, we have:

$$(7.34) \quad A_y^{(n)}(y_0, x_n, t) = A_y^{(n-1)}(y_0, x_{n-1}, t) + A_y A_x^{(n-1)} B_y^{(n-1)} \Delta^{-1},$$

with $B_y^{(n-1)} \Delta$ exponentially small with n and, using (7.28)–(7.30):

$$(7.35) \quad |A_y(y_{n-1}, x_n, t) A_x^{(n-1)}(y_0, x_{n-1}, t)| \leq c|x_0|.$$

The inequality (7.31) is now clear.

Proof of 7.32: We use here formulas (A6), (A10) of Appendix A which give

$$(7.36) \quad \begin{aligned} &|A_t^{(n)}(y_0, x_n, t) - A_t^{(n-1)}(y_0, x_{n-1}, t)| \\ &\leq C\mu^{-n} \left(|A_t(y_{n-1}, x_n, t)| + |B_t^{(n-1)}(y_0, x_{n-1}, t)| |A_y(y_{n-1}, x, t)| \right), \end{aligned}$$

$$(7.37) \quad |B_t^{(n)}(y_0, x_n, t) - B_t^{(n-1)}(y_0, x_{n-1}, t)| \leq C |B_y^{(n-1)}|_{C^0} \left(1 + |A_t^{(n-1)}|_{C^0} \right).$$

As $|B_y^{(n)}|_{C_0}$ is exponentially small, we deduce from (7.37) that

$$(7.38) \quad |B_t^{(n)}(y_0, x_n, t)| \leq C,$$

and then, from (7.36), (7.28) that (7.32) holds.

Proof of 7.33: We use formulas (A6), (A18), (A20) of Appendix A to obtain

$$(7.39) \quad A_{yy}^{(n)} = A_{yy}^{(n-1)} + 2A_{xy}^{(n-1)}X_y + A_{xx}^{(n-1)}X_y^2 + A_x^{(n-1)}X_{yy},$$

with

$$(7.40) \quad X_y = A_y B_y^{(n-1)} \Delta^{-1},$$

$$(7.41) \quad \Delta = 1 - A_y B_x^{(n-1)},$$

$$(7.42) \quad X_{yy} = B_y^{(n-1)} \Delta^{-1} \left(A_{yy} B_y^{(n-1)} \Delta^{-1} + A_y \partial_y \log |B_y^{(n-1)}| + \right. \\ \left. + A_y X_y \partial_x \log |B_y^{(n-1)}| - A_y^{(n-1)} \Delta_y \Delta^{-1} \right),$$

$$(7.43) \quad -\Delta_y = A_{yy} B_y^{(n-1)} B_x^{(n-1)} \Delta^{-1} + A_y B_{xy}^{(n-1)} + A_y B_{xx}^{(n-1)} X_y.$$

In these formulas, $A^{(n-1)}$, $B^{(n-1)}$ and their derivatives are taken at (y_0, x_{n-1}, t) , A , B and their derivatives are taken at (y_{n-1}, x_n, t) . The terms $B_x^{(n-1)}$, $B_{xx}^{(n-1)}$, $\partial_x \log |B_y^{(n-1)}|$, $\partial_y \log |B_y^{(n-1)}|$, Δ^{-1} are bounded by the uniform cone condition and the uniform distortion; the terms $B_y^{(n-1)}$, $B_{xy}^{(n-1)}$ are exponentially small. Also, from (7.28) we have:

$$(7.44) \quad |A_y(y_{n-1}, x_n, t)| \leq C|x_n|, \\ |A_{yy}(y_{n-1}, x_n, t)| \leq C|x_n|.$$

We conclude that we can write

$$(7.45) \quad A_{yy}^{(n)}(y_0, x_n, t) = A_{yy}^{(n)}(y_0, x_{n-1}, t) + \mu^{-n} x_n r_n$$

with r_n exponentially small; this leads to (7.34). □

Corollary 10. *For the special rectangle (P_s, Q_s, n_s) , we have:*

$$|A_y^{(n_s)}|_{C^0} \leq C\varepsilon_0, \\ |A_{yy}^{(n_s)}|_{C^0} \leq C\varepsilon_0, \\ |A_t^{(n_s)}|_{C^0} \leq C\varepsilon_0 \log \varepsilon_0^{-1}.$$

Proof. We have only to observe that μ^{-n_s} is of order ε_0 . □

Obviously, the same estimates hold for the other special element (P_u, Q_u, n_u) .

7.5 Proof of the Property (R4) of Section 5.3

We have to show that, for an element (P, Q, n) in $\mathcal{R}(I)$, with associated implicit representation (A, B) , we have

$$(7.46) \quad |A_y| \leq C\varepsilon_0, \quad |A_{yy}| \leq C\varepsilon_0$$

whenever $P \subset P_s$ and

$$(7.47) \quad |B_x| \leq C\varepsilon_0, \quad |B_{xx}| \leq C\varepsilon_0$$

whenever $Q \subset Q_u$. We will deal only with (7.46), the other case being symmetric.

We have already proved (7.46) when $P = P_s$, cf. the first two inequalities of Corollary 10. We will prove (7.46) by induction on the length n . Let us denote by $(\tilde{P}, \tilde{Q}, \tilde{n})$ the element of $\mathcal{R}(I)$ such that \tilde{P} is the parent of P , by (\tilde{A}, \tilde{B}) the implicit representation of this affine-like iterate. We assume now that $P \subsetneq P_s$, i.e., $\tilde{P} \subset P_s$. There are two possibilities.

First, we consider the easier case when P is a simple child of \tilde{P} . Let $(P^*, Q^*, 1)$, with associated implicit representation (A^*, B^*) , be the element such that

$$(7.48) \quad (P, Q, n) = (\tilde{P}, \tilde{Q}, \tilde{n}) * (P^*, Q^*, 1).$$

From formula (3.11) in Subsection 3.3, we have:

$$(7.49) \quad A_y = \tilde{A}_y + \tilde{A}_x \tilde{B}_y A_y^* \Delta^{-1},$$

with $|\tilde{A}_x| \leq C\varepsilon_0$, A_y^* , Δ^{-1} bounded and \tilde{B}_y satisfying the stretched exponential estimate of Proposition 13: this is fully in line with the first inequality in (7.46).

Next, we have, as in (7.39) above:

$$(7.50) \quad A_{yy} = \tilde{A}_{yy} + 2\tilde{A}_{xy} X_y + \tilde{A}_{xx} X_y^2 + \tilde{A}_x X_{yy},$$

with

$$(7.51) \quad X_y = A_y^* \tilde{B}_y \Delta^{-1},$$

$$(7.52) \quad \Delta = 1 - A_y^* \tilde{B}_x,$$

$$(7.53) \quad X_{yy} = \Delta^{-1} \tilde{B}_y \left(A_{yy}^* \tilde{B}_y \Delta^{-1} + A_y^* \partial_y \log |\tilde{B}_y| + A_y^* X_y \partial_x \log |\tilde{B}_y| - \tilde{A}_y \Delta_y \Delta^{-1} \right),$$

$$(7.54) \quad -\Delta_y = A_{yy}^* \tilde{B}_y \tilde{B}_x \Delta^{-1} + A_y^* \tilde{B}_{xy} + A_y^* \tilde{B}_{xx} X_y.$$

Here, we have:

$$(7.55) \quad |\tilde{A}_{xy}| = |\tilde{A}_x| |\partial_y \log |\tilde{A}_x|| \leq C\varepsilon_0,$$

$$(7.56) \quad |\tilde{A}_{xx}| = |\tilde{A}_x| |\partial_x \log |\tilde{A}_x|| \leq C\varepsilon_0,$$

$$(7.57) \quad |\tilde{A}_x| \leq C\varepsilon_0.$$

Moreover, \tilde{B}_y satisfies the stretched exponential estimate of Proposition 13; the uniform cone condition and uniform distortion imply that the same stretched exponential estimate holds for X_y and X_{yy} . Therefore, the estimate for $A_{yy} - \tilde{A}_{yy}$ in (7.50) is fully in line with the second inequality in (7.46). This concludes the case where P is a simple child.

We now consider the more difficult case where P is a non-simple child of \tilde{P} . By Proposition 5, we can find an element $(P^*, Q^*, n^*) \in \mathcal{R}(I)$, with associated implicit representation (A^*, B^*) , such that:

$$(7.58) \quad (P, Q, n) \in (\tilde{P}, \tilde{Q}, \tilde{n}) \square (P^*, Q^*, n^*).$$

From the formulas in Appendix A, we have:

$$(7.59) \quad A_y = \tilde{A}_y + \tilde{A}_x(X_y + X_w W_y),$$

$$(7.60) \quad X_y = X_{u,y} \tilde{B}_y \Delta_0^{-1},$$

$$(7.61) \quad X_w = X_{u,w} \Delta_0^{-1},$$

$$(7.62) \quad \Delta_0 = 1 - X_{u,y} \tilde{B}_x,$$

$$(7.63) \quad W_y = -C_y C_w^{-1},$$

$$(7.64) \quad -C_y = \theta_y \tilde{B}_y \Delta_0^{-1}.$$

Here θ , X_u , W and C are associated with the fold map G as in Subsections 2.3, 3.5.

As $|C_{ww} - 2|$ is small (cf. 3.22), the quantity C_w is related to $\delta(\tilde{Q}, P^*)$ by

$$(7.65) \quad C^{-1} \delta(\tilde{Q}, P^*)^{\frac{1}{2}} \leq |C_w| \leq C \delta(\tilde{Q}, P^*)^{\frac{1}{2}}.$$

From (R7), $\delta(\tilde{Q}, P^*)$ is always much bigger than $|\tilde{B}_y|$; therefore, from (7.63), (7.64) we have:

$$(7.66) \quad |W_y| \leq C |\tilde{B}_y|^{\frac{1}{2}}.$$

We then obtain from Proposition 13, that the term $X_y + X_w W_y$ in (7.59) satisfies a stretched exponential estimate; in view of (7.57), this concludes the proof of the first inequality in (7.46).

For the second inequality, we write, following the formulas in Appendix A:

$$(7.67) \quad \begin{aligned} A_{yy} &= \tilde{A}_{yy} + 2\tilde{A}_{xy}(X_y + X_w W_y) + \tilde{A}_{xx}(X_y + X_w W_y)^2 + \\ &\quad + \tilde{A}_x(X_{yy} + 2X_{wy}W_y + X_{ww}W_y^2 + X_w W_{yy}), \end{aligned}$$

$$(7.68) \quad W_{yy} = -C_w^{-1}(C_{ww}W_y^2 + 2C_{wy}W_y + C_{yy}),$$

$$(7.69) \quad -C_{wy} = \theta_{yy}\tilde{B}_x\tilde{B}_yX_w\Delta_0^{-1} + \theta_{xy}\tilde{B}_yA_y^*Y_{s,w}\Delta_0^{-1}\Delta_1^{-1} + \theta_y\bar{Y}_{wy},$$

$$(7.70) \quad -C_{yy} = \theta_{yy}\tilde{B}_y^2\Delta_0^{-2} + \theta_y\bar{Y}_{yy},$$

$$(7.71) \quad \bar{Y}_{wy} = \tilde{B}_{xy}X_w + \tilde{B}_{xx}X_wX_y + \tilde{B}_xX_{wy},$$

$$(7.72) \quad \bar{Y}_{yy} = \tilde{B}_{yy} + 2\tilde{B}_{xy}X_y + \tilde{B}_{xx}X_y^2 + \tilde{B}_xX_{wy},$$

$$(7.73) \quad X_{ww} = \Delta_0^{-1}(X_{u,ww} + 2\tilde{B}_xX_wX_{u,wy} + \tilde{B}_x^2X_{u,yy}X_w^2 + \tilde{B}_{xx}X_{u,y}X_w^2),$$

$$(7.74) \quad \begin{aligned} X_{wy} &= \Delta_0^{-1}\left[(X_{u,wy} + \tilde{B}_xX_{u,yy}X_w)(\tilde{B}_y + \tilde{B}_xX_y) \right. \\ &\quad \left. + X_{u,y}X_w(\tilde{B}_{xy} + \tilde{B}_{xx}X_y)\right], \end{aligned}$$

$$(7.75) \quad X_{yy} = \Delta_0^{-1}\left[X_{u,yy}(\tilde{B}_y + \tilde{B}_xX_y)^2 + X_{u,y}(\tilde{B}_{yy} + 2\tilde{B}_{xy}X_y + \tilde{B}_{xx}X_y^2)\right],$$

where $\Delta_0^{-1}, \Delta_1^{-1}$ are bounded. In formula (7.67), we have that

- $|\tilde{A}_{xy}|, |\tilde{A}_{xx}|, |\tilde{A}_x|$ are bounded by $C\varepsilon_0$, cf. (7.55)–(7.57) above;
- X_w, X_{ww} are bounded, cf. (7.61) and (7.73);
- X_y, W_y satisfy a stretched exponential estimate, as we have already seen earlier;
- X_{wy}, X_{yy} also satisfy a stretched exponential estimate, cf. (7.74) and (7.75).

The remaining term in (7.67) is $\tilde{A}_xX_wW_{yy}$, and we have to estimate W_{yy} from (7.68). From (7.75), we have, using bounded distortion:

$$(7.76) \quad |X_{yy}| \leq C|\tilde{B}_y|,$$

from which we deduce, by (7.72), that

$$(7.77) \quad |\bar{Y}_{yy}| \leq C|\tilde{B}_y|.$$

In the same way, we obtain

$$(7.78) \quad |X_{wy}| \leq C|\tilde{B}_y|,$$

$$(7.79) \quad |\bar{Y}_{wy}| \leq C|\tilde{B}_y|.$$

Plugging this into (7.69), (7.70) yields

$$(7.80) \quad |C_{wy}| \leq C|\tilde{B}_y|,$$

$$(7.81) \quad |C_{yy}| \leq C|\tilde{B}_y|.$$

We can now conclude, using (7.65), (7.66), that

$$(7.82) \quad \begin{aligned} |W_{yy}| &\leq C|\tilde{B}_y|\delta(\tilde{Q}, P^*)^{-\frac{1}{2}}, \\ &\leq C|\tilde{B}_y|^{\frac{1}{2}}. \end{aligned}$$

All terms in (7.67) are now under control, $A_{yy} - \tilde{A}_{yy}$ being bounded by $C\varepsilon_0$ times a stretched exponentially small term: this concludes the proof of the second inequality in (7.46), and so of condition (R4).

7.6 Relative Speeds of Critical Rectangles

Let I be a parameter interval, and let $(P_0, Q_0, n_0), (P_1, Q_1, n_1)$ be elements of $\mathcal{R}(I)$ such that $Q_0 \subset Q_u, P_1 \subset P_s$.

The displacements $\delta(Q_0, P_1), \delta_L(Q_0, P_1), \delta_R(Q_0, P_1), \delta_{LR}(Q_0, P_1)$ were introduced in formulas (5.5) and (5.8)–(5.10) of Subsection 5.4 (see also (3.23) in Subsection 3.5) and are the values at the four corners of the rectangle of definition of the function $\bar{C}(y_0, x_1)$ introduced in Subsection 3.5 as

$$(7.83) \quad \bar{C}(y_0, x_1) = \min_w C(w, y_0, x_1).$$

All these quantities also depend on the parameter t , and we want in this section to estimate the variation with the parameter of the displacements, which amounts to estimate the partial derivative C_t .

Let $(A_0, B_0), (A_1, B_1)$ be the implicit representations for $(P_0, Q_0, n_0), (P_1, Q_1, n_1)$ respectively. As will be seen below, an estimate for C_t depends very much on estimates for the partial derivatives $A_{1,t}, B_{0,t}$. Good estimates for these two quantities are not available for all $(P_0, Q_0, n_0), (P_1, Q_1, n_1)$. We will only consider elements satisfying conditions $(*s), (*u)$ below; fortunately, these conditions will always be satisfied whenever we are interested in the variation of the displacements.

Let (P, Q, n) be an element of $\mathcal{R}(I)$, such that $P \subset P_s$. Let I^* be the largest parameter interval such that (P, Q, n) is (the restriction of) an element of $\mathcal{R}(I^*)$. If I^* is not the starting interval $I_0 = [\varepsilon_0, 2\varepsilon_0]$, let \tilde{I}^* be the parent interval and let $(\tilde{P}, \tilde{Q}, \tilde{n})$ be the longest element of $\mathcal{R}(\tilde{I}^*)$ such that $P \subset \tilde{P}$; $(\tilde{P}, \tilde{Q}, \tilde{n})$ is the element which is denoted by (P_0, Q_0, n_0) in the structure theorem of Subsection 6.7, as $[0, \tilde{n}]$ is the maximal initial \tilde{I}^* -interval.

We say that (P, Q, n) satisfies condition $(*s)$ if either $I^* = I_0$ or $I^* \neq I_0$ and \tilde{P} is \tilde{I}^* -critical.

We define in a symmetric way a condition $(*u)$ (when $Q \subset Q_u$).

Proposition 15. *Let (P, Q, n) be an element of $\mathcal{R}(I)$ with $P \subset P_s$. Let (A, B) be the implicit representation of (P, Q, n) . If (P, Q, n) satisfies condition $(*s)$, we have*

$$|A_t| \leq \varepsilon_0^{\frac{1}{2}}.$$

If $(P, Q, n) \in \mathcal{R}(I_0)$, we have the stronger estimate:

$$|A_t| \leq C\varepsilon_0 \log \varepsilon_0^{-1}.$$

Proof. We first show for $(P, Q, n) \in \mathcal{R}(I_0)$ and $P \subset P_s$, that:

$$(7.84) \quad |A_t| \leq C\varepsilon_0 \log \varepsilon_0^{-1}, |B_t| \leq C.$$

When $P = P_s$, the first inequality is part of Corollary 10 and the second is (7.38). If $P \subsetneq P_s$, we set

$$(7.85) \quad (P, Q, n) = (\tilde{P}, \tilde{Q}, \tilde{n}) * (P^*, Q^*, 1).$$

Writing (\tilde{A}, \tilde{B}) and (A^*, B^*) for the implicit representations of $(\tilde{P}, \tilde{Q}, \tilde{n})$ and $(P^*, Q^*, 1)$, respectively, we have, from the formulas in Appendix A:

$$(7.86) \quad A_t = \tilde{A}_t + \tilde{A}_x(A_t^* + A_y^* \tilde{B}_t) \Delta^{-1},$$

$$(7.87) \quad B_t = B_t^* + B_y^*(\tilde{B}_t + \tilde{B}_x A_t^*) \Delta^{-1},$$

with, as usual, $\Delta = 1 - \tilde{B}_x A_y^*$.

This gives, assuming (7.84) for (\tilde{A}, \tilde{B}) ,

$$(7.88) \quad |A_t - \tilde{A}_t| \leq C|P|,$$

$$(7.89) \quad |B_t - \tilde{B}_t B_y \tilde{B}_y^{-1}| \leq C,$$

where we have used $B_y = \tilde{B}_y B_y^* \Delta^{-1}$. Clearly, (7.88) and (7.89) imply, by iteration, (7.84).

We now turn to the case where $I^* \neq I_0$. In this case, we introduce the integer $k \geq 1$ and the elements (P_i, Q_i, n_i) , $0 \leq i \leq k$, of $\mathcal{R}(\tilde{I}^*)$ given by the structure theorem of Subsection 6.7. We also denote by $(P^{(i)}, Q^{(i)}, n^{(i)})$ the element of $\mathcal{R}(\tilde{I}^*)$ such that $P \subset P^{(i)}$ and

$$(7.90) \quad (P^{(i)}, Q^{(i)}, n^{(i)}) \in (P_0, Q_0, n_0) \square \cdots \square (P_i, Q_i, n_i).$$

Our proof will be by induction, on the level of the parameter interval and on the integer k (for a fixed parameter interval).

We first observe that if P is I -critical, then P_0 is \tilde{I}^* -critical and P satisfies $(*s)$. Therefore, if P satisfies $(*s)$, all $P^{(i)}$ also satisfy $(*s)$; moreover, $Q^{(i)}$ satisfy $(*u)$ and P_{i+1} satisfy $(*s)$ by Theorem 1.

Fix $0 \leq i < k$ and let $(A^{(i+1)}, B^{(i+1)})$, $(A^{(i)}, B^{(i)})$ and (A^*, B^*) be the implicit representations of $(P^{(i+1)}, Q^{(i+1)}, n^{(i+1)})$, $(P^{(i)}, Q^{(i)}, n^{(i)})$ and $(P_{i+1}, Q_{i+1}, n_{i+1})$ respectively. From the formulas in

Appendix A, we have

$$(7.91) \quad A_t^{(i+1)} = A_t^{(i)} + A_x^{(i)}(X_t + X_w W_t),$$

$$(7.92) \quad X_t = (X_{u,t} + X_{u,y} B_t^{(i)}) \Delta_0^{-1},$$

$$(7.93) \quad X_w = X_{u,w} \Delta_0^{-1},$$

$$(7.94) \quad W_t = -C_t C_w^{-1}.$$

In these formulas, $X_{u,t}$, $X_{u,y}$, $X_{u,w}$, Δ_0^{-1} are uniformly bounded and $B_t^{(i)}$ is bounded by $\varepsilon_0^{1/2}$ from the induction hypothesis. The term C_w^{-1} is estimated by (7.65) above and C_t is uniformly bounded by the induction hypothesis and Corollary 11 below. We conclude that

$$(7.95) \quad \begin{aligned} |A_t^{(i+1)} - A_t^{(i)}| &\leq C |P^{(i)}| \delta(Q^{(i)}, P_{i+1})^{-\frac{1}{2}} \\ &\leq C |P^{(i)}| |I^*|^{-\frac{1}{2}}, \end{aligned}$$

as $Q^{(i)}$ and P_{i+1} are I^* -transverse. We have for $0 \leq j \leq i$

$$(7.96) \quad |P_j| \leq |\tilde{I}^*|^\beta$$

because (P_j, Q_j, n_j) is \tilde{I}^* -bicritical (we use here, for $j = 0$, the hypothesis that P_0 is \tilde{I}^* -critical).

Then, by Corollary 6, we obtain

$$(7.97) \quad |P^{(i)}| \leq C^i |\tilde{I}^*|^{(i+1)\beta - \frac{i}{2}(1+\tau)}$$

which implies

$$(7.98) \quad |A_t^{(i+1)} - A_t^{(i)}| \leq |\tilde{I}^*|^{\frac{\beta}{2}(1+i)}.$$

Summing (7.98) over i and the levels of parameter intervals leads to the estimate of the proposition. \square

Corollary 11. *Let (P_0, Q_0, n_0) , (P_1, Q_1, n_1) be elements of $\mathcal{R}(I)$ with $Q_0 \subset Q_u$, $P_1 \subset P_s$. Assume that Q_0 satisfies $(*u)$ and that P_1 satisfies $(*s)$. Then, the function C introduced in Subsection 3.5 satisfy*

$$|C_t + 1| \leq C \varepsilon_0^{\frac{1}{2}}.$$

Proof. From formula (A35) in Appendix A, using the notations there, we have

$$(7.99) \quad -C_t = \theta_x \bar{X}_t + \theta_y \bar{Y}_t + \theta_t,$$

$$(7.100) \quad \bar{X}_t = (A_{1,t} + A_{1,y} Y_{s,t}) \Delta_1^{-1},$$

$$(7.101) \quad \bar{Y}_t = (B_{0,t} + B_{0,x} X_{u,t}) \Delta_0^{-1},$$

with Δ_0^{-1} , Δ_1^{-1} uniformly bounded. The value of θ_t is taken at (\bar{X}, \bar{Y}, t) , with

$$(7.102) \quad |\bar{X}| \leq C \varepsilon_0, \quad |\bar{Y}| \leq C \varepsilon_0.$$

On the other hand, we have, in Subsection 4.2, normalized the parameter in order to have

$$(7.103) \quad \theta_t(0, 0, t) \equiv 1.$$

We, therefore, have

$$(7.104) \quad |\theta_t(\bar{X}, \bar{Y}, t) - 1| \leq C\varepsilon_0.$$

In (7.100) and (7.101), we have $|A_{1,y}| < C\varepsilon_0$, $|B_{0,x}| < C\varepsilon_0$, by (R4) and $|A_{1,t}| < \varepsilon_0^{\frac{1}{2}}$, $|B_{0,t}| < \varepsilon_0^{\frac{1}{2}}$ by Proposition 15. The Corollary follows, as θ_x , θ_y , $Y_{s,t}$, $X_{u,t}$ are uniformly bounded. \square

7.7 Variation of Width of Critical Rectangles

Our main purpose now is to prove property (R7) of Subsection 5.4:

(R7) If $(P_0, Q_0, n_0), (P_1, Q_1, n_1) \in \mathcal{R}(I)$ satisfy $Q_0 \subset Q_u$, $P_1 \subset P_s$ and $Q_0 \cap_I P_1$ holds, then, for all $t \in I$, we have

$$\delta(Q_0, P_1) \geq C^{-1} \left(|P_1|^{1-\eta} + |Q_0|^{1-\eta} \right).$$

A priori, the transversality condition gives some control through (T2), (T3) in Subsection 5.4 only for *some* values of the parameter. However, from (T1) in Subsection 5.4 and Corollary 11 above, we know that the order of magnitude of $\delta(Q_0, P_1)$ is the same through out I . Therefore, to obtain (R7), we do need to control how the widths $|P_1|$ and $|Q_0|$ vary through I . Good estimates will be obtained under the same conditions (*s) or (*u) used to obtain Proposition 15. Again, the estimates are even better for an element (P, Q, n) in $\mathcal{R}(I_0)$, involving only simple composition.

Proposition 16. *Let (P, Q, n) be an element of $\mathcal{R}(I_0)$, and let (A, B) be the implicit representation of (P, Q, n) . We have*

$$\begin{aligned} |\partial_t \log |A_x|| &\leq Cn, & |A_{yt}| &\leq C \\ |\partial_t \log |B_y|| &\leq Cn, & |B_{xt}| &\leq C. \end{aligned}$$

Proof. We first observe that we have, by the same proof as for the second inequality in (7.84):

$$(7.105) \quad |A_t| \leq C, \quad |B_t| \leq C.$$

We write

$$(7.106) \quad (P, Q, n) = (\tilde{P}, \tilde{Q}, \tilde{n}) * (P^*, Q^*, 1),$$

and denote by (\tilde{A}, \tilde{B}) , (A^*, B^*) the implicit representations of $(\tilde{P}, \tilde{Q}, \tilde{n})$, $(P^*, Q^*, 1)$, respectively.

By the formulas of Appendix A, we have

$$(7.107) \quad \begin{aligned} \partial_t \log |A_x| &= \partial_t \log |\tilde{A}_x| + \partial_t \log |A_x^*| + X_t \partial_x \log |\tilde{A}_x| \\ &\quad + Y_t \partial_y \log |A_x^*| - \Delta_t \Delta^{-1}, \end{aligned}$$

$$(7.108) \quad X_t = (A_t^* + A_y^* \tilde{B}_t) \Delta^{-1},$$

$$(7.109) \quad Y_t = (\tilde{B}_t + A_t^* \tilde{B}_x) \Delta^{-1},$$

$$(7.110) \quad \Delta = 1 - A_y^* \tilde{B}_x,$$

$$(7.111) \quad -\Delta_t = \tilde{B}_{xt} A_y^* + \tilde{B}_{xx} A_y^* X_t + \tilde{B}_x A_{yt}^* + \tilde{B}_x A_{yy}^* Y_t.$$

As $\tilde{B}_t, \tilde{B}_x, A_t^*, A_y^*, \Delta^{-1}$ are uniformly bounded, the same is true for X_t, Y_t . Using bounded distortion and the cone condition then leads to

$$(7.112) \quad |\partial_t \log |A_x| - \partial_t \log |\tilde{A}_x|| \leq C(1 + |\tilde{B}_{xt}|).$$

Still from Appendix A, we have

$$(7.113) \quad A_{yt} - \tilde{A}_{yt} = X_t \tilde{A}_{xy} + X_y \tilde{A}_{xt} + X_t X_y \tilde{A}_{xx} + \tilde{A}_x X_{yt},$$

$$(7.114) \quad X_y = A_y^* \tilde{B}_y \Delta^{-1},$$

$$(7.115) \quad \begin{aligned} X_{yt} &= \tilde{B}_y \Delta^{-1} (A_{yy}^* Y_t + A_y^* \partial_t \log |\tilde{B}_y|) + A_{yt}^* \\ &\quad + A_y^* X_t \partial_x \log |\tilde{B}_y| - A_y^* \Delta_t \Delta^{-1}. \end{aligned}$$

As X_t, X_y, Y_t are bounded and also using bounded distortion, we have

$$(7.116) \quad |X_t \tilde{A}_{xy}| \leq C |\tilde{A}_x|,$$

$$(7.117) \quad |X_y \tilde{A}_{xt}| \leq C |\tilde{A}_x| |\partial_t \log |\tilde{A}_x||,$$

$$(7.118) \quad |X_t X_y \tilde{A}_{xx}| \leq C |\tilde{A}_x|,$$

$$(7.119) \quad |X_{yt}| \leq C |\tilde{B}_y| (1 + |\partial_t \log |\tilde{B}_y|| + |\tilde{B}_{xt}|).$$

We, therefore, obtain

$$(7.120) \quad |A_{yt} - \tilde{A}_{yt}| \leq C |\tilde{A}_x| (1 + |\partial_t \log |A_x||) + C |\tilde{A}_x| |\tilde{B}_y| (|\partial_t \log |\tilde{B}_y|| + |\tilde{B}_{xt}|).$$

We have symmetric estimates for B_{xt} and $\partial_t \log |B_y|$, writing now

$$(7.121) \quad (P, Q, n) = (\hat{P}^*, \hat{Q}^*, 1) * (\hat{P}, \hat{Q}, n - 1).$$

As $|\tilde{A}_x|, |\tilde{B}_y|$ are exponentially small, the estimates (7.112), (7.120) and the other two estimates for $B_{xt}, \partial_t \log |B_y|$ lead by summation to the bounds of the proposition. \square

For the last estimate of this section, the setting is the same as that of Proposition 15.

Proposition 17. *Let (P, Q, n) be an element of $\mathcal{R}(I)$ with $P \subset P_s$. Let (A, B) be the implicit representation of (P, Q, n) . If (P, Q, n) satisfies condition $(*)$, we have*

$$(7.122) \quad |\partial_t \log |A_x|| \leq C_0 \frac{\log |P|}{|I| \log |I|},$$

$$(7.123) \quad |A_{yt}| \leq C_0.$$

Proof. The method is the same as in Proposition 15, but, as now we have to deal with second instead of first order partial derivatives, calculations are more complicated.

When $(P, Q, n) \in \mathcal{R}(I_0)$, n and $|\log |P||$ are of the same order; as $|I| |\log |I||$ is always smaller than $\varepsilon_0 \log \varepsilon_0^{-1}$, the estimates in Proposition 16 are stronger in the present case than the ones in (7.122), (7.123).

We will now assume that $(P, Q, n) \notin \mathcal{R}(I_0)$; the proof of (7.122), (7.123) is by induction on the level of the parameter interval. Let $I^* \neq I_0$ be the largest parameter interval such that $(P, Q, n) \in \mathcal{R}(I^*)$; as $|I| |\log |I|| \leq |I^*| |\log |I^*||$, (7.122) for $\mathcal{R}(I)$ follows from (7.122) for $\mathcal{R}(I^*)$. We can therefore assume that $I^* = I$ and denote by \tilde{I} the parent interval.

We apply the structure theorem of Subsection 6.7 and use the same notations as in the proof of Proposition 15: we have an integer k , elements (P_i, Q_i, n_i) in $\mathcal{R}(\tilde{I})$ for $0 \leq i \leq k$ and partial compositions of $(P^{(i)}, Q^{(i)}, n^{(i)})$ in $\mathcal{R}(I)$ for $0 \leq i \leq k$. We denote by (A, B) , (\tilde{A}, \tilde{B}) , (A^*, B^*) the implicit representations of $(P^{(i+1)}, Q^{(i+1)}, n^{(i+1)})$, $(P^{(i)}, Q^{(i)}, n^{(i)})$, $(P_{i+1}, Q_{i+1}, n_{i+1})$, respectively, for some fixed integer $0 \leq i < k$. We know that both $(P^{(i)}, Q^{(i)}, n^{(i)})$, $(P_{i+1}, Q_{i+1}, n_{i+1})$ satisfy $(*)$, and $(P^{(i)}, Q^{(i)}, n^{(i)})$ also satisfies $(*u)$. We have from formulas (A47), (A49) in Appendix A:

$$(7.124) \quad \begin{aligned} \partial_t \log |A_x| - \partial_t \log |\tilde{A}_x| &= \partial_x \log |\tilde{A}_x| (X_t + X_w W_t) + \partial_t \log |X_w| + \\ &+ W_t \partial_w \log |X_w| + \partial_t \log |W_x|, \end{aligned}$$

$$(7.125) \quad \begin{aligned} A_{yt} - \tilde{A}_{yt} &= \tilde{A}_{xy} (X_t + X_w W_t) + \tilde{A}_{xt} (X_y + X_w W_y) + \\ &+ \tilde{A}_{xx} (X_t + X_w W_t) (X_y + X_w W_y) + \\ &+ \tilde{A}_x (X_{yt} + X_{wy} W_t + X_{wt} W_y + X_{ww} W_y W_t + X_w W_{yt}). \end{aligned}$$

We need to estimate all terms in the right-hand sides of these formulas.

Terms involving \tilde{A} . By bounded distortion, $|\partial_x \log |\tilde{A}_x||$ is bounded, \tilde{A}_x , \tilde{A}_{xy} , \tilde{A}_{xx} are bounded by $C|\tilde{A}_x|$ and \tilde{A}_{xt} is part of the induction

$$(7.126) \quad |\tilde{A}_{xt}| = |\tilde{A}_x| |\partial_t \log |\tilde{A}_x||.$$

Terms involving the first order partial derivatives of X . We have dealt earlier with X_t , X_w , X_y (cf. (7.92), (7.93), (7.60) and these formulas easily give

$$(7.127) \quad |X_t| \leq C,$$

$$(7.128) \quad C^{-1} \leq |X_w| \leq C,$$

$$(7.129) \quad |X_y| \leq C|\tilde{B}_y|.$$

Terms involving the first order partial derivatives of W . We have dealt earlier with $W_y = -C_y C_w^{-1}$ and $W_t = -C_t C_w^{-1}$; the term C_w^{-1} is estimated by (7.65), the term C_y from formula (7.64) and the term C_t from Corollary 11. One obtains

$$(7.130) \quad C^{-1} \delta(Q^{(i)}, P_{i+1})^{-\frac{1}{2}} |\tilde{B}_y| \leq |W_y| \leq C \delta(Q^{(i)}, P_{i+1})^{-\frac{1}{2}} |\tilde{B}_y|,$$

$$(7.131) \quad C^{-1} \delta(Q^{(i)}, P_{i+1})^{-\frac{1}{2}} \leq |W_t| \leq C \delta(Q^{(i)}, P_{i+1})^{-\frac{1}{2}}.$$

Terms involving the second order partial derivatives of X . The formulas (A63) of Appendix A express the second order partial derivatives of X in terms of partial derivatives of first and second order of X_u (which are bounded), partial derivatives of second order of \tilde{B} (which are controlled by the distortion or by the induction hypothesis), partial derivatives of first order of X itself (see above), the bounded quantity Δ_0^{-1} , and partial derivatives of first order of \bar{Y} (defined in (A31)). These partial derivatives given by (A33) are easy to estimate: we have $\bar{Y}_w = \tilde{B}_x X_w$, $\bar{Y}_y = \tilde{B}_y \Delta_0^{-1}$, $\bar{Y}_t = (\tilde{B}_t + \tilde{B}_x X_{u,t}) \Delta_0^{-1}$, hence

$$(7.132) \quad |\bar{Y}_w| \leq C,$$

$$(7.133) \quad |\bar{Y}_y| \leq C|\tilde{B}_y|,$$

$$(7.134) \quad |\bar{Y}_t| \leq C,$$

where we have used Proposition 15 to bound \tilde{B}_t . Plugging these estimates in the formulas (A63) gives:

$$(7.135) \quad |X_{ww}| \leq C,$$

$$(7.136) \quad |X_{wy}| \leq C|\tilde{B}_y|,$$

$$(7.137) \quad |X_{yy}| \leq C|\tilde{B}_y|,$$

$$(7.138) \quad |X_{wt}| \leq C(1 + |\tilde{B}_{xt}|),$$

$$(7.139) \quad |X_{yt}| \leq C|\tilde{B}_y|(1 + |\tilde{B}_{xt}| + \partial_t \log |\tilde{B}_y|).$$

Terms involving the second order partial derivatives of W . The formulas (A55) of Appendix A express the second order partial derivatives of W in terms of C_w^{-1} (controlled by (7.65)),

the partial derivatives of first order of W (see above) and the partial derivatives of second order of C . These partial derivatives of second order of C are in turn expressed in formulas (A56)–(A60) in terms of partial derivatives of θ (which are bounded) and partial derivatives of first and second order of \bar{X} and \bar{Y} . The partial derivatives of first order of \bar{Y} have been estimated above, those of \bar{X} satisfy in the same way the inequalities

$$(7.140) \quad |\bar{X}_w| \leq C,$$

$$(7.141) \quad |\bar{X}_x| \leq C|\tilde{A}_x^*|,$$

$$(7.142) \quad |\bar{X}_t| \leq C.$$

The partial derivatives of second order of \bar{X} and \bar{Y} are expressed in formulas (A61), (A62). The formulas (A61) contain partial derivatives of first and second order of Y , which are firstly estimated in the same way, through formulas (A29), (A64), as those of X :

$$(7.143) \quad |Y_t| \leq C,$$

$$(7.144) \quad C^{-1} \leq |Y_w| \leq C,$$

$$(7.145) \quad |Y_x| \leq C|A_x^*|,$$

$$(7.146) \quad |Y_{ww}| \leq C,$$

$$(7.147) \quad |Y_{wx}| \leq C|A_x^*|,$$

$$(7.148) \quad |Y_{xx}| \leq C|A_x^*|,$$

$$(7.149) \quad |Y_{wt}| \leq C(1 + |A_{yt}^*|),$$

$$(7.150) \quad |Y_{xt}| \leq C|A_x^*|(1 + |A_{yt}^*| + |\partial_t \log |A_x^*||).$$

We can then estimate the second order partial derivatives of \bar{X} and \bar{Y} :

$$(7.151) \quad |\bar{X}_{ww}| \leq C,$$

$$(7.152) \quad |\bar{X}_{wx}| \leq C|A_x^*|,$$

$$(7.153) \quad |\bar{X}_{wt}| \leq C(1 + |A_{yt}^*|),$$

$$(7.154) \quad |\bar{X}_{xx}| \leq C|A_x^*|,$$

$$(7.155) \quad |\bar{X}_{xt}| \leq C|A_x^*|(1 + |A_{yt}^*| + |\partial_t \log |A_x^*||),$$

$$(7.156) \quad |\bar{Y}_{ww}| \leq C,$$

$$(7.157) \quad |\bar{Y}_{wy}| \leq C|\tilde{B}_y|,$$

$$(7.158) \quad |\bar{Y}_{wt}| \leq C(1 + |\tilde{B}_{xt}^*|),$$

$$(7.159) \quad |\bar{Y}_{yy}| \leq C|\tilde{B}_y|,$$

$$(7.160) \quad |\bar{Y}_{yt}| \leq C|\tilde{B}_y|(1 + |\tilde{B}_{xt}| + |\partial_t \log |\tilde{B}_y||).$$

The next step is to estimate the partial derivatives of second order of C (besides C_{ww} which is

already known to the close to 2)

$$(7.161) \quad |C_{wx}| \leq C|A_x^*|,$$

$$(7.162) \quad |C_{wy}| \leq C|\tilde{B}_y|,$$

$$(7.163) \quad |C_{wt}| \leq C(1 + |\tilde{B}_{xt}^*| + |A_{yt}^*|),$$

$$(7.164) \quad |C_{xx}| \leq C|A_x^*|,$$

$$(7.165) \quad |C_{xy}| \leq C|A_x^*||\tilde{B}_y|,$$

$$(7.166) \quad |C_{yy}| \leq C|\tilde{B}_y|,$$

$$(7.167) \quad |C_{xt}| \leq C|A_x^*|(1 + |A_{yt}^*| + |\partial_t \log |A_x^*||),$$

$$(7.168) \quad |C_{yt}| \leq C|\tilde{B}_y|(1 + |\tilde{B}_{xt}| + |\partial_t \log |\tilde{B}_y||).$$

Finally, we are able to estimate the partial derivatives of second order of W :

$$(7.169) \quad |W_{xt}| \leq C\delta^{-\frac{1}{2}}|A_x^*|(\delta^{-1} + \delta^{-\frac{1}{2}}(|\tilde{B}_{xt}| + |A_{yt}^*|) + |\partial_t \log |A_x^*||),$$

$$(7.170) \quad |W_{yt}| \leq C\delta^{-\frac{1}{2}}|\tilde{B}_y|(\delta^{-1} + \delta^{-\frac{1}{2}}(|\tilde{B}_{xt}| + |A_{yt}^*|) + |\partial_t \log |\tilde{B}_y||).$$

We have written δ for $\delta(Q^{(i)}, P_{i+1})$.

We are now ready to come back to formulas (7.124), (7.125) above. We get:

$$(7.171) \quad |\partial_t \log |A_x| - \partial_t \log |\tilde{A}_x|| \leq C(\delta^{-1} + \delta^{-\frac{1}{2}}(|\tilde{B}_{xt}| + |A_{yt}^*|) + |\partial_t \log |A_x^*||),$$

$$(7.172) \quad |A_{yt} - \tilde{A}_{yt}| \leq C\delta^{-\frac{1}{2}}|\tilde{A}_x|(1 + |\tilde{B}_y|K),$$

with

$$(7.173) \quad K = \delta^{-1} + \delta^{-\frac{1}{2}}(|\tilde{B}_{xt}| + |A_{yt}^*|) + |\partial_t \log |\tilde{A}_x| + |\partial_t \log |\tilde{B}_y||.$$

By the induction hypothesis, we have

$$(7.174) \quad |\tilde{B}_{xt}| \leq C_0,$$

$$(7.175) \quad |A_{yt}^*| \leq C_0,$$

$$(7.176) \quad |\partial_t \log |A_x^*|| \leq C_0 \frac{\log |A_x^*|}{|\tilde{I}| \log |\tilde{I}|},$$

$$(7.177) \quad |\partial_t \log |\tilde{A}_x|| \leq C_0 \frac{\log |\tilde{A}_x|}{|I| \log |I|},$$

$$(7.178) \quad |\partial_t \log |\tilde{B}_y|| \leq C_0 \frac{\log |\tilde{B}_y|}{|I| \log |I|}.$$

Here C_0 is large but independent of ε_0 . This means that the term $\delta^{-\frac{1}{2}}(|\tilde{B}_{xt}| + |A_{yt}^*|)$ in (7.171) and (7.173) is dominated by δ^{-1} . As $|I| = |\tilde{I}|^{1+\tau}$, in order to prove (7.122) by induction, we need to have, in view of (7.171):

$$(7.179) \quad C|I| |\log |I|| \delta^{-1} + C C_0 |\tilde{I}|^\tau \log |A_x^*| + C_0 |\log |\tilde{A}_x|| \leq C_0 |\log |A_x||.$$

We have here $\delta \geq 2|I|$ from the definition of the transversality relation and, by (3.27):

$$(7.180) \quad |\log |A_x|| \geq |\log |\tilde{A}_x|| + |\log |A_x^*|| - \frac{1}{2} |\log |I||.$$

Therefore, (7.179) will hold as far as

$$(7.181) \quad \left(\frac{C_0}{2} + C\right) |\log |I|| \leq C_0(1 - C|\tilde{I}|^\tau) |\log |A_x^*||.$$

From (R7), we know that $|A_x^*|$ is much smaller than δ . On the other hand, as Q_i and P_{i+1} are not \tilde{I} -transverse, δ cannot be much larger than \tilde{I} . Therefore, we must have

$$(7.182) \quad |\log |A_x^*|| \geq \log |\tilde{I}| = (1 + \tau) \log |I|,$$

from which (7.181) follows if we take $C_0 \geq 3C$. This completes the proof of the induction step for (7.122).

To do the same for (7.123), we estimate the right-hand side of (7.172). From the proof of Proposition 15, formula (7.97), we have

$$(7.183) \quad |\tilde{A}_x| \leq C^i |\tilde{I}|^{(i+1)\beta - \frac{i}{2}(1+\tau)}$$

From (R7), we have

$$(7.184) \quad |\tilde{B}_y| \leq C \delta^{(1-\eta)^{-1}}.$$

The displacement $\delta = \delta(Q^{(i)}, P_{i+1})$ satisfies

$$(7.185) \quad 2|I| \leq \delta \leq C|\tilde{I}|.$$

This gives (as $\beta > 1$)

$$(7.186) \quad \delta^{-\frac{1}{2}} |\tilde{A}_x| \leq |\tilde{I}|^{\beta/2(i+1)},$$

$$(7.187) \quad \delta^{-\frac{3}{2}} |\tilde{A}_x| |\tilde{B}_y| \leq |\tilde{I}|^{\beta/2(i+1)},$$

$$(7.188) \quad \delta^{-\frac{1}{2}} |\tilde{A}_x| |\tilde{B}_y| |\partial_t \log |\tilde{A}_x|| \leq |\tilde{I}|^{\beta/2(i+1)},$$

$$(7.189) \quad \delta^{-\frac{1}{2}} |\tilde{A}_x| |\tilde{B}_y| |\partial_t \log |\tilde{B}_y|| \leq |\tilde{I}|^{\beta/2(i+1)}.$$

This leads to:

$$(7.190) \quad |A_{yt} - \tilde{A}_{yt}| \leq C |\tilde{I}|^{\beta/2(i+1)}.$$

We can now sum over i and then over the different levels of parameter intervals to obtain (7.123). The proof of Proposition 17 is complete. \square

Let us see that property (R7) of Subsection 5.4 is a consequence of Proposition 17. Consider (P_0, Q_0, n_0) , (P_1, Q_1, n_1) to be elements of $\mathcal{R}(I)$ such that $Q_0 \subset Q_u$, $P_1 \subset P_s$ and $Q_0 \pitchfork_I P_1$ holds.

We can assume that there are no $\widehat{I} \supset I$, $(\widehat{P}_0, \widehat{Q}_0, \widehat{n}_0), (\widehat{P}_1, \widehat{Q}_1, \widehat{n}_1) \in \mathcal{R}(\widehat{I})$ with $Q_0 \subset \widehat{Q}_0$, $P_1 \subset \widehat{P}_1$, $(Q_0, P_1, I) \neq (\widehat{Q}_0, \widehat{P}_1, \widehat{I})$ and $\widehat{Q}_0 \cap_{\widehat{I}} \widehat{P}_1$: otherwise, as $|Q_0| \leq |\widehat{Q}_0|$, $|P_1| \leq |\widehat{P}_1|$ and $\delta(Q_0, P_1) \geq \delta(\widehat{Q}_0, \widehat{P}_1)$, property (R7) for (Q_0, P_1, I) would be inherited from $(\widehat{Q}_0, \widehat{P}_1, \widehat{I})$.

In view of this maximality property, we claim that Q_0 must satisfy condition $(*u)$, and P_1 must satisfy condition $(*s)$.

Let us first finish the proof of (R7) assuming the claim to be true. Indeed, we have $\delta(Q_0, P_1) \geq 2|I|$ for all $t \in I$ by (T1) of Subsection 5.4. If $|\log |Q_0||$ is much larger than $|\log |I||$ for all $t \in I$, we obviously have $\delta(Q_0, P_1) \geq |Q_0|^2$; but if $|\log |Q_0|| \leq C|\log |I||$ for some $t \in I$, we obtain from Proposition 17 that

$$(7.191) \quad \max_I |Q_0| \leq C \min_I |Q_0|.$$

We also know from Proposition 15 that

$$(7.192) \quad \max_I \delta(Q_0, P_1) \leq C \min_I \delta(Q_0, P_1).$$

It then follows from (T2) in Subsection 5.4 that

$$(7.193) \quad \delta(Q_0, P_1) \geq C^{-1}|Q_0|^{1-\eta}.$$

for all $t \in I$. We argue with P_1 in a symmetric way.

Finally, we prove the claim. Let us show, for instance, that P_1 satisfies condition $(*s)$. Let I^* be the largest parameter interval such that $(P_1, Q_1, n_1) \in \mathcal{R}(I^*)$. If I^* is the starting interval I_0 , P_1 satisfies condition $(*s)$. Assume therefore that $I^* \neq I_0$; let \widetilde{I}^* be the parent interval, $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1)$ the element of $\mathcal{R}(\widetilde{I}^*)$ such that \widetilde{P}_1 is the thinnest rectangle containing P_1 . We have to show that \widetilde{P}_1 is \widetilde{I}^* -critical. Assume by contradiction that \widetilde{P}_1 is \widetilde{I}^* -transverse. Then, there exists an \widetilde{I}^* -decomposition $(P_\alpha, Q_\alpha, n_\alpha)$ of Q_u such that, for every α , Q_α and \widetilde{P}_1 are \widetilde{I}^* -separated or \widetilde{I}^* -transverse. Let α_0 be such that Q_{α_0} and Q_0 intersect. Then, Q_{α_0} and \widetilde{P}_1 must be \widetilde{I}^* -transverse. But then \widetilde{P}_1 and Q_0 must be I -transverse, either from Proposition 2 if $Q_0 \subset Q_{\alpha_0}$, or from concavity (Proposition 4) if $Q_{\alpha_0} \subset Q_0$; this contradicts the maximality of (Q_0, P_1, I) and proves the claim.

The proof of property (R7) is complete.

The existence and properties of the classes $\mathcal{R}(I)$ are now fully justified. What we do not know at this moment is whether there exists any regular parameter interval at all! This will be the subject of Section 9. Before, we develop in the next section some results that will turn out to be essential in Sections 9 and 10.

8 Number of Children and Dimension Estimates

8.1 Estimates on the Number of Children

We start with some preliminary results.

Proposition 18. *Let $I \subset \tilde{I}$ be parameter intervals, and let $(\tilde{P}, \tilde{Q}, \tilde{n})$ be an element of $\mathcal{R}(\tilde{I})$. We assume that \tilde{Q} is \tilde{I} -transverse. Then, any element (P, Q, n) in $\mathcal{R}(I)$ such that P is a child of \tilde{P} is already an element of $\mathcal{R}(\tilde{I})$.*

Proof. We can assume that P is a non-simple child. Then (P, Q, n) is obtained by parabolic composition in $\mathcal{R}(I)$ of $(\tilde{P}, \tilde{Q}, \tilde{n})$ with some $(P_1, Q_1, n_1) \in \mathcal{R}(I)$. As \tilde{Q} is \tilde{I} -transverse, there exists an \tilde{I} -decomposition $(P_\alpha, Q_\alpha, n_\alpha)$ of P_s such that each P_α is \tilde{I} -separated or \tilde{I} -transverse with \tilde{Q} . Let α_0 be such that P_{α_0} and P_1 intersect. Then, $\tilde{Q} \pitchfork_{\tilde{I}} P_{\alpha_0}$ holds, and also $\tilde{Q} \pitchfork_I P_1$; if we had $P_1 \not\subseteq P_{\alpha_0}$, this would imply that \tilde{Q} would be I -transverse to the parent \tilde{P}_1 of P_1 and P would not be a child of \tilde{P} . Therefore, we must have $P_{\alpha_0} \subset P_1$. By coherence (Proposition 6), we have that $(P_1, Q_1, n_1) \in \mathcal{R}(\tilde{I})$. By parametric concavity (Proposition 7), from $\tilde{Q} \pitchfork_I P_1$ and $\tilde{Q} \pitchfork_{\tilde{I}} P_{\alpha_0}$, we deduce that $\tilde{Q} \pitchfork_{\tilde{I}} P_1$ also holds and $(P, Q, n) \in \mathcal{R}(\tilde{I})$. \square

Proposition 19. *Let I be a parameter interval, and let $I_1 \supset I$ be the largest parameter interval such that*

$$(8.1) \quad |I_1|^\beta < \left(\frac{1}{2} |I|\right)^{\frac{1}{1-\eta}}.$$

Let $(\tilde{P}, \tilde{Q}, \tilde{n}), (P, Q, n)$ be elements of $\mathcal{R}(I)$ such that P is a non-simple child of \tilde{P} . Let $(P_1, Q_1, n_1), (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ be the elements of $\mathcal{R}(I)$ such that

$$(P, Q, n) \in (\tilde{P}, \tilde{Q}, \tilde{n}) \square (P_1, Q_1, n_1)$$

and \tilde{P}_1 is the parent of P_1 .

Then, (P_1, Q_1, n_1) belongs to $\mathcal{R}(I_1)$, \tilde{P}_1 is I -critical, \tilde{Q}_1 is I_1 -transverse and we have

$$(8.2) \quad 2|\tilde{P}_1|^{1-\eta} > |I|$$

for all $t \in I$.

Remark. *As parabolic composition is possible, we have $I \neq I_0$; then, as $\beta > (1 - \eta)^{-1}$, we must have $I_1 \supsetneq I$ and I_1 is β -regular.*

Proof. That \tilde{P}_1 is I -critical has already been proved in Proposition 5. Also, as $\tilde{Q} \pitchfork_I P_1$ holds but $\tilde{Q} \pitchfork_{\tilde{I}} \tilde{P}_1$ does not hold (because P is a non-simple child of \tilde{P}), we deduce (8.2) from Proposition 9. Then, by definition of I_1 , we have:

$$(8.3) \quad |\tilde{P}_1| > |I_1|^\beta$$

for all $t \in I$. Let us show that $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ belongs to $\mathcal{R}(I_1)$. Otherwise, there would exist $I_2 \supset I$, with parent interval $\tilde{I}_2 \subset I_1$, such that $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ belongs to $\mathcal{R}(I_2)$ but not to $\mathcal{R}(\tilde{I}_2)$. We apply the inequality (6.74) in the proof of Corollary 9 (Subsection 6.7) to get

$$(8.4) \quad |\tilde{P}_1| \leq C |\tilde{I}_2|^{\beta + \frac{1}{3}},$$

in contradiction with (8.3). Therefore, $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ belongs to $\mathcal{R}(I_1)$. As $I_1 \supsetneq I$, I_1 is β -regular ; $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ cannot be I_1 -bicritical in view of (8.3) ; \tilde{P}_1 is I_1 -critical and hence \tilde{Q}_1 is I_1 -transverse. Proposition 18 then shows that $(P_1, Q_1, n_1) \in \mathcal{R}(I_1)$. \square

Corollary 12. *Let I be a parameter interval and let $(\tilde{P}, \tilde{Q}, \tilde{n})$ be an element of $\mathcal{R}(I)$. The number of $(P, Q, n) \in \mathcal{R}(I)$ such that P is a child of \tilde{P} is finite.*

Proof. We argue by induction on the level of the parameter interval.

If I is the starting interval I_0 , \tilde{P} has only simple children and the assertion is obvious. Assume that $I \subsetneq I_0$. The number of simple children is finite, and we have to show that the same is true for the number of non-simple children. For every non-simple child P of \tilde{P} , let I_1 , (P_1, Q_1, n_1) , $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) \in \mathcal{R}(I_1)$ be as in Proposition 19. By the induction hypothesis, there is for each fixed \tilde{P}_1 only a finite number of possibilities for P_1 . On the other hand, in view of relation (8.2), there are obviously only a finite number of possibilities for \tilde{P}_1 . The induction step is complete, and this completes the proof. \square

We want to make the finiteness assertion quantitative, and will do that in two distinct ways. In each case, we have to estimate in the proof of Corollary 12 the number of possibilities for \tilde{P}_1 , and the number of possibilities for P_1 once \tilde{P}_1 is fixed.

Proposition 20. *Let I be a parameter interval, and let $(\tilde{P}, \tilde{Q}, \tilde{n})$ be an element of $\mathcal{R}(I)$. The number of $(P, Q, n) \in \mathcal{R}(I)$ such that P is a child of \tilde{P} is at most $|I|^{-c\eta}$, where c is a constant depending only on β .*

Proof. We argue again by induction on the level of I , following the proof of Corollary 12. When $I = I_0$, the number of (simple) children is at most the number of rectangles in the Markov partition, which is much smaller than $\varepsilon_0^{-c\eta}$ when ε_0 is small enough.

When $I \neq I_0$, the number of possibilities for P_1 when \tilde{P}_1 is fixed is at most $|I_1|^{-c\eta}$ by the induction hypothesis. We have to estimate the number of possibilities for \tilde{P}_1 . We know that $\tilde{Q} \pitchfork_I \tilde{P}_1$ does not hold, but $\tilde{Q} \pitchfork_I P_1$ holds.

As \tilde{P}_1 is I -critical, it satisfies condition $(*s)$, defined just before Proposition 15 in Subsection 7.5. We have from (8.2) and Proposition 17

$$(8.5) \quad \max_I |\tilde{P}_1| \leq C \min_I |\tilde{P}_1|.$$

Lemma 3. *We have for all $t \in I$*

$$(8.6) \quad \delta(\tilde{Q}, \tilde{P}_1) \leq C |\tilde{P}_1|^{1-\eta}.$$

Proof. Let $(\hat{P}, \hat{Q}, \hat{n})$ be the element of $\mathcal{R}(I)$ with smallest \hat{n} such that $\tilde{Q} \subset \hat{Q}$ and $\hat{Q} \cap_I P_1$ holds. The parent Q^* of \hat{Q} is I -critical: otherwise, from an I -decomposition of P_s , one would find $(P_\alpha, Q_\alpha, n_\alpha)$ with $Q^* \cap_I P_\alpha$ and $P_\alpha \cap P_1 \neq \emptyset$; one would conclude that $Q^* \cap_I P_1$ from Proposition 2 (if $P_1 \subset P_\alpha$) or Proposition 4 (if $P_\alpha \subset P_1$). Thus, \hat{Q} satisfies condition $(*u)$. From Corollary 11 and Proposition 17, we get

$$(8.7) \quad \max_I |\hat{Q}| \leq C \min_I |\hat{Q}|,$$

$$(8.8) \quad \max_I \delta_{LR}(\hat{Q}, \tilde{P}_1) - \min_I \delta_{LR}(\hat{Q}, \tilde{P}_1) \leq 2|I|.$$

As $\tilde{Q} \cap_I \tilde{P}_1$ does not hold, $\hat{Q} \cap_I \tilde{P}_1$ does not hold either and at least one of the following three inequalities must hold:

$$(8.9) \quad \delta_{LR}(\hat{Q}, \tilde{P}_1) < 2|I| \text{ for some } t_0 \in I;$$

$$(8.10) \quad \delta_R(\hat{Q}, \tilde{P}_1) < 2|\hat{Q}|^{1-\eta} \text{ for all } t \in I;$$

$$(8.11) \quad \delta_L(\hat{Q}, \tilde{P}_1) < 2|\tilde{P}_1|^{1-\eta} \text{ for all } t \in I.$$

By Proposition 10, as $\hat{Q} \cap_I P_1$ holds but $\hat{Q} \cap_I \tilde{P}_1$ does not hold, we have, for some $t_1 \in I$:

$$(8.12) \quad |\tilde{P}_1| > \frac{1}{2} |\hat{Q}|.$$

We can now prove (8.6).

If (8.11) holds, we have, for all $t \in I$:

$$(8.13) \quad \delta(\tilde{Q}, \tilde{P}_1) \leq \delta_L(\tilde{Q}, \tilde{P}_1) \leq \delta_L(\hat{Q}, \tilde{P}_1) < 2|\tilde{P}_1|^{1-\eta}.$$

If (8.12) holds, we have from (5.12), (8.5), (8.7), (8.12):

$$(8.14) \quad \begin{aligned} \delta(\tilde{Q}, \tilde{P}_1) &\leq \delta_{LR}(\hat{Q}, \tilde{P}_1) \\ &\leq \delta_R(\hat{Q}, \tilde{P}_1) + c|\hat{Q}| \\ &\leq 3|\hat{Q}|^{1-\eta} \leq C|\tilde{P}_1|^{1-\eta}, \end{aligned}$$

for all $t \in I$.

Finally, if (8.9) holds, we have from (8.5), (8.8), (8.2), for all $t \in I$:

$$(8.15) \quad \begin{aligned} \delta(\tilde{Q}, \tilde{P}_1) &\leq \delta_{LR}(\hat{Q}, \tilde{P}_1) \\ &\leq 4|I| \leq 8|\tilde{P}_1|^{1-\eta}. \end{aligned}$$

□

We are now able to estimate the number of possibilities for \tilde{P}_1 and show that this number is at most

$$(8.16) \quad C |I|^{-\frac{\eta}{1-\eta}}.$$

This indeed follows from (8.6), (8.2) and the fact that if two distinct \tilde{P}_1 are not disjoint, the ratio of their widths is bounded away from 1 (so the \tilde{P}_1 at a given scale are disjoint; one then sums over scales). The total number of children is thus bounded by

$$(8.17) \quad C + 2C |I|^{-\frac{\eta}{1-\eta}} |I_1|^{-c\eta},$$

where I_1 was the largest parameter interval satisfying

$$(8.18) \quad |I_1|^\beta < \left(\frac{1}{2} |I|\right)^{\frac{1}{1-\eta}}.$$

If $|I| > 2\varepsilon_0^{\beta(1-\eta)}$, we have $I_1 = I_0$; in this case, the term $|I_1|^{-c\eta}$ in (8.17) is unnecessary because \tilde{P}_1 has only simple children. If $|I| \leq 2\varepsilon_0^{\beta(1-\eta)}$, we have

$$(8.19) \quad |I_1|^{\beta(1+\tau)^{-1}} \geq \left(\frac{1}{2} |I|\right)^{\frac{1}{1-\eta}}$$

and the term in (8.17) is bounded by $|I|^{-c\eta}$ provided that

$$(8.20) \quad c\eta > \frac{\eta}{1-\eta} + c\eta \frac{1+\tau}{1-\eta} \beta^{-1}.$$

As η, τ are very small, any choice of $c > \frac{\beta}{\beta-1}$ yields (8.20). Then, as $c > 1$, such a choice is also convenient when $|I| > 2\varepsilon_0^{\beta(1-\eta)}$, and this concludes the proof of Lemma 3. \square

In Proposition 20, we have estimated the total number of children in terms of the level of the parameter interval.

When \tilde{Q} is I -transverse, Proposition 18 guarantees that there will not be any new child of \tilde{Q} when we consider parameter intervals $\hat{I} \subset I$ of higher level.

The same is true when for all $t \in I$

$$(8.21) \quad |\tilde{Q}|^{1-\eta} \geq C|I|.$$

with some large enough constant C . Indeed, let I^* be a parameter interval strictly smaller than I , with parent interval \tilde{I}^* , and let (P, Q, n) be an element of $\mathcal{R}(I^*)$ such that P is a non-simple child of \tilde{P} . Let $(P_1, Q_1, n_1), (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ be as in the previous propositions. By Proposition 19, the element (P_1, Q_1, n_1) belongs to $\mathcal{R}(\tilde{I}^*)$. It is then easy to deduce from $\tilde{Q} \pitchfork_{I^*} P_1$ and (8.21), using (R7) and $\tilde{I}^* \subset I$, that $\tilde{Q} \pitchfork_{\tilde{I}^*} P_1$ also holds. This proves by induction that (P, Q, n) belongs to $\mathcal{R}(I)$.

In the next proposition, we are interested, not in the total number of children, but in the number of children of a given width. The estimate is independent on the level of the parameter interval.

Proposition 21. *Let I be a parameter interval, and let $(\tilde{P}, \tilde{Q}, \tilde{n})$ be an element of $\mathcal{R}(I)$. For any $\varepsilon > 0$, the number of elements $(P, Q, n) \in \mathcal{R}(I)$ such that P is a non-simple child of \tilde{P} satisfying $|P| \geq \varepsilon |\tilde{P}|$ for some $t \in I$, is at most $\varepsilon^{-c'}$, where c' is a constant depending only on the regularity parameter β .*

Proof. Let $\varepsilon > 0$, and let (P, Q, n) be an element of $\mathcal{R}(I)$ such that P is a non-simple child of \tilde{P} . We assume, for some $t_0 \in I$, that:

$$(8.22) \quad |P| \geq \varepsilon |\tilde{P}|.$$

Let $(P_1, Q_1, n_1), (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) \in \mathcal{R}(I)$ be as in Proposition 19. From (3.27), we have, for all $t \in I$:

$$(8.23) \quad |P| \leq C |\tilde{P}| |P_1| \delta(\tilde{Q}, P_1)^{-\frac{1}{2}}.$$

Property (R7) guarantees that, for all $t \in I$

$$(8.24) \quad \delta(\tilde{Q}, P_1) \geq C^{-1} |P_1|^{1-\eta}.$$

Combining (8.22), (8.23), (8.24), we have, for some $t_0 \in I$

$$(8.25) \quad \delta(\tilde{Q}, P_1) \geq C^{-1} \varepsilon^2 \frac{1-\eta}{1+\eta}.$$

As we always have

$$(8.26) \quad \delta(\tilde{Q}, P_1) \leq C \varepsilon_0,$$

there is no non-simple child satisfying (8.22) unless $\varepsilon < \varepsilon_0^{\frac{1}{2}}$; we will assume that this holds in the sequel.

From Lemma 3 above, we have, for all $t \in I$:

$$(8.27) \quad \delta(\tilde{Q}, \tilde{P}_1) \leq C |\tilde{P}_1|^{1-\eta},$$

and thus, from (5.11), also

$$(8.28) \quad \begin{aligned} \delta(\tilde{Q}, P_1) &\leq \delta_L(\tilde{Q}, \tilde{P}_1) \\ &\leq \delta(\tilde{Q}, \tilde{P}_1) + C |\tilde{P}_1| \\ &\leq C |\tilde{P}_1|^{1-\eta}. \end{aligned}$$

Combining (8.25) and (8.28), we get, for some $t_0 \in I$

$$(8.29) \quad |\tilde{P}_1| \geq C^{-1} \varepsilon^{\frac{2}{1+\eta}},$$

an inequality which actually holds for all $t \in I$ in view of (8.5).

As in the proof of Proposition 20, the number of $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ for which both (8.28), (8.29) hold is easily seen to be at most

$$(8.30) \quad C \varepsilon^{-\frac{2\eta}{1+\eta}}.$$

On the other hand, let $\hat{I} = I$ if $|I| \geq \varepsilon^{2(\beta+1/3)^{-1}}$; otherwise, let \hat{I} be the largest parameter interval which contains I and satisfies $|\hat{I}| < \varepsilon^{2(\beta+1/3)^{-1}}$. As \tilde{P}_1 is I -critical, the same argument as in the proof of Corollary 9 in Subsection 6.7 shows, from (8.29), that we must have

$$(8.31) \quad (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) \in \mathcal{R}(\hat{I}).$$

Now, \tilde{P}_1 is \hat{I} -critical. If $\hat{I} \neq I$, we have, from (8.29) and the definition of \hat{I} , that

$$(8.32) \quad |\tilde{P}_1| > |\hat{I}|^\beta, \quad \text{for all } t \in I.$$

Therefore, $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$ cannot be \hat{I} -bicritical, Q_1 is \hat{I} -transverse and we conclude from Proposition 18 that (P_1, Q_1, n_1) also belongs to $\mathcal{R}(\hat{I})$. The same is also obviously true when $\hat{I} = I$. We apply Proposition 20: for each fixed $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$, the number of children P_1 is at most $|\hat{I}|^{-c\eta}$. But we have

$$(8.33) \quad |\hat{I}| = |I| \geq \varepsilon^{2(\beta+1/3)^{-1}} \quad \text{if } |I| \geq \varepsilon^{2(\beta+1/3)^{-1}}$$

$$(8.34) \quad |\hat{I}| = \varepsilon_0 \quad \text{if } \varepsilon^{2(\beta+1/3)^{-1}} > \varepsilon_0$$

$$(8.35) \quad |\hat{I}| \geq \varepsilon^{2(1+\tau)(\beta+1/3)^{-1}} \quad \text{if } \varepsilon_0 \geq \varepsilon^{2(\beta+1/3)^{-1}} > |I|$$

In all cases, this gives

$$(8.36) \quad |\hat{I}|^{-c\eta} \geq \varepsilon^{-c_0\eta}.$$

Combining Proposition 20 with the previous estimate in (8.30) for the number of possibilities for \tilde{P}_1 gives therefore the required estimate. \square

8.2 A Dimension Estimate

The goal of this subsection is to obtain a bound on the number of elements (P, Q, n) in $\mathcal{R}(I)$ with width $|P|$ bounded from below. This is a first step towards estimating the transverse dimension of the stable set $W^s(\Lambda)$, which is necessary in order to achieve our parameter selection in Section 9.

Let I be a parameter interval, and let (P^*, Q^*, n^*) be an element of $\mathcal{R}(I)$. We introduce, in the spirit of Laplace, Dirichlet and Poincaré, the series

$$(8.37) \quad \Theta(P^*, I, s) = \sum |P|^s,$$

where the sum runs over elements $(P, Q, n) \in \mathcal{R}(I)$ such that $P \subset P^*$. Here s is a complex variable and the series is at first a formal object, but we will soon see that it is uniformly convergent in a

half-plane $\{\operatorname{Re} s > \sigma_0\}$. The goal of this subsection is to obtain a nice estimate for σ_0 and for Θ in this half-plane.

The width $|P|$ and therefore also the series Θ , depend on the parameter $t \in I$; but the estimate that we will get is uniform with respect to the parameter. The dependence of the estimate on P^* is also quite straightforward, through a simple scaling factor.

Let us recall that we denote by d_s^0 the transverse Hausdorff dimension of the stable foliation $W^s(K)$ of the initial horseshoes K for the value 0 of the parameter. It is well-known that this transverse Hausdorff dimension depends smoothly on the parameter, and it controls in a precise way the number of cylinders (for the Markov partition) of a given size; more precisely, as these cylinders correspond exactly to the elements of $\mathcal{R}(I_0)$, we know that, for all $t \in I_0$ and all $\varepsilon > 0$ the number of $(P, Q, n) \in \mathcal{R}(I_0)$ such that $|P| \geq \varepsilon$ is at most

$$(8.38) \quad C \varepsilon^{-(d_s^0 + C\varepsilon_0)}.$$

This shows that for $\Theta(P^*, I_0; s)$ we could take $\sigma_0 = d_s^0 + C\varepsilon_0$. For smaller parameter intervals, we have to allow a slightly larger margin with relation to the initial value d_s^0 .

Proposition 22. *The series $\Theta(P^*, I, s)$ is uniformly convergent in the half-plane $\{\operatorname{Re} s \geq d_s^0 + \varepsilon_0^{\frac{1}{3}}\}$. When $(P^*, Q^*, n^*) \in \mathcal{R}(I_0)$, we have for $\operatorname{Re} s \geq d_s^0 + \varepsilon_0^{\frac{1}{5}}$*

$$|\Theta(P^*, I; s) - \Theta(P^*, I_0; s)| \geq |P^*|^s \varepsilon_0^{\frac{1}{20} d_s^0}.$$

Proof. Let (P, Q, n) be an element of $\mathcal{R}(I)$ with $P \subset P^*$. Consider the intermediary rectangles

$$(8.39) \quad P^* = P(0) \subset P(1) \subset \cdots \subset P(\ell) = P$$

with $P(i)$ the parent of $P(i+1)$. Let

$$(8.40) \quad \ell_0 < \ell_1 < \cdots < \ell_{k-1}$$

be the indices such that $P(\ell_j + 1)$ is a non-simple child of $P(\ell_j)$.

We also define for $0 \leq j \leq k$ elements $(P^{(j)}, Q^{(j)}, n^{(j)}) \in \mathcal{R}(I_0)$ by the following properties

$$(8.41) \quad (P(\ell_0), Q(\ell_0), n(\ell_0)) = (P(0), Q(0), n(0)) * (P^{(0)}, Q^{(0)}, n^{(0)}),$$

$$(8.42) \quad (P(\ell_j), Q(\ell_j), n(\ell_j)) = (P(\ell_{j-1} + 1), Q(\ell_{j-1} + 1), n(\ell_{j-1} + 1)) * (P^{(j)}, Q^{(j)}, n^{(j)}),$$

$$(8.43) \quad (P, Q, n) = (P(\ell_{k-1} + 1), Q(\ell_{k-1} + 1), n(\ell_{k-1} + 1)) * (P^{(k)}, Q^{(k)}, n^{(k)}).$$

We now estimate the widths from (3.10):

$$(8.44) \quad |P(\ell_0)| \leq C|P^*| |P^{(0)}|,$$

$$(8.45) \quad |P(\ell_j)| \leq C|P(\ell_{j-1} + 1)| |P^{(j)}|,$$

$$(8.46) \quad |P| \leq C|P(\ell_{k-1} + 1)| |P^{(k)}|.$$

From (3.27) and property (R7), we also have:

$$(8.47) \quad |P(\ell_j + 1)| < \varepsilon_0^{\frac{1}{2}} |P(\ell_j)|.$$

Define m_j for $0 \leq j < k$ to be the largest integer such that, for all $t \in I$

$$(8.48) \quad |P(\ell_j + 1)| \leq 2^{-m_j} \varepsilon_0^{\frac{1}{2}} |P(\ell_j)|.$$

From Proposition 21, for each fixed $P(\ell_j)$, the number of non-simple children $P(\ell_j + 1)$ satisfying (8.48) is at most

$$(8.49) \quad \left(2^{m_j+1} \varepsilon_0^{-\frac{1}{2}}\right)^{c'\eta}.$$

Combining (8.44), (8.45), (8.46) and (8.48), we also have

$$(8.50) \quad |P| \leq C^{k+1} |P^*| \left(\prod_0^k |P^{(j)}| \right) \varepsilon_0^{\frac{k}{2}} 2^{-\Sigma m_j}.$$

We will take this to the power s and sum over P . We introduce (corresponding to the term $|P^{(j)}|^s$)

$$(8.51) \quad \begin{aligned} \Theta_0(s) &:= \sum_{(\hat{P}, \hat{Q}, \hat{n}) \in \mathcal{R}(I_0)} |\hat{P}|^s \\ &= \sum_a \Theta(R_a, I_0, s), \end{aligned}$$

and also

$$(8.52) \quad \theta(s) := \sum_{m \geq 0} (C \varepsilon_0^{\frac{1}{2}} 2^{-m})^s (2^{m+1} \varepsilon_0^{-\frac{1}{2}})^{c'\eta}.$$

The function Θ_0 is controlled by (8.38), while θ satisfies

$$(8.53) \quad \theta(s) = 2^{c'\eta} C^s \varepsilon_0^{\frac{1}{2}(s-c'\eta)} \left(1 - 2^{-(s-c'\eta)}\right)^{-1},$$

and therefore, for $C^{-1} < s < C$:

$$(8.54) \quad C^{-1} \varepsilon_0^{\frac{1}{2}(s-c'\eta)} \leq \theta(s) \leq C \varepsilon_0^{\frac{1}{2}(s-c'\eta)}.$$

From (8.38), we have, for $s > d_s^0 + C\varepsilon_0$:

$$(8.55) \quad \Theta_0(s) \leq C (s - d_s^0 - C\varepsilon_0)^{-1}.$$

In particular, for $s \geq d_s^0 + \varepsilon_0^{1/3} d_s^0$, we have

$$(8.56) \quad \Theta_0(s) \leq C \varepsilon_0^{-\frac{1}{3} d_s^0},$$

$$(8.57) \quad \Theta_0(s) \theta(s) \leq C \varepsilon_0^{\frac{1}{10} d_s^0}.$$

But, from (8.50), we have for real s

$$(8.58) \quad \Theta(P^*, I, s) \leq C^s |P^*|^s \sum_{k \geq 0} \Theta_0^{k+1}(s) \theta^k(s),$$

and therefore we deduce from (8.57) that the series defining Θ is uniformly convergent in the half plane $\{\operatorname{Re} s \geq d_s^0 + \varepsilon_0^{\frac{1}{3} d_s^0}\}$.

Assume now that $(P^*, Q^*, n^*) \in \mathcal{R}(I_0)$; the difference $\Theta(P^*, I, s) - \Theta(P^*, I_0, s)$ consists of the sum of $|P|^s$ over those P for which $k > 0$. We get, for real s

$$(8.59) \quad \Theta(P^*, I, s) - \Theta(P^*, I_0, s) \leq C^s |P^*|^s \sum_{k > 0} \Theta_0^{k+1}(s) \theta^k(s).$$

For $s > d_s^0 + \varepsilon_0^{\frac{1}{5} d_s^0}$, we have, from (8.55), (8.54):

$$(8.60) \quad \Theta_0(s) \leq C \varepsilon_0^{-\frac{1}{5} d_s^0},$$

$$(8.61) \quad \Theta_0^2 \theta(s) \leq C \varepsilon_0^{\frac{1}{15} d_s^0},$$

which gives the second part of the proposition. □

8.3 Transfer to Parameter Space

8.3.1 Our goal in this subsection will be to prove the following result, which expresses a transfer of the dimension estimate of Subsection 8.2 to parameter space.

Proposition 23. *Let \tilde{I} be a regular parameter interval. Let (P^*, Q^*, n^*) be an element of $\mathcal{R}(\tilde{I})$ such that Q^* is \tilde{I} -critical and*

$$(8.62) \quad |Q^*| \leq \frac{1}{2} |\tilde{I}|^{(1+\tau)(1-\eta)^{-1}}$$

for all $t \in \tilde{I}$. The, the number of candidates $I \subset \tilde{I}$ of the next level, such that Q^ is I -critical, is at most $|\tilde{I}|^{-\tau d_s^+}$, where $d_s^+ = d_s^0 + C\eta\tau^{-1}$ can be made arbitrarily close to d_s^0 .*

Recall that the total number of candidates is $|\tilde{I}|^{-\tau}$. Proposition 23 is the key estimate that will allow us in Section 9 to proceed with the selection process for parameters. The rest of the section is devoted to the proof of Proposition 23.

8.3.2 We start with some general observations, that could have been made much earlier, but are only useful now.

Let I be a parameter interval. Let (P, Q, n) , (P_0, Q_0, n_0) , (P_1, Q_1, n_1) , be elements of $\mathcal{R}(I)$ such that $Q \subset Q_u$, $P_0 \subset P_s$, $P_1 \subset P_s$ and $P_0 \cap P_1 = \emptyset$. Associated to the pair (Q, P_0) (resp. (Q, P_1)), we have defined in Subsection 3.5 a function $C_0(w, y, x)$ (resp. $C_1(w, y, x)$) which is the basis for the definition of $\delta(Q, P_0)$, $\delta_{LR}(Q, P_0), \dots$ (resp. $\delta(Q, P_1)$, $\delta_{LR}(Q, P_1), \dots$). It follows immediately from the monotonicity of the function θ that, since P_0 and P_1 are disjoint, we have either

$$(8.63) \quad C_0(w, y, x_0) > C_1(w, y, x_1) \quad \text{for all } w, y, x_0, x_1, t, \text{ or}$$

$$(8.64) \quad C_0(w, y, x_0) < C_1(w, y, x_1) \quad \text{for all } w, y, x_0, x_1, t.$$

In the first case, we have, from the definitions in Subsection 5.4:

$$(8.65) \quad \begin{aligned} \delta_R(Q, P_0) &< \delta(Q, P_1), \\ \delta_{LR}(Q, P_0) &< \delta_L(Q, P_1), \end{aligned}$$

while in the second the same inequalities hold exchanging P_0 and P_1 .

We see in particular that if (8.63) holds and Q, P_1 are I -separated, then Q, P_0 are also I -separated.

On the other hand, we will prove (see figure 8).

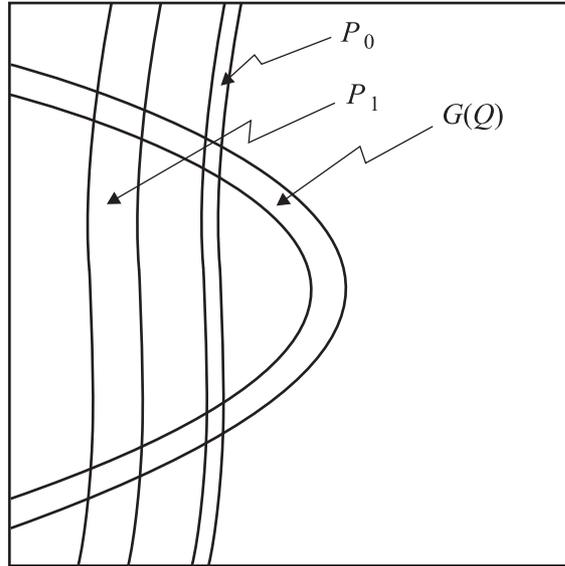


Figure 8

Proposition 24. *Assume that (8.63) holds and that $|P_1|^{1-\eta} \leq |I|$ for some $t \in I$. Then, if Q and P_0 are I -transverse, Q and P_1 are also I -transverse.*

Proof. Let $\tilde{I} \supset I$ and $(\tilde{P}, \tilde{Q}, \tilde{n}), (\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0) \in \mathcal{R}(\tilde{I})$ be such that $\tilde{Q} \supset Q, \tilde{P}_0 \supset P_0$ and $\tilde{Q} \bar{\cap}_{\tilde{I}} \tilde{P}_0$ holds. If $P_1 \subset \tilde{P}_0$, we immediately conclude that Q and P_1 are I -transverse. We assume, therefore, that $P_1 \cap \tilde{P}_0 = \emptyset$; replacing (P_0, Q_0, n_0) by $(\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0)$, and (P, Q, n) by $(\tilde{P}, \tilde{Q}, \tilde{n})$, we can also assume that $(P, Q, n), (P_0, Q_0, n_0) \in \mathcal{R}(\tilde{I})$ and $Q \bar{\cap}_{\tilde{I}} P_0$ holds.

Let $(\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1)$, be the element of $\mathcal{R}(\tilde{I})$ with $P_1 \subset \tilde{P}_1$ and smallest \tilde{P}_1 . We will prove that Q and \tilde{P}_1 are \tilde{I} -transverse. We have, for all $t \in \tilde{I}$, that

$$(8.66) \quad \delta_{LR}(Q, \tilde{P}_1) > \delta_{LR}(Q, P_0) \geq 2|\tilde{I}|,$$

and also for some $t_0 \in \tilde{I}$,

$$(8.67) \quad \delta_R(Q, \tilde{P}_1) > \delta_R(Q, P_0) \geq 2|Q|^{1-\eta}.$$

Therefore, assuming that $Q \bar{\cap}_{\tilde{I}} \tilde{P}_1$ does not hold, we must have, for all $t \in \tilde{I}$, that

$$(8.68) \quad \delta_L(Q, \tilde{P}_1) < 2|\tilde{P}_1|^{1-\eta}.$$

We cannot have in this case $\tilde{P}_1 = P_1$, because, for all $t \in I$,

$$(8.69) \quad \delta_L(Q, P_1) > \delta_{LR}(Q, P_0) \geq 2|\tilde{I}|,$$

and (8.68), (8.69) together would contradict the hypothesis of the proposition. Therefore, \tilde{P}_1 strictly contains P_1 and \tilde{I} strictly contains I . But, then, applying the structure theorem of Subsection 6.7 to the child of \tilde{P}_1 which contains P_1 , we obtain that \tilde{Q}_1 is \tilde{I} -critical. As \tilde{I} is β -regular, it then follows from (8.68), (8.69) and (5.14) that \tilde{P}_1 is \tilde{I} -transverse. This implies that there exists $(P', Q', n') \in \mathcal{R}(\tilde{I})$ with $Q \cap Q' \neq \emptyset$ such that $Q' \bar{\cap}_{\tilde{I}} \tilde{P}_1$ holds. If $Q \subset Q'$, it follows that $Q \bar{\cap}_{\tilde{I}} \tilde{P}_1$ holds. When $Q' \subset Q$ we use both $Q \bar{\cap}_{\tilde{I}} P_0$ and $Q' \bar{\cap}_{\tilde{I}} \tilde{P}_1$ to conclude, as in the proof of Proposition 4, that $Q \bar{\cap}_{\tilde{I}} \tilde{P}_1$. \square

8.3.3 We now switch back to the setting of Proposition 23.

Let $(P, Q, n) \in \mathcal{R}(\tilde{I})$ with $P \subset P_s$. We say that P is *eventually \tilde{I} -separated* from Q^* if there exists an \tilde{I} -decomposition $(P_\alpha, Q_\alpha, n_\alpha)$ of P such that Q^* and P_α are \tilde{I} -separated for every α . We say that P is *eventually \tilde{I} -transverse* to Q^* if there exists an \tilde{I} -decomposition $(P_\alpha, Q_\alpha, n_\alpha)$ of P such that Q^* and P_α are \tilde{I} -transverse for every α . We say that P is *eventually \tilde{I} - Q^* -critical* if it is neither eventually \tilde{I} -separated from Q^* nor eventually \tilde{I} -transverse to Q^* .

Lemma 4. *If P is eventually \tilde{I} -transverse to Q^* and $2|P|^{1-\eta} \leq |\tilde{I}|$ holds for some $t \in \tilde{I}$, then $Q^* \bar{\cap}_{\tilde{I}} P$ holds.*

Proof. This is an immediate consequence of Proposition 9. \square

Lemma 5. *If P is eventually \tilde{I} - Q^* -critical, then P is \tilde{I} -critical.*

Proof. If P was \tilde{I} -transverse, there would exist $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{R}(\tilde{I})$ with $Q^* \cap Q_\alpha \neq \emptyset$ and $Q_\alpha \pitchfork_{\tilde{I}} P$. But, then, from Proposition 2 if $Q_\alpha \supset Q^*$ or Proposition 9 if $Q_\alpha \subset Q^*$, we would deduce that $Q^* \pitchfork_{\tilde{I}} P$ also holds. \square

Lemma 6. *If P is eventually \tilde{I} - Q^* -critical and $|P| > |\tilde{I}|^\beta$ holds for some $t \in \tilde{I}$, then some child of P is also eventually \tilde{I} - Q^* -critical.*

Proof. By Lemma 5 and Corollary 4 in Subsection 6.2, P is \tilde{I} -decomposable. If all children of P were eventually \tilde{I} -transverse to Q^* (resp. eventually \tilde{I} -separated from Q^*), we would put together the corresponding \tilde{I} -decompositions and obtain that P is eventually \tilde{I} -transverse to Q^* (resp. eventually \tilde{I} -separated from Q^*). Therefore, we shall assume that some child P_0 is eventually \tilde{I} -transverse to Q^* , and some child P_1 is eventually \tilde{I} -separated from Q^* .

By contradiction, we assume that none of the children is eventually \tilde{I} - Q^* -critical and we will show that Q^* is \tilde{I} -transverse. We construct an \tilde{I} -decomposition $(P_\alpha, Q_\alpha, n_\alpha)$ of P_s such that every P_α is either \tilde{I} -separated from Q^* or \tilde{I} -transverse to Q^* in the following way.

Actually, it is sufficient to have an \tilde{I} -decomposition such that every P_α is either eventually \tilde{I} -separated from Q^* or eventually \tilde{I} -transverse to Q^* . Starting from the trivial decomposition of P_s , we have at step i an \tilde{I} -decomposition $(P_\alpha^{(i)}, Q_\alpha^{(i)}, n_\alpha^{(i)})$. As long as there is one $(P_\alpha^{(i)}, Q_\alpha^{(i)}, n_\alpha^{(i)})$ with $P \subset P_\alpha^{(i)}$, we observe that $P_\alpha^{(i)}$ is \tilde{I} -critical and therefore \tilde{I} -decomposable and break it into its children to go to step $i + 1$.

After a finite number of steps, each $P_\alpha^{(i)}$ is either a child of P or disjoint from P . Comparing $P_\alpha^{(i)}$ with P_0 and P_1 , we conclude from Proposition 24 or from the remark before this proposition that $P_\alpha^{(i)}$ is either eventually \tilde{I} -transverse to Q^* or \tilde{I} -separated from Q^* . Thus, we have constructed the required \tilde{I} -decomposition and the lemma is proved. \square

Lemma 7. *If P_0, P_1 are eventually \tilde{I} - Q^* -critical and disjoint, then we have $|P_0| \leq C_0 |\tilde{I}|$, $|P_1| \leq C_0 |\tilde{I}|$ for all $t \in \tilde{I}$.*

Proof. From Lemma 6, we can find $(\hat{P}_0, \hat{Q}_0, \hat{n}_0), (\hat{P}_1, \hat{Q}_1, \hat{n}_1)$ in $\mathcal{R}(\tilde{I})$ with $\hat{P}_0 \subset P_0, \hat{P}_1 \subset P_1$, such that both \hat{P}_0, \hat{P}_1 are eventually \tilde{I} - Q^* -critical and we have

$$(8.70) \quad |\hat{P}_0| \leq |\tilde{I}|^\beta, \quad |\hat{P}_1| \leq |\tilde{I}|^\beta \quad \text{for all } t \in \tilde{I}.$$

If we had, for all $t \in \tilde{I}$ and $i = 0$ or 1 ,

$$(8.71) \quad \delta(Q^*, \hat{P}_i) \geq 2|\tilde{I}|,$$

then, from (8.62) and (8.70), one would have that $Q^* \pitchfork_{\tilde{I}} \hat{P}_i$ holds and \hat{P}_i would not be eventually \tilde{I} - Q^* -critical. We have, therefore,

$$(8.72) \quad \delta_{LR}(Q^*, \hat{P}_0) < 2|\tilde{I}| \quad \text{for some } t_0 \in \tilde{I} \text{ and}$$

$$(8.73) \quad \delta_{LR}(Q^*, \widehat{P}_1) < 2|\widetilde{I}| \quad \text{for some } t_1 \in I.$$

In the same way, we must have

$$(8.74) \quad \delta_{LR}(Q^*, \widehat{P}_0) \geq 0 \quad \text{for some } t'_0 \in \widetilde{I} \text{ and also.}$$

$$(8.75) \quad \delta'_{LR}(Q^*, \widehat{P}_1) \geq 0 \quad \text{for some } t'_1 \in \widetilde{I}.$$

Let (P', Q', n') , (P'_0, Q'_0, n'_0) , (P'_1, Q'_1, n'_1) be the elements of $\mathcal{R}(\widetilde{I})$ such that $Q^* \subset Q'$, $\widehat{P}_0 \subset P'_0$, $\widehat{P}_1 \subset P'_1$, and

$$(8.76) \quad |Q'| < |\widetilde{I}|, \quad |P'_0| < |\widetilde{I}|, \quad |P'_1| < |\widetilde{I}|, \quad \forall t \in \widetilde{I}$$

and which are maximal with these properties.

Then, P'_0 satisfies condition $(*s)$ of Subsection 7.6: this follows easily from (8.72). Similarly, P'_1 satisfies condition $(*s)$ and Q' satisfies condition $(*u)$. We are therefore allowed to apply the estimate of Corollary 11 in Subsection 7.6 to conclude that

$$(8.77) \quad \max_{\widetilde{I}} \delta_{LR}(Q', P'_i) - \min_{\widetilde{I}} \delta_{LR}(Q', P'_i) \leq 2|\widetilde{I}|$$

for $i = 0, 1$, and similarly for δ , δ_L , δ_R . We either have (8.63) or (8.64). Assume, for instance, that (8.63) holds. If $P_1 \subset P'_1$, we have $|P_1| \leq |\widetilde{I}|$ from (8.76). If $P'_1 \subset P_1$, we have

$$(8.78) \quad \begin{aligned} \max_{\widetilde{I}} \delta_L(Q^*, P_1) &\leq \max_{\widetilde{I}} \delta_L(Q', P'_1) \\ &\leq \min_{\widetilde{I}} \delta_L(Q', P'_1) + 2|\widetilde{I}| \\ &\leq \min_{\widetilde{I}} (\delta_L(Q^*, P_1) + c|Q'|) + 2|\widetilde{I}| \\ &\leq c|\widetilde{I}|. \end{aligned}$$

Similarly, we have $|P_0| \leq |\widetilde{I}|$ from (8.76) if $P_0 \subset P'_0$ and otherwise

$$(8.79) \quad \begin{aligned} \min_{\widetilde{I}} \delta_{LR}(Q^*, P_0) &\geq \min_{\widetilde{I}} \delta_{LR}(Q', P'_0) - c|\widetilde{I}| \\ &\geq \max_{\widetilde{I}} \delta_{LR}(Q', P'_0) - (c+2)|\widetilde{I}| \\ &\geq -c|\widetilde{I}|. \end{aligned}$$

Observe that, for all $t \in \widetilde{I}$, the inequalities

$$(8.80) \quad c|\widetilde{I}| \geq \delta_L(Q^*, P_1) > \delta_{LR}(Q^*, P_0) \geq -c|\widetilde{I}|$$

are also valid when $P_1 \subset P'_1$ or $P_0 \subset P'_0$. However, there is a constant $c > 0$ such that, for any child P_1^* of P_1 , any child P_0^* of P_0 , we have, for all $t \in I$, that

$$(8.81) \quad \delta_L(Q^*, P_1^*) \geq \delta_L(Q^*, P_1) + c|P_1|,$$

$$(8.82) \quad \delta_{LR}(Q^*, P_0^*) \leq \delta_{LR}(Q^*, P_0) - c|P_0|.$$

As we also must have, for some children P_1^*, P_0^* , that

$$(8.83) \quad \delta_L(Q^*, P_1^*) \leq c|\tilde{I}|,$$

$$(8.84) \quad \delta_{LR}(Q^*, P_0^*) \geq -c|\tilde{I}|,$$

because these children are eventually \tilde{I} - Q^* -critical, we obtain the conclusion of the lemma. \square

8.3.4 Consider the set Π of elements $(P, Q, n) \in \mathcal{R}(\tilde{I})$ which are eventually \tilde{I} - Q^* -critical, satisfy

$$(8.85) \quad |P| \leq |\tilde{I}|^{1+\tau}$$

for all $t \in \tilde{I}$ and are maximal (in P) with respect to these two properties.

Lemma 8. *We have*

$$\#\Pi \leq \frac{1}{C_1} |\tilde{I}|^{-\tau d_s^+},$$

where $d_s^+ = d_s^0 + C\eta\tau^{-1}$ is as in the statement of Proposition 24.

Proof. From Lemma 7, there exists a unique element $(P_0, Q_0, n_0) \in \mathcal{R}(\tilde{I})$ with the following properties:

- $P \subset P_0$ for all $(P, Q, n) \in \Pi$
- $|P_0| > C_0|\tilde{I}|$ for some $t \in \tilde{I}$
- every child P_1 of P_0 which contains a rectangle P with $(P, Q, n) \in \Pi$ satisfies $|P_1| \leq C_0|\tilde{I}|$ for all $t \in \tilde{I}$.

As P_0 is eventually \tilde{I} - Q^* -critical and $|P_0| > C_0|\tilde{I}|$ for some $t \in \tilde{I}$, P_0 must be \tilde{I} -critical. From Proposition 20, the number of children of P_0 is at most $|\tilde{I}|^{-c\eta}$.

For every $(P, Q, n) \in \Pi$, either P is a child of P_0 or the parent \tilde{P} of P is contained in a child P_1 of P_0 . Observe that we have by definition of Π

$$(8.86) \quad |\tilde{P}| > |\tilde{I}|^{1+\tau} \quad \text{for some } t \in \tilde{I}.$$

As \tilde{P} satisfies condition $(*s)$ of Subsection 7.6 since P_0 is \tilde{I} -critical, we also have

$$(8.87) \quad |\tilde{P}| \geq c^{-1} |\tilde{I}|^{1+\tau} \quad \text{for all } t \in \tilde{I}.$$

As we also have

$$(8.88) \quad |P_1| \leq C_0|\tilde{I}| \quad \text{for all } t \in \tilde{I},$$

the number of possible \tilde{P} , given P_1 , is from Proposition 22 at most,

$$(8.89) \quad (C|\tilde{I}|^\tau)^{-d_s^*}$$

with $d_s^* = d_s^0 + \varepsilon_0^{\frac{1}{5}} d_s^0$. The number of children of \tilde{P} is at most $|\tilde{I}|^{-c\eta}$. This gives a bound for the cardinality of Π by

$$(8.90) \quad |\tilde{I}|^{-c\eta} + |\tilde{I}|^{-2c\eta} (C|\tilde{I}|^\tau)^{-d_s^*},$$

in accordance with the statement of Lemma 8. □

8.3.5 Proof of Proposition 23: By Lemma 5 and Corollary 4 in Subsection 6.2, every $(P, Q, n) \in \mathcal{R}(\tilde{I})$ such that P is eventually \tilde{I} - Q^* -critical and $|P| \geq |\tilde{I}|^\beta$ (for some $t \in \tilde{I}$) is \tilde{I} -decomposable.

Therefore, there exists an \tilde{I} -decomposition $(P_\alpha, Q_\alpha, n_\alpha)$ of P_s such that every $(P_\alpha, Q_\alpha, n_\alpha)$ is either eventually \tilde{I} -separated from Q^* or eventually \tilde{I} -transverse to Q^* or an element of Π .

Let $I \subset \tilde{I}$ be a candidate interval of the next level, i.e. $|I| = |\tilde{I}|^{1+\tau}$, such that Q^* is I -critical.

We claim that there exists $(P, Q, n) \in \Pi$ such that P is eventually I - Q^* -critical.

Indeed, every $(P_\alpha, Q_\alpha, n_\alpha)$ which is eventually \tilde{I} -transverse to Q^* (resp. eventually \tilde{I} -separated from Q^*) is a fortiori I -transverse to Q^* (resp. I -separated from Q^*). If every $(P_\alpha, Q_\alpha, n_\alpha) \in \Pi$ was also either eventually I -transverse to Q^* or eventually I -separated from Q^* , we would obtain a decomposition of P_s which expresses that Q^* is I -transverse.

On the other hand, fix $(P, Q, n) \in \Pi$. We show that there are at most C_1 candidates $I \subset \tilde{I}$ such that (P, Q, n) is I -critical. Together with Lemma 8, this will imply the statement of Proposition 23.

Choose $(P', Q', n') \in \mathcal{R}(\tilde{I})$, with $Q^* \subset Q'$, $|Q'| < |\tilde{I}|$ for all $t \in \tilde{I}$, and maximal with this property as in the proof of Lemma 7. Then, Q' satisfies condition $(*u)$ of Subsection 7.6; we already know that P satisfies condition $(*s)$. Then, by Corollary 11 of Subsection 7.6, we have

$$(8.91) \quad \left| \frac{d}{dt} \delta_{LR}(Q', P) - 1 \right| \leq C \varepsilon_0^{\frac{1}{2}}.$$

But, if we have for all $t \in I$

$$(8.92) \quad \delta_{LR}(Q', P) < 0,$$

then Q' and P are I -separated and a fortiori Q^* and P are I -separated.

On the other hand, if we have

$$(8.93) \quad \delta(Q', P) > 3|I| \quad \text{for all } t \in I,$$

it is easy to conclude from (8.62) and Proposition 9, that P is eventually I -transverse to Q^* . But, by (5.11)–(5.14), we have

$$(8.94) \quad \delta_{LR}(Q', P) < \delta(Q', P) + c|I|$$

and our claim then follows from (8.91). Proposition 23 is proved. □

9 Strong Regularity and Parameter Selection

9.1 Partitions of the Critical Locus

As it was mentioned in Subsection 4.5, regularity is a rather qualitative property which is not appropriate for the quantitative estimates needed for parameter selection. Therefore, we will introduce later in this section, a stronger quantitative property, that we call *strong regularity*.

This property is built up from several bounds: the number of bicritical elements of a given size (including, of course, that there are no "fat" bicritical elements), and also sizes of the "critical locus" (see below), which should be of approximate dimension $d_s^0 + d_u^0 - 1$. These last bounds are more elementary and will be taken care of first.

In all of Section 9, we fix a parameter interval \tilde{I} which is strongly regular, i.e. satisfies all bounds that will be stated shortly. We will prove at some point that the starting interval satisfies this strong regularity. As mentioned before, and as will be proved in Subsection 9.6, strong regularity implies β -regularity for some $\beta > 1$. We denote by I any parameter interval contained in \tilde{I} of the next level, i.e. $|I| = |\tilde{I}|^{1+\tau}$; such a candidate interval is $\bar{\beta}$ -regular (Corollary 9 in Subsection 6.7), with $\bar{\beta} = \beta(1 + \tau)^{-1}$. The aim of this section is to estimate how many of the candidates I fail to be strongly regular.

We denote by $\mathcal{C}_+(\tilde{I})$ the set of (P, Q, n) in $\mathcal{R}(\tilde{I})$ such that P is \tilde{I} -critical, $|P| \leq |\tilde{I}|^{1+\tau}$ for all $t \in \tilde{I}$, and P is maximal with this property: the parent \tilde{P} of P satisfies $|\tilde{P}| > |\tilde{I}|^{1+\tau}$ for some $t \in \tilde{I}$.

Obviously, if $(P, Q, n), (P', Q', n')$ are distinct elements in $\mathcal{C}_+(\tilde{I})$, P and P' are disjoint. Moreover, if $(\hat{P}, \hat{Q}, \hat{n})$ belongs to $\mathcal{R}(\tilde{I})$, \hat{P} is \tilde{I} -critical, and $|\hat{P}| \leq |\tilde{I}|^{1+\tau}$ for all $t \in \tilde{I}$, there is a unique $(P, Q, n) \in \mathcal{C}_+(\tilde{I})$ such that $\hat{P} \subset P$.

Exchanging P 's and Q 's, we define $\mathcal{C}_-(\tilde{I})$ in a similar way. The sets $\mathcal{C}_+(\tilde{I}), \mathcal{C}_-(\tilde{I})$ correspond to the \tilde{I} -critical locus at the $|\tilde{I}|^{1+\tau}$ scale. We will also need to consider this at the $|\tilde{I}|$ scale, as follows. We define $\hat{\mathcal{C}}_+(\tilde{I})$ to be the set of $(P, Q, n) \in \mathcal{R}(\tilde{I})$ such that P contains some P' with $(P', Q', n') \in \mathcal{C}_+(\tilde{I})$, $|P| \leq |\tilde{I}|$ for all $t \in \tilde{I}$, and P is maximal with this property. We define similarly $\hat{\mathcal{C}}_-(\tilde{I})$.

We will need in the sequel to consider \hat{I} -criticality (for some parameter interval $\hat{I} \supset \tilde{I}$) for rectangles in $\mathcal{R}(\tilde{I})$ but not in $\mathcal{R}(\hat{I})$. The following definition will be useful.

Definition. Let $I_\alpha \supset I$ be parameter intervals, and let $(P, Q, n) \in \mathcal{R}(I)$. We say that P is *thin I_α -critical* if there exists $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{R}(I_\alpha)$ with $P \subset P_\alpha$, P_α is I_α -critical and

$$|P_\alpha|^{1-\eta} \leq 2|I_\alpha|$$

for some $t \in I_\alpha$.

The main justification for this definition is the following lemma, which is an immediate consequence of the structure theorem of Subsection 6.7 and Proposition 11 in Subsection 6.6.

Lemma 9. *Let I be a parameter interval with parent \tilde{I} , and let (P, Q, n) be an element of $\mathcal{R}(I)$ which is not the restriction of an element of $\mathcal{R}(\tilde{I})$. Let (P_i, Q_i, n_i) , for $0 \leq i \leq k$, be the elements of $\mathcal{R}(\tilde{I})$ given by the structure theorem of Subsection 6.7. Then P_i is thin \tilde{I} -critical for $0 < i \leq k$, and Q_i is thin \tilde{I} -critical for $0 \leq i < k$.*

9.2 Size of the Critical Locus

We will state several inequalities related to the size of the sets $\mathcal{C}_+(I)$, $\mathcal{C}_-(I)$, $\hat{\mathcal{C}}_+(I)$, $\hat{\mathcal{C}}_-(I)$. All these inequalities are part of the definition of strong regularity: they have to be satisfied by a strongly regular parameter interval. We will then see that if these inequalities are satisfied by the parent interval \tilde{I} , most of the candidates $I \subset \tilde{I}$ also satisfy these inequalities.

Recall that d_s^0, d_u^0 are the transverse Hausdorff dimensions of $W^s(K)$, $W^u(K)$ at $t = 0$. In Proposition 23 of Subsection 8.3, we introduced $d_s^+ = d_s^0 + C\eta\tau^{-1}$. Let also $d_u^+ = d_u^0 + C\eta\tau^{-1}$.

In Proposition 22 of Subsection 8.2, we have used $d_s^* := d_s^0 + \varepsilon_0^{\frac{1}{5}} d_s^0 < d_s^+$. Define similarly

$$d_u^* = d_u^0 + \varepsilon_0^{\frac{1}{5}} d_u^0.$$

We will control the cardinalities of $\mathcal{C}_+(I)$, $\mathcal{C}_-(I)$, $\hat{\mathcal{C}}_+(I)$, $\hat{\mathcal{C}}_-(I)$, through:

$$(SR1)_s \quad \#\mathcal{C}_+(I) \leq C \left(\frac{|I|}{\varepsilon_0} \right)^{1-d_s^+-d_u^+-\tau} \varepsilon_0^{-\tau d_s^0},$$

$$(SR1)_{\hat{s}} \quad \#\hat{\mathcal{C}}_+(I) \leq C \left(\frac{|I|}{\varepsilon_0} \right)^{1-d_s^+-d_u^+-\tau},$$

$$(SR1)_u \quad \#\mathcal{C}_-(I) \leq C \left(\frac{|I|}{\varepsilon_0} \right)^{1-d_s^+-d_u^+-\tau} \varepsilon_0^{-\tau d_u^0},$$

$$(SR1)_{\hat{u}} \quad \#\hat{\mathcal{C}}_-(I) \leq C \left(\frac{|I|}{\varepsilon_0} \right)^{1-d_s^+-d_u^+-\tau}.$$

We will also control a weighted cardinality through

$$(SR2)_s \quad \sum_{\mathcal{C}_+(I)} |Q|^{d_u^*} \leq C |Q_s|^{d_u^*} \left(\frac{|I|}{\varepsilon_0} \right)^{1-d_u^+-\tau},$$

$$(SR2)_{\hat{s}} \quad \sum_{\hat{\mathcal{C}}_+(I)} |Q|^{d_u^*} \leq C |Q_s|^{d_u^*} \left(\frac{|I|}{\varepsilon_0} \right)^{1-d_u^+-\tau},$$

$$(SR2)_u \quad \sum_{\mathcal{C}_-(I)} |P|^{d_s^*} \leq C |P_u|^{d_s^*} \left(\frac{|I|}{\varepsilon_0} \right)^{1-d_s^+-\tau},$$

$$(SR2)_{\hat{u}} \quad \sum_{\hat{\mathcal{C}}_-(I)} |P|^{d_s^*} \leq C |P_u|^{d_s^*} \left(\frac{|I|}{\varepsilon_0} \right)^{1-d_s^+-\tau}.$$

The heuristics behind the second set of inequalities is the following: in the mean, one expects that elements of $\mathcal{R}(I)$ more or less satisfy

$$(9.1) \quad |P|^{d_s^0} \sim |Q|^{d_u^0}$$

and, for $(P, Q, n) \in \widehat{\mathcal{C}}_+(I)$ (for instance), one should have

$$(9.2) \quad |P| \sim |I|$$

which explains the relation between $(\text{SR1})_{\widehat{s}}$ and $(\text{SR2})_{\widehat{s}}$.

Let us check that the starting interval I_0 satisfies these eight inequalities.

Then $\mathcal{C}_+(I_0)$ (resp. $\mathcal{C}_-(I_0)$) consists of the elements $(P, Q, n) \in \mathcal{R}(I_0)$ with $P \cap P_s \neq \emptyset$ (resp. $Q \cap Q_u \neq \emptyset$), $|P| \leq \varepsilon_0^{1+\tau}$ for all $t \in I_0$ (resp. $|Q| \leq \varepsilon_0^{1+\tau}$ for all $t \in I_0$) and maximal with this property. Then, $(\text{SR1})_s$ follows from (8.38) and $(\text{SR1})_u$ is similar. Writing $(P, Q, n) = (P_s, Q_s, n_s) * (P', Q', n')$, the inequality $(\text{SR2})_s$ becomes

$$(9.3) \quad \sum_{\mathcal{C}_+(I)} |Q'|^{d_u^0 + C\varepsilon_0} \leq C$$

which is a standard property of uniformly hyperbolic horseshoes. The proof of $(\text{SR2})_u$ is similar. The case of $\widehat{\mathcal{C}}_+(I_0)$, $\widehat{\mathcal{C}}_-(I_0)$ is even simpler. For the induction step, we have:

Proposition 25. *If the parent interval \widetilde{I} is β -regular and satisfies one of the eight inequalities (SR) above, then all candidates $I \subset \widetilde{I}$ satisfy the same inequality except perhaps for a proportion not larger than $C|\widetilde{I}|^{\tau^2}$.*

Notation. Let $(P, Q, n) \in \mathcal{C}_+(\widetilde{I})$. We denote by $Cr(P)$ the set of candidates $I \subset \widetilde{I}$ such that P contains a thin I -critical rectangle.

Lemma 10. *For any $(P, Q, n) \in \mathcal{C}_+(\widetilde{I})$, we have*

$$\#Cr(P) \leq C|\widetilde{I}|^{-\tau d_u^+}.$$

Proof. Let $(P^*, Q^*, n^*) \in \mathcal{R}(\widetilde{I})$ be an element such that $P^* \subset P$ and

$$(9.4) \quad |P^*| \leq \frac{1}{2} |\widetilde{I}|^{(1+\tau)(1-\eta)^{-1}}$$

for all $t \in \widetilde{I}$. By Proposition 23 in Subsection 8.3, there are at most $|\widetilde{I}|^{-\tau d_u^+}$ candidates $I \subset \widetilde{I}$ such that P^* is I -critical. Let $I \in Cr(P)$. By definition, there exists $(P_0, Q_0, n_0) \in \mathcal{R}(I)$ with $P_0 \subset P$, P_0 I -critical and

$$(9.5) \quad |P_0|^{1-\eta} \leq 2|I| \text{ for some } t_0 \in I.$$

But if P_0 is I -critical and (9.5) holds, there must exist $(P', Q', n') \in \mathcal{R}(I)$ and $t_1 \in I$ such that

$$(9.6) \quad 0 \leq \delta_{LR}(Q', P_0) \leq 2|I|.$$

As P_0, P^* are contained in P and $|P| \leq |I|$ for all $t \in \tilde{I}$, we also have

$$(9.7) \quad |\delta_{LR}(Q', P^*)| \leq C|I|.$$

Proceeding as in the Proof of Lemma 7, Subsection 8.3, we deduce from Corollary 11 in Subsection 7.6, that there exists $I' \subset \tilde{I}$ at distance $c|I|$ of I such that P^* is I' -critical (except perhaps when I is very close to the boundary of \tilde{I}). This proves the lemma. \square

Proof of Proposition 25. We will deal with $(\text{SR1})_s, (\text{SR1})_{\hat{s}}, (\text{SR2})_s, (\text{SR2})_{\hat{s}}$, the other four being symmetric.

For each $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(\tilde{I})$, we consider the elements $(P_{\alpha,i}, Q_{\alpha,i}, n_{\alpha,i}) \in \mathcal{R}(\tilde{I})$ which satisfy $P_{\alpha,i} \subset P_\alpha$, $|P_{\alpha,i}| \leq |\tilde{I}|^{(1+\tau)^2}$ for all $t \in \tilde{I}$, and which are maximal with respect to this property.

This gives an \tilde{I} -decomposition of P_α : indeed, it is easy to see that for each $(P'_\alpha, Q'_\alpha, n'_\alpha) \in \mathcal{R}(\tilde{I})$ with $P'_\alpha \subset P_\alpha$ and $|P'_\alpha| > |\tilde{I}|^{(1+\tau)^2}$ for some $t \in \tilde{I}$, Q'_α is \tilde{I} -transverse and therefore P'_α is \tilde{I} -decomposable.

Using Propositions 20 and 22, we argue as in the proof of Lemma 8, Subsection 8.3, to see that for each P_α , the number of $P_{\alpha,i}$ is not larger than

$$(9.8) \quad |\tilde{I}|^{-c\eta} |\tilde{I}|^{-\tau(1+\tau)d_s^*},$$

with $d_s^* = d_s^0 + \varepsilon_0^{\frac{1}{5}} d_s^0$.

Obviously, if $(P, Q, n) \in \mathcal{C}_+(I)$, there exists α, i such that $P \supset P_{\alpha,i}$ and I must belong to $Cr(P_\alpha)$. We, therefore, have

$$(9.9) \quad \begin{aligned} \sum_{I \subset \tilde{I}} \#\mathcal{C}_+(I) &\leq C |\tilde{I}|^{-c\eta - \tau(1+\tau)d_s^*} \sum_{\mathcal{C}_+(\tilde{I})} \#Cr(P_\alpha) \\ &\leq \#\mathcal{C}_+(\tilde{I}) |\tilde{I}|^{-c\eta - \tau(1+\tau)d_s^*} \max \#Cr(P_\alpha). \end{aligned}$$

We have

$$(9.10) \quad d_s^+ \geq d_s^*(1 + \tau) + C\eta\tau^{-1}$$

and therefore, using also Lemma 10:

$$(9.11) \quad \sum_{I \subset \tilde{I}} \#\mathcal{C}_+(I) \leq C \#\mathcal{C}_+(\tilde{I}) |\tilde{I}|^{-\tau(d_s^+ + d_u^+)}.$$

The induction step for $(\text{SR1})_s$ follows immediately. In the same way as in (9.8), we obtain from Propositions 20 and 22 that

$$(9.12) \quad \#\mathcal{C}_+(\tilde{I}) \leq |\tilde{I}|^{-C\eta} |\tilde{I}|^{-\tau d_s^*} \#\hat{\mathcal{C}}_+(\tilde{I}).$$

On the other hand, for every $(P, Q, n) \in \widehat{\mathcal{C}}_+(I)$, there exists a $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(\tilde{I})$ with $P_\alpha \subset P$ and $I \in Cr(P_\alpha)$. Therefore, we have

$$(9.13) \quad \sum_{I \in \tilde{I}} \#\widehat{\mathcal{C}}_+(I) \leq \sum_{\mathcal{C}_+(\tilde{I})} \#Cr(P_\alpha).$$

Putting together (9.12), (9.13) and Lemma 10, we get

$$(9.14) \quad \sum \#\widehat{\mathcal{C}}_+(I) \leq C|\tilde{I}|^{-C\eta - \tau d_s^* - \tau d_u^+} \#\widehat{\mathcal{C}}_+(\tilde{I})$$

and we conclude as above.

Let us now consider $(SR2)_s$. We must sum $|Q|^{d_u^*}$ over elements (P, Q, n) of $\mathcal{C}_+(I)$. Fix $(P_\alpha, Q_\alpha, n_\alpha)$ in $\mathcal{C}_+(\tilde{I})$ such that $I \in Cr(P_\alpha)$. When we consider the partial sum over those Q contained in Q_α , we get, from Proposition 22 in Subsection 8.2.

$$(9.15) \quad \sum_{Q \subset Q_\alpha} |Q|^{d_u^*} \leq C|Q_\alpha|^{d_u^*}.$$

As every Q is contained in such a Q_α , we will have:

$$(9.16) \quad \sum_{\mathcal{C}_+(I)} |Q|^{d_u^*} \leq C \sum_{I \in Cr(P_\alpha)} |Q_\alpha|^{d_u^*}.$$

Summing then over candidates I and using Lemma 10, we get

$$(9.17) \quad \begin{aligned} \sum_I \sum_{\mathcal{C}_+(I)} |Q|^{d_u^*} &\leq C \sum_{\mathcal{C}_+(\tilde{I})} |Q_\alpha|^{d_u^*} \#Cr(P_\alpha) \\ &\leq C|\tilde{I}|^{-\tau d_u^+} \sum_{\mathcal{C}_+(\tilde{I})} |Q_\alpha|^{d_u^*}, \end{aligned}$$

which allow us to conclude as above the induction step for $(SR2)_s$. The argument for $(SR2)_{\tilde{s}}$ is similar and left to the reader. \square

9.3 Classes of Bicritical Rectangles

Now that the size of the critical locus is under control, we must pay attention to the number of bicritical rectangles, which represent the returns of the critical locus to itself under the dynamics.

In order to have an appropriate induction scheme, we need to bound the number of bicritical rectangles according to all width scales and also according to the level of criticality (i.e., the distance to critical locus) of both P and Q . As we will see in the next subsection, the number of bicritical elements experiments a "phase transition" which is crucial for our argument but brings a lot of complications.

Let I be a candidate interval as above, and let I_α, I_ω be parameter intervals such that $I \subset I_\alpha \cap I_\omega$. Let also x be a positive number.

Definition. We denote by $Bi_+(I, I_\alpha, I_\omega; x)$ the set of elements $(P, Q, n) \in \mathcal{R}(I)$ such that P is thin I_α -critical, Q is thin I_ω -critical and $|P| \geq x$ for some $t \in I$.

Similarly, $Bi_-(I, I_\alpha, I_\omega; x)$ is the set of elements $(P, Q, n) \in \mathcal{R}(I)$ such that P is thin I_α -critical, Q is thin I_ω -critical and $|Q| \geq x$ for some $t \in I$.

We denote by $Bi_\pm^{new}(I, I_\alpha, I_\omega; x)$ the set of elements $(P, Q, n) \in Bi_\pm(I, I_\alpha, I_\omega; x)$ that do not belong to $\mathcal{R}(\tilde{I})$, \tilde{I} being as above the parent interval of I .

We will estimate the cardinality of all sets $Bi_\pm(I, I_\alpha, I_\omega; x)$, by induction on the level of the parameter interval I . The easy case is when I is strictly smaller than I_α and I_ω . In this case, no parameter selection is needed. An element of $Bi_+(I, I_\alpha, I_\omega; x)$ is either in $Bi_+^{new}(I, I_\alpha, I_\omega; x)$ or in $Bi_+(\tilde{I}, I_\alpha, I_\omega; x)$, where \tilde{I} is the parent interval of I . We will see that the cardinality of $Bi_+^{new}(I, I_\alpha, I_\omega; x)$ can be estimated from the induction hypothesis and the structure theorem of Subsection 6.7, and this cardinality is much smaller than the cardinality of $Bi_+(\tilde{I}, I_\alpha, I_\omega; x)$ which is controlled by the induction hypothesis.

When I is equal to I_α , but strictly smaller than I_ω (or in the symmetric case $I = I_\omega \subsetneq I_\alpha$), a parameter selection is needed in order to obtain satisfactory estimates. An element of $Bi_+(I, I_\alpha, I_\omega; x)$ is either in $Bi_+^{new}(I, I_\alpha, I_\omega; x)$, whose cardinality can be again estimated from the structure theorem and the induction hypothesis, or it belongs to $Bi_+(\tilde{I}, \tilde{I}, I_\omega; x)$. But in this last case, the vertical rectangle P , which is known to be \tilde{I} -critical, is only I -critical for a small fraction of candidates I which is controlled by Proposition 23 in Subsection 8.3. Averaging, like in Subsection 9.2, allows us to get the required estimate.

By far the most difficult case occurs when $I = I_\alpha = I_\omega$. When x is large, we have a set of estimates, which is taken care of in the same way as for $I = I_\alpha \neq I_\omega$. But when x is small, the parameter selection process is much more subtle and will be explained in Subsection 9.8.

9.4 Number of Bicritical Rectangles

In this subsection, we will state, and comment, the estimates for the cardinalities of the sets Bi_\pm introduced above. The rest of Section 9 will then be devoted to the proof of these estimates under appropriate parameter selection.

At this point, we have to break the symmetry between past and future, P 's and Q 's, stable and unstable direction: the estimates are indeed not symmetric, except when $d_s^0 = d_u^0$, i.e., in the conservative case of area-preserving diffeomorphisms.

We will assume that $d_s^0 \geq d_u^0$ (and $d_s^0 + d_u^0 > 1$). The case $d_u^0 \geq d_s^0$ is obviously symmetric.

For I, I_α, I_ω, x as above we want to have

$$(SR3)_s \quad \# Bi_+(I, I_\alpha, I_\omega; x) \leq CB,$$

with

$$(9.18) \quad B = \max(B_0, B_1),$$

$$(9.19) \quad B_0 = \left(\frac{x}{\varepsilon_0 |P_u|} \right)^{-\rho_0} \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1} \left(\frac{|I_\omega|}{\varepsilon_0} \right)^{\sigma_0},$$

$$(9.20) \quad B_1 = \left(\frac{x}{\varepsilon_0 |P_u|} \right)^{-\rho_1} \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\sigma_1} \left(\min \left(\frac{|I_\alpha|}{\varepsilon_0}, \frac{|I_\omega|}{\varepsilon_0} \right) \right)^{\sigma_0}.$$

Here $|P_u|$ denotes the supremum over I_0 of the width of P_u : the exponents $\rho_0, \rho_1, \sigma_0, \sigma_1$ will be specified more precisely later, but anyway they satisfy

$$(9.21) \quad \rho_0 = d_s^0 + o(1),$$

$$(9.22) \quad \rho_1 = \frac{d_s^0}{d_s^0 + d_u^0} (2d_s^0 + d_u^0 - 1) + o(1),$$

$$(9.23) \quad \sigma_0 = 1 - d_s^0 + o(1),$$

$$(9.24) \quad \sigma_1 = d_s^0 - d_u^0 + o(1).$$

The meaning of the $o(1)$ terms in these formulas is that they become arbitrarily small when $\tau \gg \eta \gg \varepsilon_0$ are small enough.

For the Bi_- sets, we should have:

$$(SR3)_u \quad \# Bi_-(I, I_\alpha, I_\omega; x) \leq CB',$$

with

$$(9.25) \quad B' = \max(B'_0, B'_1),$$

$$(9.26) \quad B'_0 = \left(\frac{x}{\varepsilon_0 |Q_s|} \right)^{-\rho'_0} \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1} \left(\frac{|I_\omega|}{\varepsilon_0} \right)^{\sigma_0},$$

$$(9.27) \quad B'_1 = \left(\frac{x}{\varepsilon_0 |Q_s|} \right)^{-\rho'_1} \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\sigma_1} \left(\min \left(\frac{|I_\alpha|}{\varepsilon_0}, \frac{|I_\omega|}{\varepsilon_0} \right) \right)^{\sigma_0},$$

$$(9.28) \quad \rho'_0 = \frac{d_u^0}{d_s^0} \rho_0 = d_u^0 + o(1),$$

$$(9.29) \quad \rho'_1 = \frac{d_u^0}{d_s^0} \rho_1 = \frac{d_u^0}{d_s^0 + d_u^0} (2d_s^0 + d_u^0 - 1) + o(1).$$

Observe that the formulas (9.22), (9.29) for ρ_1, ρ'_1 are *not* symmetric.

Definition. A parameter interval is *strongly regular* if it satisfies (SR3)_s, (SR3)_u and the eight conditions (SR1), (SR2) of Subsection 8.2.

Remark. *The definition will only be complete when we specify precisely the exponents ρ_0, ρ_1, \dots .*

We now comment on the inequalities above. First, observe that B does not depend on I : this reflects the fact mentioned above that $Bi_+^{new}(I, I_\alpha, I_\omega; x)$ is small compared with $Bi_+(\tilde{I}, I_\alpha, I_\omega; x)$.

From the formulas (9.21), (9.22), we have

$$(9.30) \quad \rho_1 < \rho_0.$$

Set

$$(9.31) \quad x_{cr} := \varepsilon_0 |P_u| \left(\max \left(\frac{|I_\alpha|}{\varepsilon_0}, \frac{|I_\omega|}{\varepsilon_0} \right) \right)^{\frac{\sigma_0}{\rho_0 - \rho_1}}.$$

Then, we have $B = B_0$ for $x \leq x_{cr}$ and $B = B_1$ for $x \geq x_{cr}$: this is the "phase transition" mentioned earlier. We have

$$(9.32) \quad \rho_0 - \rho_1 = \frac{d_s^0(1 - d_s^0)}{d_s^0 + d_u^0} + o(1),$$

$$(9.33) \quad \frac{\sigma_0}{\rho_0 - \rho_1} = \frac{d_s^0 + d_u^0}{d_s^0} + o(1) > 1.$$

For $x = x_{cr}$, we have

$$(9.34) \quad B = B_{cr} := \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1} \left(\frac{|I_\omega|}{\varepsilon_0} \right)^{\sigma_0} \left(\max \left(\frac{|I_\alpha|}{\varepsilon_0}, \frac{|I_\omega|}{\varepsilon_0} \right) \right)^{-\frac{\rho_0 \sigma_0}{\rho_0 - \rho_1}}.$$

Assume $I_\alpha = I_\omega$; we then have

$$(9.35) \quad B_{cr} = \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\sigma_1 + \sigma_0} \frac{\rho_0 - 2\rho_1}{\rho_0 - \rho_1}$$

Here, the exponent satisfies

$$(9.36) \quad \sigma_1 + \sigma_0 \frac{\rho_0 - 2\rho_1}{\rho_0 - \rho_1} = 2 - 2d_s^0 - 2d_u^0 + o(1) < 0.$$

As $|I_\alpha| \leq \varepsilon_0$, we have $B_{cr} \geq 1$. As B is a decreasing function of x , we have $B < 1$ (in which case (SR3)_s means that the Bi_+ set is empty!) iff $B_1 < 1$ which corresponds to

$$(9.37) \quad x > \bar{x} := \varepsilon_0 |P_u| \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\frac{\sigma_0 + \sigma_1}{\rho_1}}.$$

The exponent here satisfies

$$(9.38) \quad \frac{\sigma_0 + \sigma_1}{\rho_1} = \frac{1 - d_u^0}{d_s^0} \frac{d_s^0 + d_u^0}{2d_s^0 + d_u^0 - 1} + o(1)$$

We are finally able to justify the assumption (H4) of our Main Theorem stated in Subsection 1.2! Indeed, with $d_s^0 \geq d_u^0$, it means that

$$(H4) \quad 2(d_s^0)^2 + (d_u^0)^2 + 2d_s^0 d_u^0 < 2d_s^0 + d_u^0$$

and this is exactly what is needed to guarantee that

$$(9.39) \quad \frac{\sigma_0 + \sigma_1}{\rho_1} > 1.$$

We will choose the constant β (related to the regularity property) in order to have

$$(9.40) \quad 1 < \beta < \frac{\sigma_0 + \sigma_1}{\rho_1}$$

and also

$$(9.41) \quad |P_u| < \varepsilon_0^{\beta-1}.$$

Then, from (9.37), we will have that

$$(9.42) \quad \bar{x} < |I_\alpha|^\beta.$$

Summarizing: if $(\text{SR3})_s$ holds, and if $(P, Q, n) \in \mathcal{R}(I)$ is such that P and Q are thin I_α -critical, we must have

$$(9.43) \quad |P| < \bar{x} < |I_\alpha|^\beta$$

for all $t \in I$. This is almost what we need for β -regularity. The full proof is given below.

The discussion for $(\text{SR3})_u$ is similar; the critical threshold is

$$(9.44) \quad x'_{cr} := \varepsilon_0 |Q_s| \left(\max \left(\frac{|I_\alpha|}{\varepsilon_0}, \frac{|I_\omega|}{\varepsilon_0} \right) \right)^{\frac{\sigma_0}{\rho'_0 - \rho'_1}},$$

with

$$(9.45) \quad \rho'_0 - \rho'_1 = \frac{d_u^0(1 - d_s^0)}{d_s^0 + d_u^0} + o(1) = \frac{d_u^0}{d_s^0} (\rho_0 - \rho_1)$$

$$(9.46) \quad \frac{\sigma_0}{\rho'_0 - \rho'_1} = \frac{\sigma_0}{\rho_0 - \rho_1} \frac{d_s^0}{d_u^0} = \frac{d_s^0 + d_u^0}{d_u^0} + o(1) > 1.$$

When $I_\alpha = I_\omega$, we have

$$(9.47) \quad B'_{cr} := \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\sigma_1 + \sigma_0} \frac{\rho'_0 - 2\rho'_1}{\rho'_0 - \rho'_1} = B_{cr} \geq 1.$$

Thus, we have $B' < 1$ if

$$(9.48) \quad x > \bar{x}' := \varepsilon_0 |Q_s| \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\frac{\sigma_0 + \sigma_1}{\rho'_1}}$$

We have here

$$(9.49) \quad \frac{\sigma_0 + \sigma_1}{\rho'_1} = \frac{\sigma_0 + \sigma_1}{\rho_1} \frac{d_s^0}{d_u^0} > \beta$$

and we choose β close enough to one to have

$$(9.50) \quad |Q_s| < \varepsilon_0^{\beta-1}.$$

Then, we are again able to conclude that, if $(\text{SR3})_u$ holds, any $(P, Q, n) \in \mathcal{R}(I)$ such that both P and Q are thin I_α -critical must satisfy, for all $t \in I$

$$(9.51) \quad |Q| < \bar{x}' < |I_\alpha|^\beta.$$

Proposition 26. *If a candidate interval satisfies $(\text{SR3})_s$ and $(\text{SR3})_u$, then it is β -regular. In particular, strong regularity implies regularity.*

Proof. We argue by induction on the level of the parameter interval. For the starting interval I_0 , we use (for the first time!) the assumption (H1) of Subsection 1.2 that the periodic points p_s, p_u do not belong to the same periodic orbit. Then, if $(P, Q, n) \in \mathcal{R}(I_0)$ is such that $P \subset P_s, Q \subset Q_u$, we must have, for all $t \in I_0$

$$(9.52) \quad |P| < \varepsilon_0^\beta, \quad |Q| < \varepsilon_0^\beta$$

if we take β close enough to one. This proves that I_0 is β -regular (independently of $(\text{SR3})_s, (\text{SR3})_u$). Assume that $I \neq I_0$ satisfies $(\text{SR3})_s, (\text{SR3})_u$ and that $(P, Q, n) \in \mathcal{R}(I)$ is I -bicritical. Assume also, for instance, that

$$(9.53) \quad \max_I |Q| \leq \max_I |P|$$

and, by contradiction that

$$(9.54) \quad \max_I |P| \geq |I|^\beta.$$

From the proof of Corollary 9 in Subsection 6.7, we know that $(P, Q, n) \in \mathcal{R}(\tilde{I})$ (\tilde{I} being the parent of I). As (P, Q, n) is \tilde{I} -bicritical, we must have

$$(9.55) \quad \max_I |P| < |\tilde{I}|^\beta.$$

Therefore, P would be thin I -critical; similarly Q would be thin I -critical. But $(\text{SR3})_s$, (see (9.43)), says that such a (P, Q, n) satisfying (9.54) does not exist. \square

Remark. *While there are only eight inequalities $(\text{SR1}), (\text{SR2})$ for each parameter interval I , the inequalities $(\text{SR3})_s, (\text{SR3})_u$ form a family parametrized not only by I , but also by the parameter intervals $I_\alpha \supset I$ and $I_\omega \supset I$ and the real number $x > 0$. Because each inequality, at least when $I = I_\alpha$ or $I = I_\omega$, is only obtained after parameter selection, we will discretize the continuous variable x by considering only the values $x = 2^{-N}$, $N \geq 0$. There is still an infinite number of inequalities, but we will see that they are trivially satisfied if N is large enough.*

9.5 The Starting Interval

Our main purpose in this subsection is to prove the following fact.

Proposition 27. *The starting interval is strongly regular.*

Proof. We have already checked in Subsection 8.2 that I_0 satisfies the eight inequalities (SR1), (SR2). We have therefore to prove that (SR3)_s, (SR3)_u are also satisfied. We clearly have $I_\alpha = I_\omega = I_0$. Then $Bi_+(I_0, I_0, I_0; x)$ is the set of $(P, Q, n) \in \mathcal{R}(I_0)$ such that $P \subset P_s$, $Q \subset Q_u$ and $|P| \geq x$ for some $t \in I_0$. Writing (P, Q, n) as a simple composition

$$(9.56) \quad (P, Q, n) = (P_s, Q_s, n_s) * (P', Q', n') * (P_u, Q_u, n_u)$$

(here we use again assumption (H1)), we, then, use the standard estimate (8.38) to obtain:

$$(9.57) \quad \#Bi_+(I_0, I_0, I_0; x) \leq C \left(\frac{x}{\varepsilon_0 |P_u|} \right)^{-(d_s^0 + C\varepsilon_0)},$$

and similarly

$$(9.58) \quad \#Bi_-(I_0, I_0, I_0; x) \leq C \left(\frac{x}{\varepsilon_0 |Q_s|} \right)^{-(d_u^0 + C\varepsilon_0)}.$$

Therefore, we obtain (SR3)_s and (SR3)_u if we have:

$$(9.59) \quad \rho_0 > d_s^0 + C\varepsilon_0,$$

$$(9.60) \quad \rho'_0 > d_u^0 + C\varepsilon_0,$$

which is compatible with (9.21), (9.28). □

The other part of (SR3)_s, (SR3)_u which can be taken care of right now is the case where x is extremely small.

In this case, we will just forget about the criticality conditions for P and Q and bound the cardinality of $Bi_+(I, I_\alpha, I_\omega; x)$ by the cardinality of the set of $(P, Q, n) \in \mathcal{R}(I)$ for which $|P| \geq x$ for at least some $t \in I$.

This cardinality was estimated in Proposition 22 of Subsection 8.2. Actually, the estimation was given for fixed parameter but it is easy to check that the same proof gives the same estimate as in Proposition 22 for the set we are considering. One obtains

$$(9.61) \quad \#Bi_+(I, I_0, I_0; x) \leq \left(\frac{x}{\varepsilon_0} \right)^{-d_s^*}$$

with $d_s^* = d_s^0 + \varepsilon_0^{\frac{1}{5}} d_s^0$ as above.

We want to have (for all $I_\alpha, I_\omega \supset I$)

$$(9.62) \quad \left(\frac{x}{\varepsilon_0}\right)^{-d_s^*} \leq B_0,$$

which is equivalent to

$$(9.63) \quad \left(\frac{x}{\varepsilon_0}\right)^{\rho_0 - d_s^*} \leq |P_u|^{\rho_0} \left(\frac{|I|}{\varepsilon_0}\right)^{2\sigma_0 + \sigma_1}.$$

This will be satisfied if

$$(9.64) \quad x \leq x_{\min} := |\tilde{I}|^{C(\rho_0 - d_s^*)^{-1}}.$$

So, we need to have

$$(9.65) \quad \rho_0 > d_s^*$$

which is compatible with (9.59), (9.21). The exponent $C(\rho_0 - d_s^*)^{-1}$ will be very large. We have proved

Proposition 28. *The estimate $(SR3)_s$ is satisfied for all candidates I , all $I_\alpha, I_\omega \supset I$, as soon as $x \leq x_{\min}$. A similar statement holds for $(SR3)_u$, with a threshold*

$$(9.66) \quad x'_{\min} := |\tilde{I}|^{C(\rho'_0 - d_u^*)^{-1}}$$

9.6 New Bicritical Rectangles

We consider in this section the set $Bi_+^{new}(I, I_\alpha, I_\omega; x)$ of bicritical rectangles which were not defined over the parent \tilde{I} of I : cf. Subsection 9.3. We assume that $\tilde{I} \subset I_\alpha \cap I_\omega$. We apply to each element (P, Q, n) in this set the structure theorem (Theorem 1) of Subsection 6.7. We obtain an integer $k > 0$, elements $(P_0, Q_0, n_0), \dots, (P_k, Q_k, n_k)$ of $\mathcal{R}(\tilde{I})$ such that

$$(9.67) \quad (P, Q, n) \in (P_0, Q_0, n_0) \square \cdots \square (P_k, Q_k, n_k).$$

Moreover, P_i is \tilde{I} -critical for $0 < i \leq k$ and Q_i is \tilde{I} -critical for $0 \leq i < k$. On the other hand, P_0 is I_α -critical because P is I_α -critical, and Q_k is I_ω -critical because Q is I_ω -critical. Denote by $x_i = 2^{-n_i}$ the largest integral negative power of 2 such that

$$(9.68) \quad |P_i| \geq x_i \quad \text{for some } t \in \tilde{I}.$$

Lemma 11. *We have*

$$\begin{aligned} (P_0, Q_0, n_0) &\in Bi_+(\tilde{I}, I_\alpha, \tilde{I}; x_0), \\ (P_k, Q_k, n_k) &\in Bi_+(\tilde{I}, \tilde{I}, I_\omega; x_k), \\ (P_i, Q_i, n_i) &\in Bi_+(\tilde{I}, \tilde{I}, \tilde{I}; x_i) \end{aligned}$$

for $0 < i < k$.

Remark. This lemma is the reason why we need to consider different levels of criticality for P and Q .

Proof. For $0 < i \leq k$, P_i is \tilde{I} -critical and it is even thin \tilde{I} -critical by Corollary 7 of Subsection 6.7. Similarly, for Q_j when $0 \leq j < k$. Finally, P_0 is thin P_α -critical because P is thin I_α -critical and P_0 is the thinnest rectangle containing P which is defined over \tilde{I} (which is contained in I_α), so the thinnest rectangle containing P and defined over I_α also contains P_0 . Similarly, Q_k is thin I_ω -critical. \square

The widths $|P_i|$ are related to the width of P by Corollary 6 in Subsection 6.7 which gives

$$(9.69) \quad x \leq C^k |I|^{-\frac{k}{2}} \prod_0^k x_i.$$

Let us write

$$\begin{aligned} \#(x_0) &:= \#Bi_+(\tilde{I}, I_\alpha, \tilde{I}; x_0), \\ \#(x_k) &:= \#Bi_+(\tilde{I}, \tilde{I}, I_\omega; x_k), \\ \#(x_i) &:= \#Bi_+(\tilde{I}, \tilde{I}, \tilde{I}; x_i) \text{ for } 0 < i < k. \end{aligned}$$

Then, as each parabolic composition produces two elements, we have

$$(9.70) \quad \#Bi_+^{new}(I, I_\alpha, I_\omega; x) \leq \sum_{k>0} 2^k \sum_{x_0, \dots, x_k} \prod_0^k \#(x_i).$$

The term $\#(x_i)$ is estimated by the induction hypothesis (SR3)_s for \tilde{I} . In view of the threshold (9.31), we divide Bi_+^{new} into two parts. In the first, we put the elements for which every x_i is above the threshold $x_{i,cr}$ given by (9.31). In the second, at least one of the x_i is below $x_{i,cr}$.

Let us consider the first part. Then all $\#(x_i)$ are estimated by B_1 and we have

$$(9.71) \quad \prod_0^k \#(x_i) \leq C^{k+1} \left(\prod_0^k x_i \right)^{-\rho_1} (\varepsilon_0 |P_u|)^{(k+1)\rho_1} \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\sigma_1} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{(k+1)\sigma_0 + k\sigma_1}.$$

In view of (9.69), the right-hand side is bounded by

$$(9.72) \quad \left(\frac{x}{\varepsilon_0 |P_u|} \right)^{-\rho_1} \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\sigma_1} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{\sigma_0} Z^k,$$

with

$$(9.73) \quad Z := \left(C \varepsilon_0 |P_u| |I|^{-\frac{1}{2}} \right)^{\rho_1} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1}.$$

For all $0 \leq i \leq k$, from (9.31) we have

$$(9.74) \quad x_{i,cr} \geq \varepsilon_0 |P_u| \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{\frac{\sigma_0}{\rho_0 - \rho_1}}$$

(with equality when $i \neq 0, k$). Therefore, the number of $(k+1)$ -tuples (n_0, \dots, n_k) such that $2^{-n_i} \geq x_{i,cr}$ for $0 \leq i \leq k$ is at most $(C \log |\tilde{I}|^{-1})^{k+1} \leq (C \log |\tilde{I}|^{-1})^{2k}$. We conclude that the cardinality of the first part of Bi_+^{new} is bounded by

$$(9.75) \quad \left(\frac{x}{\varepsilon_0 |P_u|} \right)^{-\rho_1} \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\sigma_1} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{\sigma_0} \sum_{k>0} Z_1^k,$$

with

$$(9.76) \quad Z_1 := 2(C \log |\tilde{I}|^{-1})^2 Z.$$

In view of (9.73), we have

$$(9.77) \quad Z_1 < \frac{1}{2} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1 - \rho_1} |\tilde{I}|^{\frac{1}{3} \rho_1},$$

with $\sigma_0 + \sigma_1 - \rho_1 > 0$ (cf. (9.39)).

As $Z_1 < \frac{1}{2}$, the bound (9.75) is smaller than

$$(9.78) \quad \left(\frac{x}{\varepsilon_0 |P_u|} \right)^{-\rho_1} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{2\sigma_0 + \sigma_1 - \rho_1} \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\sigma_1} |\tilde{I}|^{\frac{1}{3} \rho_1}.$$

Thus, (9.78) is a bound for the cardinality of the first part of $Bi_+^{new}(I, I_\alpha, I_\omega; x)$. Let us turn to the second part. Let J be the non-empty subset of indices $i \in \{0, \dots, k\}$ for which $x_i < x_{i,cr}$, and write $j = \#J$.

We first estimate the product $\prod_i \#(x_i)$. As $\rho_0 > \rho_1$, we have from (9.69)

$$(9.79) \quad \prod_J x_i^{-\rho_0} \prod_{J^c} x_i^{-\rho_1} \leq \left(C^{-k} |I|^{\frac{k}{2}} x \right)^{-\rho_0}.$$

As we also have $\tilde{I} \subset I_\alpha$, $\tilde{I} \subset I_\omega$, we obtain

$$(9.80) \quad \prod_i \#(x_i) \leq C^{k+1} \left(\frac{x}{\varepsilon_0 |P_u|} \right)^{-\rho_0} \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1} \left(\frac{|I_\omega|}{\varepsilon_0} \right)^{\sigma_0} Y_0^{j-1} Y_1^k,$$

with

$$(9.81) \quad Y_0 = (\varepsilon_0 |P_u|)^{\rho_0 - \rho_1} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{\sigma_0},$$

$$(9.82) \quad Y_1 = \left(C^{-1} |I|^{\frac{1}{2}} \right)^{-\rho_0} (\varepsilon_0 |P_u|)^{\rho_1} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1}.$$

We have

$$(9.83) \quad Y_0 < 1,$$

and we can rewrite Y_1 as

$$(9.84) \quad Y_1 = C^{\rho_0} \left(\frac{\tilde{I}}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1 - \frac{1}{2} \rho_0 (1+\tau)} \varepsilon_0^{\rho_1 - \frac{1}{2} \rho_0 (1+\tau)} |P_u|^{\rho_1}.$$

The exponents satisfy

$$(9.85) \quad \sigma_0 + \sigma_1 - \frac{1}{2} \rho_0(1 + \tau) = 1 - d_u^0 - \frac{1}{2} d_s^0 + o(1),$$

$$(9.86) \quad \rho_1 - \frac{1}{2} \rho_0(1 + \tau) = \frac{d_s^0(3d_s^0 + d_u^0 - 2)}{2(d_s^0 + d_u^0)} + o(1).$$

From (H4), we have $1 - d_u^0 - \frac{1}{2} d_s^0 > 0$; from $d_s^0 + d_u^0 \geq 1$, $d_s^0 \geq d_u^0$, we have $3d_s^0 + d_u^0 - 2 \geq 0$. We can, therefore, find $\sigma_2 > 0$ such that

$$(9.87) \quad Y_1 \leq |\tilde{I}|^{2\sigma_2}.$$

Consider now the number of $(k+1)$ -tuples (n_0, \dots, n_k) for this second part. From (9.69), we have

$$(9.88) \quad \prod_0^k x_i \geq C^{-k} x |I|^{\frac{k}{2}}$$

and we have only to consider the case

$$(9.89) \quad x \geq x_{\min} := |\tilde{I}|^{C(\rho_0 - d_s^*)^{-1}}.$$

From Corollary 7 in Subsection 6.7, we also have, for $0 < i \leq k$,

$$(9.90) \quad x_i < |\tilde{I}|.$$

We conclude that the number of (n_0, \dots, n_k) is smaller than

$$(9.91) \quad \left(C(\rho_0 - d_s^*)^{-1} \log |\tilde{I}|^{-1} \right)^{k+1}.$$

Using $k+1 \leq 2k$ for $k > 0$, we carry this to (9.70) to obtain a bound for the second part of Bi_+^{new} , which is equal to

$$(9.92) \quad \left(\frac{x}{\varepsilon_0 |P_u|} \right)^{-\rho_0} \left(\frac{|I_\alpha|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1} \left(\frac{|I_\omega|}{\varepsilon_0} \right)^{\sigma_0} \sum_{k>0} Y^k$$

with

$$(9.93) \quad Y = 2C(\rho_0 - d_s^*)^{-2} (\log |\tilde{I}|^{-1})^2 Y_1.$$

Here, $(\rho_0 - d_s^*)^{-2}$ is large but independent of ε_0 , which is always assumed to be as small as necessary.

In view of (9.87), we have

$$(9.94) \quad \sum_{k>0} Y^k < |\tilde{I}|^{\sigma_2}.$$

We have now estimated the cardinalities of the two parts of Bi_+^{new} . Taking σ_2 smaller if necessary, we have

$$(9.95) \quad 0 < \sigma_2 \leq \frac{1}{3} \rho_1.$$

In (9.78), as $\sigma_0 + \sigma_1 - \rho_1 > 0$, we have,

$$(9.96) \quad \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{2\sigma_0 + \sigma_1 - \rho_1} \leq \left(\min \frac{|I_\alpha|}{\varepsilon_0}, \frac{|I_\omega|}{\varepsilon_0} \right)^{\sigma_0}.$$

Thus, we have proved

Proposition 29. *Assume that $(SR3)_s$ is satisfied by the parent interval \tilde{I} . Then, for all $x \geq x_{\min}$ and all candidates $I \subset \tilde{I}$, all I_α, I_ω containing \tilde{I} , we have*

$$\#Bi_+^{new}(I, I_\alpha, I_\omega; x) \leq 2B|\tilde{I}|^{\sigma_2},$$

with B given by (9.18).

The presence of the term $|\tilde{I}|^{\sigma_2}$ puts Bi_+^{new} well under control.

Corollary 13. *Assume that $\tilde{I} \subset I_\alpha \cap I_\omega$ and that $(SR3)_s$ holds for $\hat{I} = I_\alpha \cap I_\omega$. Then, $(SR3)_s$ holds for all candidates $I \subset \tilde{I}$.*

Proof. An element $(P, Q, n) \in Bi_+(I, I_\alpha, I_\omega; x)$ either belongs to $Bi_+(\hat{I}, I_\alpha, I_\omega; x)$ or to $Bi_+^{new}(I^*, I_\alpha, I_\omega; x)$ for some interval I^* with $\hat{I} \subsetneq I^* \subset I$. As the series $\sum_{I^*} |I^*|^{\sigma_2}$ is bounded (actually very small), the Corollary follows. \square

The conclusions of Proposition 29 and Corollary 13 also hold for the Bi_- sets; let us review briefly the proof of Proposition 29.

In the formulas for B'_0, B'_1 the exponents σ_0, σ_1 are the same as for B_0, B_1 but we have to replace the exponents ρ_0, ρ_1 by $\rho'_0 = \frac{d_u^0}{d_s^0} \rho_0, \rho'_1 = \frac{d_u^0}{d_s^0} \rho_1$ (with $d_s^0 \geq d_u^0$).

Therefore, we have $\sigma_0 + \sigma_1 - \rho'_1 > 0$ (cf. (9.77), (9.78)), $\sigma_0 + \sigma_1 - \frac{1}{2} \rho'_0 > 0$ (cf. (9.85)), $\rho'_1 - \frac{1}{2} \rho'_0 \geq 0$ (cf. (9.86)). Therefore, Proposition 29 holds for the set $Bi_-^{new}(I, I_\alpha, I_\omega; x)$ with an appropriate choice of σ_2 and Corollary 13 holds for $(SR3)_u$.

The induction step in the case where I is distinct from I_α and I_ω is complete.

9.7 The Case $I = I_\alpha \neq I_\omega$

We consider here the set $Bi_+(I, I, I_\omega; x)$ with $I_\omega \supset \tilde{I}$. We will estimate the size of this set for most candidates $I \subset \tilde{I}$.

Let (P, Q, n) be an element of $Bi_+(I, I, I_\omega; x)$. Either it belongs to $Bi_+^{new}(I, I, I_\omega; x)$ or to $Bi_+(\tilde{I}, \tilde{I}, I_\omega; x)$. In the second case, as P is thin I -critical, there exists $(P^*, Q^*, n^*) \in \mathcal{C}_+(\tilde{I})$ such that $P \subset P^*$ and, moreover, we have $I \in Cr(P^*)$ by definition of this set (Subsection 9.1). We will denote by $Bi_+(P^*)$ the set of $(P, Q, n) \in Bi_+(\tilde{I}, \tilde{I}, I_\omega; x)$ such that $P \subset P^*$. We, thus, have

$$(9.97) \quad \#Bi_+(I, I, I_\omega; x) \leq \#Bi_+^{new}(I, I, I_\omega; x) + \sum_{\substack{(P^*, Q^*, n^*) \in \mathcal{C}_+(\tilde{I}) \\ I \in Cr(P^*)}} \#Bi_+(P^*)$$

As the sets $Bi_+(P^*)$ are disjoint, we have

$$(9.98) \quad \sum_{\mathcal{C}_+(\tilde{I})} \#Bi_+(P^*) \leq \#Bi_+(\tilde{I}, \tilde{I}, I_\omega; x).$$

We also have

$$(9.99) \quad Bi_+^{new}(I, I, I_\omega; x) \subset Bi_+^{new}(\tilde{I}, \tilde{I}, I_\omega; x).$$

We assume that $(SR3)_s$ holds for parameter intervals containing \tilde{I} . Write $\tilde{B} = \max(\tilde{B}_0, \tilde{B}_1)$ for the values of (9.18)–(9.20) with $I_\alpha = \tilde{I}$. Thus, we have

$$(9.100) \quad \#Bi_+(\tilde{I}, \tilde{I}, I_\omega; x) \leq C\tilde{B},$$

and, from Proposition 29,

$$(9.101) \quad \#Bi_+^{new}(I, \tilde{I}, I_\omega; x) \leq 2\tilde{B}|\tilde{I}|^{\sigma_2}.$$

We now sum over candidates $I \subset \tilde{I}$ the estimate (9.97). We obtain

$$(9.102) \quad \sum_{I \subset \tilde{I}} \#Bi_+(I, I, I_\omega; x) \leq 2\tilde{B}|\tilde{I}|^{\sigma_2}|\tilde{I}|^{-\tau} + \sum_{\mathcal{C}_+(\tilde{I})} \#Bi_+(P^*)\#Cr(P^*),$$

and then, using Lemma 10 and (9.100), the same sum is bounded by

$$(9.103) \quad 2\tilde{B}|\tilde{I}|^{\sigma_2-\tau} + C'\tilde{B}|\tilde{I}|^{-\tau d_u^+} \leq C''\tilde{B}|\tilde{I}|^{-\tau d_u^+}.$$

Write $B = \max(B_0, B_1)$ for the values of (9.18)–(9.20) with $I_\alpha = I$. The number of candidates I such that $\#Bi_+(I, I, I_\omega; x) > B$ is at most

$$(9.104) \quad C''\tilde{B}B^{-1}|\tilde{I}|^{-\tau d_u^+}.$$

From formulas (9.18)–(9.20), we see that

$$(9.105) \quad \tilde{B}B^{-1} = |\tilde{I}|^{-\tau(\sigma_0+\sigma_1)}.$$

The exponents σ_0, σ_1 will be chosen in order to have

$$(9.106) \quad \sigma_0 + \sigma_1 + d_u^+ < 1 - 2\tau,$$

which is compatible with (9.23), (9.24). Then, we obtain that $(SR3)_s$ holds except for a proportion of candidates no greater than $C|\tilde{I}|^{2\tau^2}$. This assertion has been proved for a fixed value of x and a fixed parameter interval I_ω . But (in view of the constant C in $(SR3)_s$), it is sufficient to prove $(SR3)_s$ when $x = 2^{-n}$, n a nonnegative integer, and $x \geq x_{\min}$; and the number of intervals $I_\omega \supset \tilde{I}$ is at most

$$(9.107) \quad C\tau^{-1} \log \left(\frac{\log |\tilde{I}|}{\log \varepsilon_0} \right).$$

Thus, the total number of cases that we have to consider is much smaller than $|\tilde{I}|^{-\tau^2}$. We obtain

Proposition 30. *Assume that $(SR3)_s$ holds for parameter intervals containing \tilde{I} . Then, $(SR3)_s$ holds for $Bi_+(I, I, I_\omega; x)$, for all $I_\omega \supset \tilde{I}$ and all x , except perhaps for a proportion of candidates no greater than $|\tilde{I}|^{-\tau^2}$.*

Proposition 30 is also valid exchanging I_α and I_ω . There are also two similar statements involving $(SR3)_u$ and Bi_- . We review briefly the slight difference in the proof.

For $Bi_-(I, I, I_\omega; x)$ (with $I_\omega \supset \tilde{I}$), we proceed exactly as for $Bi_+(I, I, I_\omega; x)$. For $Bi_+(I, I_\alpha, I; x)$, we use $\mathcal{C}_-(\tilde{I})$ instead of $\mathcal{C}_+(\tilde{I})$ to subdivide the set of bicritical elements.

We now have, by the dual version of Lemma 10

$$(9.108) \quad \#Cr(Q^*) \leq C|\tilde{I}|^{-\tau d_s^+}.$$

On the other hand, from formulas (9.18)–(9.20), we must replace (9.105) by

$$(9.109) \quad \tilde{B}B^{-1} = |\tilde{I}|^{-\tau\sigma_0}.$$

Thus, we will choose σ_0 in order to have

$$(9.110) \quad \sigma_0 + d_s^+ < 1 - 2\tau.$$

Obviously, this is compatible with (9.23), (9.24), (9.106). The end of the argument is the same as before.

This completes the proof of the induction step except for the case $I = I_\alpha = I_\omega$.

In this last case, we can actually apply the same argument as above to complete the induction step when x is large. When $I_\alpha = I_\omega$ is equal to I or \tilde{I} , indeed we have

$$(9.111) \quad B_0(x) = \left(\frac{x}{\varepsilon_0|P_u|}\right)^{-\rho_0} \left(\frac{|I|}{\varepsilon_0}\right)^{2\sigma_0+\sigma_1},$$

$$(9.112) \quad B_1(x) = \left(\frac{x}{\varepsilon_0|P_u|}\right)^{-\rho_1} \left(\frac{|I|}{\varepsilon_0}\right)^{\sigma_0+\sigma_1},$$

$$(9.113) \quad \tilde{B}_0(x) = \left(\frac{x}{\varepsilon_0|P_u|}\right)^{-\rho_0} \left(\frac{|\tilde{I}|}{\varepsilon_0}\right)^{2\sigma_0+\sigma_1},$$

$$(9.114) \quad \tilde{B}_1(x) = \left(\frac{x}{\varepsilon_0|P_u|}\right)^{-\rho_1} \left(\frac{|\tilde{I}|}{\varepsilon_0}\right)^{\sigma_0+\sigma_1}.$$

For $x \geq \tilde{x}_{cr}$, we, therefore, have

$$(9.115) \quad \tilde{B}(x)B(x)^{-1} = |\tilde{I}|^{-\tau(\sigma_0+\sigma_1)}$$

and the same argument as before applies.

For $x \leq x_{cr}$, we have

$$(9.116) \quad \tilde{B}(x)B(x)^{-1} = |\tilde{I}|^{-\tau(2\sigma_0+\sigma_1)},$$

with $2\sigma_0 + \sigma_1 = 2 - d_s^0 - d_u^0 + o(1)$: this case will be the object of the rest of Section 9.

In the intermediate range $x_{cr} \leq x \leq \tilde{x}_{cr}$, we write $x = \tilde{x}_{cr}y$, where we have, in view of (9.31),

$$(9.117) \quad 1 \geq y \geq |\tilde{I}|^{\frac{\tau\sigma_0}{\rho_0 - \rho_1}}.$$

We now have in this range:

$$(9.118) \quad \begin{aligned} \tilde{B}(x)B(x)^{-1} &= \tilde{B}_0(x)B_1(x)^{-1} \\ &= y^{\rho_1 - \rho_0} |\tilde{I}|^{-\tau(\sigma_0 + \sigma_1)}. \end{aligned}$$

Therefore, the proportion of bad candidates will not be greater than $C|\tilde{I}|^{2\tau^2}$, as long as we have

$$(9.119) \quad y^{\rho_0 - \rho_1} \geq |\tilde{I}|^{\tau(1 - \sigma_0 - \sigma_1 - d_u^+ - 2\tau)}.$$

We recall that $\rho_0 - \rho_1 > 0$ and that $1 - \sigma_0 - \sigma_1 - d_u^+ - 2\tau$ is positive according to (9.106), but it has to be small according to (9.23), (9.24). We state

Proposition 31. *Assume that $(SR3)_s$ holds for all parameter intervals containing \tilde{I} . Then, it holds for $Bi_+(I, I, I; x)$ in the range*

$$(9.120) \quad x \geq \tilde{x}_{cr} |\tilde{I}|^{\frac{\tau}{\rho_0 - \rho_1} (1 - \sigma_0 - \sigma_1 - d_u^+ - 2\tau)}$$

except for a proportion of candidates no greater than $|\tilde{I}|^{\tau^2}$.

For $Bi_-(I, I, I; x)$, the corresponding range is

$$(9.121) \quad x \geq \tilde{x}'_{cr} |\tilde{I}|^{\frac{\tau}{\rho_0' - \rho_1'} (1 - \sigma_0 - \sigma_1 - d_u^+ - 2\tau)}.$$

9.8 The Case $I = I_\alpha = I_\omega$: General Overview

We explain in this subsection the strategy which will be pursued in the case $I = I_\alpha = I_\omega$, x small.

The exponent $2\sigma_0 + \sigma_1$ in (9.116) means that we have take into account the criticality of both P and Q in the selection process.

We start as before. An element (P, Q, n) in $Bi_+(I, I, I; x)$ is either in $Bi_+^{new}(I, I, I; x)$ or in $Bi_+(\tilde{I}, \tilde{I}, \tilde{I}; x)$. The cardinality of the first set is bounded as before by

$$(9.122) \quad \#Bi_+^{new}(I, I, I; x) \leq 2\tilde{B}|\tilde{I}|^{\sigma_2}.$$

On the other hand, let $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(\tilde{I})$ and $(P_\omega, Q_\omega, n_\omega) \in \mathcal{C}_-(\tilde{I})$. We denote by $Bi_+(P_\alpha, Q_\omega)$ the set of $(P, Q, n) \in Bi_+(\tilde{I}, \tilde{I}, \tilde{I}; x)$ such that $P \subset P_\alpha$ and $Q \subset Q_\omega$. These subsets are disjoint and their union contain the elements (P, Q, n) of $Bi_+(I, I, I; x)$ which belong to $\mathcal{R}(\tilde{I})$ because P and

Q are thin I -critical. Moreover, only those P_α, Q_ω such that $I \in Cr(P_\alpha) \cap Cr(Q_\omega)$ will actually appear.

Thus, in analogy with (9.97), we have

$$(9.123) \quad \#Bi_+(I, I, I; x) \leq 2\tilde{B}|\tilde{I}|^{\sigma_2} + \sum \#Bi_+(P_\alpha, Q_\omega),$$

where the sum is over pairs $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(\tilde{I}), (P_\omega, Q_\omega, n_\omega) \in \mathcal{C}_-(\tilde{I})$ such that $I \subset Cr(P_\alpha)$ and $I \subset Cr(Q_\omega)$.

Unfortunately, we are not able to estimate the size of the intersection $Cr(P_\alpha) \cap Cr(Q_\omega)$ directly in a satisfactory way: while these two sets seem "independent" in the naive sense, it is another matter to translate this intuition in a quantitative estimate in the style of Lemma 10.

Instead of this approach, we will use some degree of independence, when x is small, of the variables P_α and Q_ω in $Bi_+(P_\alpha, Q_\omega)$. To explain the technique, consider first the unrealistic model case where we would have

$$(9.124) \quad \#Bi_+(P_\alpha, Q_\omega) = b_+(P_\alpha)b_-(Q_\omega),$$

for some functions b_+, b_- on $\mathcal{C}_+(\tilde{I}), \mathcal{C}_-(\tilde{I})$, respectively.

The formula (9.123) now gives

$$(9.125) \quad \#Bi_+(I, I, I; x) \leq 2\tilde{B}|\tilde{I}|^{\sigma_2} + \phi_+(I)\phi_-(I),$$

with

$$(9.126) \quad \phi_+(I) = \sum_{I \in Cr(P_\alpha)} b_+(P_\alpha),$$

$$(9.127) \quad \phi_-(I) = \sum_{I \in Cr(Q_\omega)} b_-(Q_\omega).$$

We now average *separately* ϕ_+ and ϕ_- . We obtain

$$(9.128) \quad \sum_I \phi_+(I) \leq \left(\max_{\tilde{I}} \#Cr(P_\alpha) \right) \sum_{\mathcal{C}_+(\tilde{I})} b_+(P_\alpha),$$

$$(9.129) \quad \sum_I \phi_-(I) \leq \left(\max_{\tilde{I}} \#Cr(Q_\omega) \right) \sum_{\mathcal{C}_-(\tilde{I})} b_-(Q_\omega),$$

where $Cr(P_\alpha)$ and $Cr(Q_\omega)$ are estimated by Lemma 10 and we have

$$(9.130) \quad \sum b_+(P_\alpha) \sum b_-(Q_\omega) = \sum \#Bi_+(P_\alpha, Q_\omega) \leq \tilde{B}.$$

It is then sufficient to eliminate candidates for which either ϕ_+ or ϕ_- is much above its average value to be able to conclude the proof.

As (9.124) does not hold, we will find an appropriate substitute as follows.

We will subdivide each class $Bi_+(P_\alpha, Q_\omega)$ into subclasses $Bi_+(P_\alpha, Q_\omega, \ell)$; the index ℓ runs through a finite large set L dependent on I and x but independent on P_α and Q_ω . Moreover, we will have functions $b_+(P_\alpha, \ell)$, $b_-(Q_\omega, \ell)$ on $\mathcal{C}_+(\tilde{I}) \times L$, $\mathcal{C}_-(\tilde{I}) \times L$, respectively, such that,

$$(9.131) \quad \#Bi_+(P_\alpha, Q_\omega, \ell) \leq b_+(P_\alpha, \ell), b_-(Q_\omega, \ell).$$

We then set, for each $\ell \in L$:

$$(9.132) \quad \phi_{+,\ell}(I) = \sum_{I \in Cr(P_\alpha)} b_+(P_\alpha, \ell),$$

$$(9.133) \quad \phi_{-,\ell}(I) = \sum_{I \in Cr(Q_\omega)} b_-(Q_\omega, \ell).$$

We average each of these functions to get, in view of Lemma 10,

$$(9.134) \quad \sum_I \phi_{+,\ell}(I) \leq C|\tilde{I}|^{-\tau d_u^+} b_+(\ell),$$

$$(9.135) \quad \sum_I \phi_{-,\ell}(I) \leq C|\tilde{I}|^{-\tau d_s^+} b_-(\ell),$$

with

$$(9.136) \quad b_+(\ell) = \sum_{\mathcal{C}_+(\tilde{I})} b_+(P_\alpha, \ell),$$

$$(9.137) \quad b_-(\ell) = \sum_{\mathcal{C}_-(\tilde{I})} b_-(Q_\omega, \ell).$$

For each ℓ , we will have

$$(9.138) \quad \phi_{+,\ell}(I) \leq |\tilde{I}|^{\tau(1-d_u^+-3\tau)} b_+(\ell)$$

$$(9.139) \quad \phi_{-,\ell}(I) \leq |\tilde{I}|^{\tau(1-d_s^+-3\tau)} b_-(\ell)$$

except for a proportion of candidates not greater than $C|\tilde{I}|^{3\tau^2}$. Set

$$(9.140) \quad \widehat{B} = \sum_L b_+(\ell)b_-(\ell).$$

Because we need to eliminate candidates for each ℓ , L should not be too large. We will have, see Proposition 33 in Subsection 9.10, that

$$(9.141) \quad \#L \leq C|\tilde{I}|^{-\tau^2}.$$

Taking into account that we must eliminate candidates for each $x = 2^{-n} \geq x_{\min}$, the total proportion of the failed candidates is at most $|\tilde{I}|^{\tau^2}$. On the other hand, for the surviving candidates, the discussion above gives

$$(9.142) \quad \begin{aligned} \sum_{I \in Cr(P_\alpha) \cap Cr(Q_\omega)} \#Bi_+(P_\alpha, Q_\omega) &\leq \sum_L \sum_{I \in Cr(P_\alpha)} \sum_{I \in Cr(Q_\omega)} b_+(P_\alpha, \ell) b_-(Q_\omega, \ell) \\ &= \sum_L \phi_{+,\ell}(I) \phi_{-,\ell}(I) \leq |\tilde{I}|^{\tau(2-d_s^+-d_u^+-6\tau)} \widehat{B} \end{aligned}$$

$$(9.143) \quad \#Bi_+(I, I, I; x) \leq 2\tilde{B}|\tilde{I}|^{\sigma_2} + \widehat{B}|\tilde{I}|^{\tau(2-d_s^+ - d_u^+ - 6\tau)}.$$

We definitely have $\tilde{B}|\tilde{I}|^{\sigma_2} \ll B$. In order to obtain (SR3)_s, we need

$$(9.144) \quad \widehat{B}|\tilde{I}|^{\tau(2-d_s^+ - d_u^+ - 6\tau)} \leq CB.$$

As $2 - d_s^+ - d_u^+ - 6\tau = 2\sigma_0 + \sigma_1 + o(1)$, we see from (9.116) that \widehat{B} cannot be much larger than \tilde{B} .

In the next three subsections, we will

- define precisely L and the subclasses $Bi_+(P_\alpha, Q_\omega, \ell)$;
- bound the cardinality of L (to obtain (9.141));
- obtain an appropriate estimate for \widehat{B} (cf. (9.144)).

9.9 Subclasses of Bicritical Elements

9.9.1 Bound elements.

Definition. Let $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(\tilde{I})$, $(P_\omega, Q_\omega, n_\omega) \in \mathcal{C}_-(\tilde{I})$. An element $(P, Q, n) \in Bi_+(P_\alpha, Q_\omega)$ is *bound* if $n \leq n_\alpha + n_\omega$. Otherwise, we say that (P, Q, n) is *free*. We will denote by $Bi_+(P_\alpha, Q_\omega, \diamond)$ the subset of bound elements of $Bi_+(P_\alpha, Q_\omega)$.

Thus, \diamond is an element of L . On the other hand, free elements will correspond to many elements of L . Recall that we have $x \leq \tilde{x}_{cr}$. When $x \ll \tilde{x}_{cr}$, most elements are free. When $x \gg \tilde{x}_{cr}$, on the opposite, most elements are bound.

Proposition 32. *For any $(P_\alpha, Q_\alpha, n_\alpha)$, $(P_\omega, Q_\omega, n_\omega) \in \mathcal{R}(\tilde{I})$, and any $n \leq n_\alpha + n_\omega$, there is at most one element $(P, Q, n) \in \mathcal{R}(\tilde{I})$ of length n such that $P \subset P_\alpha$, $Q \subset Q_\omega$.*

Proof. We argue by induction on the level of the parameter interval.

When \tilde{I} is the starting interval I_0 , no parabolic composition is involved and the result follows from usual symbolic dynamics: as $n \leq n_\alpha + n_\omega$, the word associated to a bound element is determined by its initial and final parts.

Assume that the result holds for parameter intervals strictly larger than \tilde{I} . Denote by \tilde{I}_1 the parent interval of \tilde{I} .

Assume first that both $(P_\alpha, Q_\alpha, n_\alpha)$ and $(P_\omega, Q_\omega, n_\omega)$ belong to $\mathcal{R}(\tilde{I}_1)$. We claim that any bound element also belongs to $\mathcal{R}(\tilde{I}_1)$, which allow us to conclude the proof by the induction hypothesis. Indeed, if (P, Q, n) satisfies $P \subset P_\alpha$, $Q \subset Q_\omega$ and does not belong to $\mathcal{R}(\tilde{I}_1)$, we apply the structure

theorem in Subsection 6.7: it gives elements $(P_0, Q_0, n_0), (P_k, Q_k, n_k) \in \mathcal{R}(\tilde{I}_1)$ such that P_0 is the thinnest rectangle containing P defined over \tilde{I}_1 , Q_k is the thinnest rectangle containing Q defined over \tilde{I}_1 , and $n \geq n_0 + n_k + N_0$. Therefore, $n_0 \geq n_\alpha$, $n_k \geq n_\omega$ and $n > n_\alpha + n_\omega$.

We now consider the case when, for instance, $(P_\alpha, Q_\alpha, n_\alpha)$ does not belong to $\mathcal{R}(\tilde{I}_1)$. We now apply the structure theorem of Subsection 6.7 to $(P_\alpha, Q_\alpha, n_\alpha)$ and also to an element $(P, Q, n) \in \mathcal{R}(\tilde{I})$ with $P \subset P_\alpha$, $Q \subset Q_\omega$, $n \leq n_\alpha + n_\omega$. We obtain integers $0 < j \leq k$, elements $(P_i, Q_i, n_i) \in \mathcal{R}(\tilde{I}_1)$ for $0 \leq i \leq k$ such that

$$(9.145) \quad (P, Q, n) \in (P_0, Q_0, n_0) \square \cdots \square (P_k, Q_k, n_k)$$

and also $(\tilde{P}_j, \tilde{Q}_j, \tilde{n}_j) \in \mathcal{R}(\tilde{I}_1)$ such that

$$(9.146) \quad (P_\alpha, Q_\alpha, n_\alpha) \in (P_0, Q_0, n_0) \square \cdots \square (P_{j-1}, Q_{j-1}, n_{j-1}) \square (\tilde{P}_j, \tilde{Q}_j, \tilde{n}_j).$$

There also exists m with $0 \leq m \leq j$ and $(\tilde{P}'_m, \tilde{Q}'_m, \tilde{n}'_m) \in \mathcal{R}(\tilde{I}_1)$ such that

- either $m = j = k$, $(\tilde{P}'_m, \tilde{Q}'_m, \tilde{n}'_m) = (P_\omega, Q_\omega, n_\omega)$;
- or $m < k$ and we have,

$$(9.147) \quad (P_\omega, Q_\omega, n_\omega) \in (\tilde{P}'_m, \tilde{Q}'_m, \tilde{n}'_m) \square \cdots \square (P_k, Q_k, n_k).$$

If $m < j$, the sequence (P_i, Q_i, n_i) and the choice of the result (out of two possibilities) in each parabolic composition are completely determined by P_α and Q_ω : the assertion of the proposition follows.

When $m = j$, the sequence (P_i, Q_i, n_i) for $i \neq m = j$ is determined by P_α , Q_ω ; but we also have $P_j \subset \tilde{P}_j$, $Q_j = Q_m \subset \tilde{Q}_m$ and $n_j \leq \tilde{n}_j + \tilde{n}'_m$, so by the induction hypothesis (P_j, Q_j, n_j) is also determined by P_α , Q_ω . Again, the choices of the results in the parabolic compositions are also determined by P_α , Q_ω . The proof of the proposition is complete. \square

Recall that, by Proposition 13 in Subsection 7.1, we have

$$(9.148) \quad |P| \leq C \exp(-n^\gamma)$$

for any $(P, Q, n) \in \mathcal{R}(\tilde{I})$, any $t \in \tilde{I}$, with $\gamma = \log \frac{3}{2} / \log 2$. Any bound element in $Bi_+(P_\alpha, Q_\omega, \diamond)$ must satisfy $|P| \geq x$ for some $t \in \tilde{I}$, and we therefore have

$$(9.149) \quad n \leq \left(\log \frac{C}{x} \right)^{\frac{1}{\gamma}}$$

For $x \geq x_{\min}$, we have

$$(9.150) \quad \left(\log \frac{C}{x} \right)^{\frac{1}{\gamma}} \leq (\log |\tilde{I}|)^2.$$

We thus shall define

$$(9.151) \quad b_+(P_\alpha, \diamond) = b_-(Q_\omega, \diamond) = \log |\tilde{I}|^{-1},$$

and we will indeed have, from Proposition 32,

$$(9.152) \quad \#Bi_+(P_\alpha, Q_\omega, \diamond) \leq b_+(P_\alpha, \diamond) b_-(Q_\omega, \diamond).$$

9.9.2 Decomposition of a free element.

Let $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(\tilde{I})$, $(P_\omega, Q_\omega, n_\omega) \in \mathcal{C}_-(\tilde{I})$ and let $(P, Q, n) \in Bi_+(P_\alpha, Q_\omega)$ be a *free* element. We will analyze with respect to the structure theorem of Subsection 6.7 the way in which (P, Q, n) , $(P_\alpha, Q_\alpha, n_\alpha)$, $(P_\omega, Q_\omega, n_\omega)$ have been created. This will allow us in the sequel to define various subclasses of free elements.

Denote by \hat{I}_0 the largest parameter interval such that $(P, Q, n) \in \mathcal{R}(\hat{I}_0)$. Elements (P, Q, n) for which \hat{I}_0 is the starting interval I_0 are said to have depth 0. They form a first subclass of $Bi_+(P_\alpha, Q_\omega)$ denoted by $Bi_+(P_\alpha, Q_\omega, 0)$.

We now assume that $\hat{I}_0 \neq I_0$ and denote by \tilde{I}_0 the parent interval of \hat{I}_0 . We apply the structure theorem of Subsection 6.7. We obtain an integer $k > 0$, elements $(P_0, Q_0, n_0), \dots, (P_k, Q_k, n_k)$ in $\mathcal{R}(\tilde{I}_0)$ such that

$$(9.153) \quad (P, Q, n) \in (P_0, Q_0, n_0) \square \cdots \square (P_k, Q_k, n_k).$$

As in the proof of Proposition 32, we find $0 \leq j \leq k$ and $(\tilde{P}_j, \tilde{Q}_j, \tilde{n}_j) \in \mathcal{R}(\tilde{I}_0)$ such that either $j = 0$, $(\tilde{P}_j, \tilde{Q}_j, \tilde{n}_j) = (P_\alpha, Q_\alpha, n_\alpha)$ (if $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{R}(\tilde{I}_0)$) or $j > 0$ and

$$(9.154) \quad (P_\alpha, Q_\alpha, n_\alpha) \in (P_0, Q_0, n_0) \square \cdots \square (\tilde{P}_j, \tilde{Q}_j, \tilde{n}_j).$$

Similarly, we find $0 \leq m \leq k$ and $(\tilde{P}'_m, \tilde{Q}'_m, \tilde{n}'_m) \in \mathcal{R}(\tilde{I}_0)$ such that either $m = k$, $(\tilde{P}'_m, \tilde{Q}'_m, \tilde{n}'_m) = (P_\omega, Q_\omega, n_\omega)$ or $m < k$ and

$$(9.155) \quad (P_\omega, Q_\omega, n_\omega) \in (\tilde{P}'_m, \tilde{Q}'_m, \tilde{n}'_m) \square \cdots \square (P_k, Q_k, n_k).$$

We also must have $P_j \subset \tilde{P}_j$, $Q_m \subset \tilde{Q}'_m$. Moreover, as (P, Q, n) is free, we must have $j \leq m$ and, when $j = m$, we must also have $n_j = n_m > \tilde{n}_j + \tilde{n}'_m$.

We say that (P, Q, n) is *fully decomposed* if one has here $j < m$ or $j = m$ and $(P_j, Q_j, n_j) \in \mathcal{R}(I_0)$. Such elements are said to have depth one.

Assume that (P, Q, n) is not fully decomposed. Then, we have $j = m$, $P_j \subset \tilde{P}_j$, $Q_j \subset \tilde{Q}'_j$ and the largest parameter interval \hat{I}_1 for which $(P_j, Q_j, n_j) \in \mathcal{R}(\hat{I}_1)$ is not the starting interval I_0 . We denote by \tilde{I}_1 the parent interval. We rewrite

$$(9.156) \quad \begin{aligned} (P^1, Q^1, n^1) &:= (P_j, Q_j, n_j), \\ (P_\alpha^1, Q_\alpha^1, n_\alpha^1) &:= (\tilde{P}_j, \tilde{Q}_j, \tilde{n}_j), \\ (P_\omega^1, Q_\omega^1, n_\omega^1) &:= (\tilde{P}'_j, \tilde{Q}'_j, \tilde{n}'_j), \end{aligned}$$

and proceed with these elements as we did with (P, Q, n) , $(P_\alpha, Q_\alpha, n_\alpha)$, $(P_\omega, Q_\omega, n_\omega)$: we will find integers $0 \leq j_1 \leq m_1 \leq k_1$ (with $k_1 > 0$), elements (P_i^1, Q_i^1, n_i^1) for $0 \leq i \leq k_1$ and also $(P_\alpha^2, Q_\alpha^2, n_\alpha^2)$, $(P_\omega^2, Q_\omega^2, n_\omega^2)$, all in $\mathcal{R}(\tilde{I}_1)$, such that

$$(9.157) \quad \begin{aligned} (P^1, Q^1, n^1) &\in (P_0^1, Q_0^1, n_0^1) \square \cdots \square (P_{k_1}^1, Q_{k_1}^1, n_{k_1}^1), \\ (P_\alpha^1, Q_\alpha^1, n_\alpha^1) &\in (P_0^1, Q_0^1, n_0^1) \square \cdots \square (P_{j_1-1}^1, Q_{j_1-1}^1, n_{j_1-1}^1) \square (P_\alpha^2, Q_\alpha^2, n_\alpha^2), \\ (P_\omega^1, Q_\omega^1, n_\omega^1) &\in (P_\omega^2, Q_\omega^2, n_\omega^2) \square (P_{m_1+1}^1, Q_{m_1+1}^1, n_{m_1+1}^1) \square \cdots \square (P_{k_1}^1, Q_{k_1}^1, n_{k_1}^1). \end{aligned}$$

Again, we say that (P^1, Q^1, n^1) is fully decomposed if either $j_1 < m_1$ or $j_1 = m_1$ and $(P_{j_1}^1, Q_{j_1}^1, n_{j_1}^1)$ is defined over the starting interval I_0 ; otherwise we set

$$(9.158) \quad (P^2, Q^2, n^2) := (P_{j_1}^1, Q_{j_1}^1, n_{j_1}^1),$$

and we go on. The sequence of parameter intervals $\hat{I}_0 \subset \hat{I}_1 \subset \cdots$ is strictly increasing and therefore the process will stop. We define inductively the *depth* of (P, Q, n) to be the depth of (P^1, Q^1, n^1) plus one.

9.9.3 Size of the subclass of depth 0

We will define in this subsection $b_+(P_\alpha, 0)$, $b_-(Q_\omega, 0)$ in order to have

$$(9.159) \quad \#Bi_+(P_\alpha, Q_\omega, 0) \leq b_+(P_\alpha, 0) b_-(Q_\omega, 0).$$

Let $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(\tilde{I})$, $(P_\omega, Q_\omega, n_\omega) \in \mathcal{C}_-(\tilde{I})$ and let $(P, Q, n) \in Bi_+(P_\alpha, Q_\omega, 0)$. Then (P, Q, n) is obtained by a simple composition

$$(9.160) \quad (P, Q, n) = (P_\alpha, Q_\alpha, n_\alpha) * (P', Q', n') * (P_\omega, Q_\omega, n_\omega).$$

We have here, for all $t \in \tilde{I}$

$$(9.161) \quad |P| \leq C |P_\alpha| |P'| |P_\omega|,$$

$$(9.162) \quad |P_\alpha| \leq |\tilde{I}|^{1+\tau} \quad (\text{cf. Subsection 9.1}),$$

and also, for some $t_0 \in \tilde{I}$

$$(9.163) \quad |P| \geq x.$$

This gives, for this value t_0 :

$$(9.164) \quad |P'| \geq C^{-1} |\tilde{I}|^{-(1+\tau)} \left(\max_{\tilde{I}} |P_\omega| \right)^{-1} x.$$

We observe that, as $(P_\omega, Q_\omega, n_\omega)$ belongs to $\mathcal{R}(I_0)$ and $|Q_\omega|$ is of the order of $|\tilde{I}|^{1+\tau}$, we have

$$(9.165) \quad \max_{\tilde{I}} |P_\omega| \leq C \min_{\tilde{I}} |P_\omega|.$$

From the estimate (8.38) in Subsection 8.2, we can thus define, as $d_s^0 + C\varepsilon_0 < d_s^*$,

$$(9.166) \quad b_+(P_\alpha, 0) = \begin{cases} (C|\tilde{I}|^{1+\tau}x^{-1})^{d_s^*} & \text{if } (P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{R}(I_0), \\ 0 & \text{otherwise,} \end{cases}$$

$$(9.167) \quad b_-(Q_\omega, 0) = \begin{cases} (\min_{\tilde{I}} |P_\omega|)^{d_s^*} & \text{if } (P_\omega, Q_\omega, n_\omega) \in \mathcal{R}(I_0), \\ 0 & \end{cases}$$

Then, (9.159) is satisfied.

9.9.4 Subclasses of higher depth

Let $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(\tilde{I})$, $(P_\omega, Q_\omega, n_\omega) \in \mathcal{C}_-(\tilde{I})$ and let $(P, Q, n) \in Bi_+(P_\alpha, Q_\omega)$ be an element of depth $s > 0$.

Let us first restate and extend somewhat the notations and the setting of 9.9.2. We set

$$(9.168) \quad \begin{aligned} (P^0, Q^0, n^0) &:= (P, Q, n), \\ (P_\alpha^0, Q_\alpha^0, n_\alpha^0) &:= (P_\alpha, Q_\alpha, n_\alpha), \\ (P_\omega^0, Q_\omega^0, n_\omega^0) &:= (P_\omega, Q_\omega, n_\omega). \end{aligned}$$

We have

- a strictly increasing sequence of parameter intervals

$$(9.169) \quad \hat{I}_0 \subset \hat{I}_1 \subset \cdots \subset \hat{I}_{s-1} \subset \hat{I}_s = I_0$$

with $\tilde{I} \subset \hat{I}_0$; we denote by \tilde{I}_r the parent interval of \hat{I}_r for $0 \leq r < s$;

- a sequence (P^r, Q^r, n^r) , $0 \leq r \leq s$ such that (P^r, Q^r, n^r) belongs to $\mathcal{R}(\hat{I}_r)$ but not to $\mathcal{R}(\tilde{I}_r)$ for $r < s$; also $(P^s, Q^s, n^s) \in \mathcal{R}(\hat{I}_{s-1})$;
- two sequences $(P_\alpha^r, Q_\alpha^r, n_\alpha^r)$, $(P_\omega^r, Q_\omega^r, n_\omega^r)$, $0 \leq r \leq s$; for each $r < s$, resp. $r = s$, the two elements belonging to $\mathcal{R}(\hat{I}_r)$, resp. $\mathcal{R}(\tilde{I}_{s-1})$;
- two sequences (P_+^r, Q_+^r, n_+^r) , (P_-^r, Q_-^r, n_-^r) , $0 < r \leq s$; for each r , the two elements belonging to $\mathcal{R}(\hat{I}_{r-1})$.

These data are related by the following properties: for each $0 < r \leq s$, we have

$$(9.170) \quad (P^{r-1}, Q^{r-1}, n^{r-1}) \in (P_-^r, Q_-^r, n_-^r) \square (P^r, Q^r, n^r) \square (P_+^r, Q_+^r, n_+^r),$$

$$(9.171) \quad (P_\alpha^{r-1}, Q_\alpha^{r-1}, n_\alpha^{r-1}) \in (P_-^r, Q_-^r, n_-^r) \square (P_\alpha^r, Q_\alpha^r, n_\alpha^r),$$

$$(9.172) \quad (P_\omega^{r-1}, Q_\omega^{r-1}, n_\omega^{r-1}) \in (P_\omega^r, Q_\omega^r, n_\omega^r) \square (P_+^r, Q_+^r, n_+^r).$$

The process stops at step s because of one of the two following cases occur

a) (P^s, Q^s, n^s) does not belong to $\mathcal{R}(\tilde{I}_{s-1})$; then, by the structure theorem of Subsection 6.7, there exists an integer $h > 0$, elements $(P_0^s, Q_0^s, n_0^s) \cdots (P_h^s, Q_h^s, n_h^s)$ in $\mathcal{R}(\tilde{I}_{s-1})$ with

$$(9.173) \quad (P^s, Q^s, n^s) \in (P_0^s, Q_0^s, n_0^s) \square \cdots \square (P_h^s, Q_h^s, n_h^s)$$

and also

$$(9.174) \quad P_0^s \subset P_\alpha^s, \quad Q_h^s \subset Q_\omega^s.$$

b) (P^s, Q^s, n^s) belongs to $\mathcal{R}(I_0)$; in this case we set $h = 0$.

We also observe that the parabolic compositions in (9.170) through (9.172) take place in $\mathcal{R}(\hat{I}_{r-1})$ but not in $\mathcal{R}(\tilde{I}_{r-1})$; in (9.173), they take place in $\mathcal{R}(\hat{I}_{s-1})$ but not in $\mathcal{R}(\tilde{I}_{s-1})$.

A subclass $Bi_+(P_\alpha, Q_\omega, \ell)$, i.e. an element of L , distinct from the two $(\diamond, 0)$ that we already know is determined by the following data

- the depth $s (> 0)$;
- the sequence $\hat{I}_0 \subset \cdots \subset \hat{I}_s = I_0$;
- the integer $h \geq 0$;
- for each $0 < i < h$, the smallest integer u_i such that $|P_i^s| \geq 2^{-u_i} =: x_i$ for some $t \in \tilde{I}$ when $h > 1$;
- when $h > 0$, the smallest integers u_0, u_h such that $|P_-| \geq 2^{-u_0} =: x_0$ for some $t_- \in \tilde{I}$, $|P_+| \geq 2^{-u_h} =: x_h$ for some $t_+ \in \tilde{I}$; here, the elements (P_-, Q_-, n_-) , (P_+, Q_+, n_+) are determined by $P \subset P_-, Q \subset Q_+$ and

$$(9.175) \quad (P_-, Q_-, n_-) \in (P_-^1, Q_-^1, n_-^1) \square \cdots \square (P_-^s, Q_-^s, n_-^s) \square (P_0^s, Q_0^s, n_0^s),$$

$$(9.176) \quad (P_+, Q_+, n_+) \in (P_+^s, Q_+^s, n_+^s) \square (P_+^s, Q_+^s, n_+^s) \square \cdots \square (P_+^1, Q_+^1, n_+^1).$$

Thus, we group together in a subclass $Bi_+(P_\alpha, Q_\omega, \ell)$ the elements of $Bi_+(P_\alpha, Q_\omega)$ who share the same data; the elements of L , distinct from $\diamond, 0$, are the sets of data for which at least one subclass $Bi_+(P_\alpha, Q_\omega, \ell)$ is non-empty, for some $(P_\alpha, Q_\alpha, n_\alpha)$ in $\mathcal{C}_+(\tilde{I})$, $(P_\omega, Q_\omega, n_\omega)$ in $\mathcal{C}_-(\tilde{I})$.

The definition of the set L is now complete.

9.9.5 Sizes of subclasses of higher depth

The context and notations are the same as above. We want to define $b_+(P_\alpha, \ell)$ and $b_-(Q_\omega, \ell)$ in order to satisfy (9.131) in Subsection 9.8.

We first observe that $(P_\alpha, Q_\alpha, n_\alpha)$ determines $(P_-^1, Q_-^1, n_-^1), \dots, (P_-^s, Q_-^s, n_-^s), (P_\alpha^s, Q_\alpha^s, n_\alpha^s)$ and the result of parabolic compositions between these elements. Similarly, $(P_\omega, Q_\omega, n_\omega)$ determines $(P_+^1, Q_+^1, n_+^1), \dots, (P_+^s, Q_+^s, n_+^s), (P_\omega^s, Q_\omega^s, n_\omega^s)$ and the result of parabolic compositions between these elements. Therefore, the only "freedom" for the element (P, Q, n) in the subclass $Bi_+(P_\alpha, Q_\omega, \ell)$ is through (P^s, Q^s, n^s) , and this freedom is constrained by the relations $P^s \subset P_\alpha^s, Q^s \subset Q_\omega^s$.

Consider first a subclass with $h = 0$, i.e., $(P^s, Q^s, n^s) \in \mathcal{R}(I_0)$. The widths of the strips are related as follows: for every $t \in \tilde{I}$, we have

$$(9.177) \quad C^{-1} \frac{|P^s|}{|P_\alpha^s| |P_\omega^s|} \leq \frac{|P|}{|P_\alpha| |P_\omega|} \leq C \frac{|P^s|}{|P_\alpha^s| |P_\omega^s|}.$$

This allows us to take, as in the case of depth 0,

$$(9.178) \quad b_+(P_\alpha, \ell) = \begin{cases} (C|\tilde{I}|^{1+\tau}x^{-1})^{d_s^*} & \text{if } (P_\alpha^s, Q_\alpha^s, n_\alpha^s) \in \mathcal{R}(I_0), \\ 0 & \text{otherwise,} \end{cases}$$

$$(9.179) \quad b_-(Q_\omega, \ell) = \begin{cases} |P_\omega^s|^{d_s^*} & \text{if } (P_\omega^s, Q_\omega^s, n_\omega^s) \in \mathcal{R}(I_0), \\ 0 & \end{cases}$$

Consider now a subclass with $h > 0$, i.e., case a) in Subsection 9.9.4 above.

By the structure theorem in Subsection 6.7, see Lemma 11 in Subsection 9.6, for $0 < i < h$, the element (P_i^s, Q_i^s, n_i^s) belongs to $Bi_+(\tilde{I}_{s-1}, \tilde{I}_{s-1}, \tilde{I}_{s-1}; x_i)$. From Corollary 6 in Subsection 6.7, we have

$$(9.180) \quad x \leq C \left(C|\hat{I}_{s-1}|^{-\frac{1}{2}} \right)^h x_0 x_1 \cdots x_h.$$

Thus, the data of every subclass must satisfy (9.180). Assuming that (9.180) holds, we set $b_+(P_\alpha, \ell) = 0$ if $(P_\alpha, Q_\alpha, n_\alpha) \notin \mathcal{R}(\hat{I}_0)$. When $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{R}(\hat{I}_0)$, we set

$$(9.181) \quad b_+(P_\alpha, \ell) = 2^h \left(\prod_{0 < i < h} \#Bi_+(\tilde{I}_{s-1}, \tilde{I}_{s-1}, \tilde{I}_{s-1}; x_i) \right) \#Bi_+(P_\alpha, \tilde{I}_{s-1}; x_0).$$

Here, $Bi_+(P_\alpha, \tilde{I}_{s-1}, x_0)$ is by definition the set of elements (P_-, Q_-, n_-) in $\mathcal{R}(\tilde{I})$ such that $P_- \subset P_\alpha$, Q_- is thin \tilde{I}_{s-1} -critical and $|P_-| \geq x_0$ for some $t \in \tilde{I}$.

Similarly, when (9.180) holds, we set $b_-(Q_\omega, \ell) = 0$ if $(P_\omega, Q_\omega, n_\omega) \notin \mathcal{R}(\hat{I}_0)$. When $(P_\omega, Q_\omega, n_\omega) \in \mathcal{R}(\hat{I}_0)$, we set

$$(9.182) \quad b_-(Q_\omega, \ell) = \#Bi_+(\tilde{I}_{s-1}, Q_\omega; x_h),$$

where now $Bi_+(\tilde{I}_{s-1}, Q_\omega; x_h)$ is the set of elements (P_+, Q_+, n_+) in $\mathcal{R}(\tilde{I})$ such that $Q_+ \subset Q_\omega$, P_+ is thin \tilde{I}_{s-1} -critical and $|P_+| \geq x_h$ for some $t \in \tilde{I}$.

The factor 2^h in (9.181) takes care of the possible results of the "free" parabolic compositions, i.e., those compositions which are not constrained by $(P_\alpha, Q_\alpha, n_\alpha)$ or $(P_\omega, Q_\omega, n_\omega)$.

The definition of L, b_+, b_- is now complete, and relation (9.131) is satisfied.

9.10 The Size of the Index set L

It is not difficult from (9.180) to see that the index set L is finite, but we need an explicit bound on its cardinality (cf. (9.141)).

Proposition 33. *The index set L satisfies*

$$\#L \leq C|\tilde{I}|^{-\tau^2}.$$

Proof. In the first part of the proof, we fix the depth s and the sequence of intervals $\hat{I}_0 \subset \cdots \subset \hat{I}_{s-1} \subset \hat{I}_s = I_0$. There is one subclass with $h = 0$ and we will estimate the number of subclasses with $h > 0$, i.e., the number of $(h + 1)$ -tuples (u_0, \dots, u_h) such that (9.180) is satisfied; the integer h itself is *not* fixed.

By Corollary 7 in Subsection 6.7, we have

$$(9.183) \quad x_i < |\tilde{I}_{s-1}|^\beta \quad \text{for } 0 < i < h,$$

$$(9.184) \quad x_h < C|\tilde{I}_{s-1}|^{(1-\eta)^{-1}}.$$

As $P_- \subset P_\alpha$, we also have, for a non-empty subclass

$$(9.185) \quad x_0 < |\tilde{I}|^{1+\tau}.$$

We rewrite (9.180) as

$$(9.186) \quad \frac{x_0}{|\tilde{I}|^{1+\tau}} \left(\prod_{0 < i < h} \frac{x_i}{|\tilde{I}_{s-1}|^\beta} \right) \frac{x_h}{C|\tilde{I}_{s-1}|^{\frac{1}{1-\eta}}} \geq \frac{x}{C|\tilde{I}|^{1+\tau} |\tilde{I}_{s-1}|^{\beta(h-1)+(1-\eta)^{-1}}} \left(C^{-1} |\hat{I}_{s-1}|^{\frac{1}{2}} \right)^h.$$

Using $\beta > 1$, and taking base-two logarithms, it is sufficient to bound the number of non-negative integral solutions of

$$(9.187) \quad n_0 + \cdots + n_h \leq A_0 - A_1 h,$$

with

$$(9.188) \quad A_0 = \log_2(|\tilde{I}|x^{-1}),$$

$$(9.189) \quad A_1 = \frac{1}{3} \log_2 |\hat{I}_{s-1}|^{-1}.$$

As $x \leq \tilde{x}_{cr} \ll |\tilde{I}|$, both A_0 and A_1 are large; by taking A_0 slightly larger and A_1 slightly smaller, we can assume that both A_0, A_1 are integers. The number of non-negative integral solutions of (9.187) is then the coefficient of z^{A_0} in the power series for

$$(9.190) \quad \chi(z) := \sum_{h \geq 0} z^{A_1 h} (1-z)^{-h-2} = (1-z)^{-1} (1-z-z^{A_1})^{-1}.$$

We estimate this coefficient by a Cauchy integral on the circle $|z| = 1 - 2A_1^{-1} \log A_1$. On this circle, we have

$$(9.191) \quad |z^{A_1}| < A_1^{-1},$$

$$(9.192) \quad |\chi(z)| < \frac{1}{2} A_1^2 (\log A_1)^{-2}.$$

The number of solutions of (9.187) is, therefore, not greater than

$$(9.193) \quad A_1^2 (\log A_1)^{-2} (1 - 2A_1^{-1} \log A_1)^{-A_0}.$$

In view of (9.189), this quantity is smaller than

$$(9.194) \quad (\log |\widehat{I}_{s-1}|)^2 \exp \left(CA_0 \frac{\log |\log |\widehat{I}_{s-1}||}{|\log |\widehat{I}_{s-1}||} \right).$$

This is a bound for the number of subclasses with fixed depth s and sequence $\widehat{I}_0 \subset \cdots \subset \widehat{I}_{s-1}$. We have now to sum over these remaining data. Observe that (9.194) depends on $|\widehat{I}_{s-1}|$, *not* on the depth s and the intervals \widehat{I}_r , $0 \leq r < s - 1$.

Fix an interval \widehat{I} with $\widetilde{I} \subset \widehat{I} \subset I_0$, $\widehat{I} \neq I_0$. Let $S(\widehat{I})$ be the number of parameter intervals I^* with $\widetilde{I} \subset I^* \subset \widehat{I}$, $I^* \neq \widetilde{I}$. Every I^* in this range may or may not be one of the \widehat{I}_r , for a sequence $\widehat{I}_0 \subset \cdots \subset \widehat{I}_{s-1}$ terminating with $\widehat{I}_{s-1} = \widehat{I}$; in other terms, there are exactly $2^{S(\widehat{I})}$ such sequences (of various lengths). This means that the total number of subclasses is bounded by

$$(9.195) \quad \sum_{\widehat{I}} 2^{S(\widehat{I})} (\log |\widehat{I}|)^2 \exp \left(CA_0 \frac{\log |\log |\widehat{I}||}{|\log |\widehat{I}||} \right).$$

We have here

$$(9.196) \quad \frac{\log |\log |\widehat{I}||}{|\log |\widehat{I}||} \leq \frac{\log \log \varepsilon_0^{-1}}{\log \varepsilon_0^{-1}},$$

$$(9.197) \quad |\log |\widetilde{I}|| = (1 + \tau)^{S(\widehat{I})} \log |\widehat{I}|,$$

$$(9.198) \quad S(\widehat{I}) \leq 2\tau^{-1} \log \left(\frac{\log |\widetilde{I}|^{-1}}{\log \varepsilon_0^{-1}} \right) =: S_{\max}.$$

The sum (9.195) is thus bounded by

$$(9.199) \quad \begin{aligned} & C 2^{S_{\max}} (\log \varepsilon_0^{-1})^2 \exp \left(CA_0 \frac{\log \log \varepsilon_0^{-1}}{\log \varepsilon_0^{-1}} \right) \\ & \leq (\log |\widetilde{I}|^{-1})^{2\tau^{-1}} \exp \left(CA_0 \frac{\log \log \varepsilon_0^{-1}}{\log \varepsilon_0^{-1}} \right). \end{aligned}$$

As $x \geq x_{\min} := |\widetilde{I}|^{C(\rho_0 - d_s^*)^{-1}}$ (cf. (9.64)), we have

$$(9.200) \quad A_0 \leq C(\rho_0 - d_s^*)^{-1} \log |\widetilde{I}|^{-1}.$$

We choose the exponent ρ_0 in order to have

$$(9.201) \quad \rho_0 > d_s^* + \tau.$$

As ε_0 can be chosen arbitrarily small, we have

$$(9.202) \quad CA_0 \frac{\log \log \varepsilon_0^{-1}}{\log \varepsilon_0^{-1}} < \frac{1}{2} \tau^2 \log |\tilde{I}|^{-1}.$$

We conclude, then, that with ε_0 small enough, the term in (9.199) is indeed smaller than $|\tilde{I}|^{-\tau^2}$. \square

9.11 The Size of \widehat{B}

According to the roadmap exposed in Subsection 9.8, we have now to estimate the quantity set in Subsection 9.8

$$(9.140) \quad \widehat{B} = \sum_L b_+(\ell) b_-(\ell)$$

with

$$(9.136) \quad b_+(\ell) = \sum_{c_+(\tilde{I})} b_+(P_\alpha, \ell),$$

$$(9.137) \quad b_-(\ell) = \sum_{c_-(\tilde{I})} b_-(Q_\omega, \ell).$$

Consider first the bound elements. In view of (9.150), we have:

$$(9.203) \quad \begin{aligned} b_+(\diamond) &= \#\mathcal{C}_+(\tilde{I}) \log |\tilde{I}|^{-1}, \\ b_-(\diamond) &= \#\mathcal{C}_-(\tilde{I}) \log |\tilde{I}|^{-1}. \end{aligned}$$

Consider next the class of depth 0, and also the classes of higher depth with $h = 0$: in view of (9.166)–(9.167) and (9.178)–(9.179), we have in these cases

$$(9.204) \quad b_+(\ell) \leq (C|\tilde{I}|^{1+\tau} x^{-1})^{d_s^*} \#\mathcal{C}_+(\tilde{I}),$$

$$(9.205) \quad b_-(\ell) \leq \sum_{c_-(\tilde{I})} (\max_{\tilde{I}} |P_\omega|)^{d_s^*}.$$

Also, the number of such classes, according to the discussion in the proof of Proposition 33 is not larger than

$$(9.206) \quad 2^{S_{\max}} \leq \left(\frac{\log |\tilde{I}|^{-1}}{\log \varepsilon_0^{-1}} \right)^{2\tau^{-1}}.$$

The remaining subclasses are more complicated! Formulas (9.181), (9.182) suggest an induction. We thus assume that $(\text{SR3})_s$ is satisfied for all parameter intervals containing \tilde{I} . We have, for a class of depth $s > 0$ with $h > 0$:

$$(9.207) \quad b_+(\ell) \leq 2^h \left(\prod_{0 < i < h} (\#Bi_+(\tilde{I}_{s-1}, \tilde{I}_{s-1}, \tilde{I}_{s-1}; x_i)) \right) \#Bi_+(\tilde{I}, \tilde{I}, \tilde{I}_{s-1}; x_0)$$

$$(9.208) \quad b_-(\ell) \leq \#Bi_+(\tilde{I}, \tilde{I}_{s-1}, \tilde{I}; x_h).$$

Observe that, from (9.31), the critical value x_{cr} in each of the Bi_+ sets above in the same and equal to

$$(9.209) \quad x_{cr} := \varepsilon_0 |P_u| \left(\frac{|\tilde{I}_{s-1}|}{\varepsilon_0} \right)^{\frac{\sigma_0}{\rho_0 - \rho_1}}.$$

As in Subsection 9.6, we separate the subclasses into two parts: those for which every x_i is above the critical value x_{cr} and the others. In the first case, we have from $(\text{SR3})_s$

$$(9.210) \quad \#Bi_+(\tilde{I}_{s-1}, \tilde{I}_{s-1}, \tilde{I}_{s-1}; x_i) \leq C \left(\frac{x_i}{\varepsilon_0 |P_u|} \right)^{-\rho_1} \left(\frac{|\tilde{I}_{s-1}|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1}, \text{ for } 0 < i < h,$$

$$(9.211) \quad \#Bi_+(\tilde{I}, \tilde{I}, \tilde{I}_{s-1}; x_0) \leq C \left(\frac{x_0}{\varepsilon_0 |P_u|} \right)^{-\rho_1} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1},$$

$$(9.212) \quad \#Bi_+(\tilde{I}, \tilde{I}_{s-1}, \tilde{I}; x_h) \leq C \left(\frac{x_h}{\varepsilon_0 |P_u|} \right)^{-\rho_1} \left(\frac{|\tilde{I}_{s-1}|}{\varepsilon_0} \right)^{\sigma_1} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{\sigma_0}.$$

Multiplying these inequalities, we obtain, taking (9.180) into account

$$(9.213) \quad b_+(\ell) b_-(\ell) \leq A_2^h A_3$$

with

$$(9.214) \quad A_2 = 2 \left(\frac{|\tilde{I}_{s-1}|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1} \left(C \varepsilon_0 |P_u| |\tilde{I}_{s-1}|^{-\frac{1}{2}} \right)^{\rho_1}$$

$$(9.215) \quad A_3 = C^{\rho_1} \left(\frac{x}{\varepsilon_0 |P_u|} \right)^{-\rho_1} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1}$$

In the second case, as $\rho_0 > \rho_1$, we have

$$(9.216) \quad \#Bi_+(\tilde{I}_{s-1}, \tilde{I}_{s-1}, \tilde{I}_{s-1}; x_i) \leq C \left(\frac{x_i}{\varepsilon_0 |P_u|} \right)^{-\rho_0} \left(\frac{|\tilde{I}_{s-1}|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1},$$

$$(9.217) \quad \#Bi_+(\tilde{I}, \tilde{I}, \tilde{I}_{s-1}; x_0) \leq C \left(\frac{x_0}{\varepsilon_0 |P_u|} \right)^{-\rho_0} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1},$$

$$(9.218) \quad \#Bi_+(\tilde{I}, \tilde{I}_{s-1}, \tilde{I}; x_h) \leq C \left(\frac{x_h}{\varepsilon_0 |P_u|} \right)^{-\rho_0} \left(\frac{|\tilde{I}_{s-1}|}{\varepsilon_0} \right)^{\sigma_1} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{\sigma_0}.$$

Moreover, comparing the σ exponents in (9.19) and (9.20), we see that if $x_i \leq x_{cr}$, $0 \leq i \leq h$, the corresponding inequality is still true after multiplying the right-hand side by $\left(\frac{|\tilde{I}_{s-1}|}{\varepsilon_0} \right)^{\sigma_0}$. As this happens at least once, we get, by multiplying the three inequalities together:

$$(9.219) \quad b_+(\ell) b_-(\ell) \leq \tilde{A}_2^h \tilde{A}_3,$$

now with

$$(9.220) \quad \tilde{A}_2 = 2 \left(\frac{|\tilde{I}_{s-1}|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1} \left(C \varepsilon_0 |P_u| |\hat{I}_{s-1}|^{-\frac{1}{2}} \right)^{\rho_0}$$

$$(9.221) \quad \tilde{A}_3 = C^{\rho_0} \left(\frac{x}{\varepsilon_0 |P_u|} \right)^{-\rho_0} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{2\sigma_0 + \sigma_1}.$$

With \tilde{B} as in Subsection 9.8, we have $\max(A_3, \tilde{A}_3) \leq C\tilde{B}$.

We observe that in both (9.213) and (9.219), our estimate for $b_+(\ell)$, $b_-(\ell)$ depends on the class ℓ only through \hat{I}_{s-1} and h . We first sum over subclasses with a fixed depth s and sequence $\hat{I}_0 \subset \dots \subset \hat{I}_{s-1}$, using the same method of generating series as in the proof of Proposition 33. To deal with the two cases at the same time, we first observe that

$$(9.222) \quad \sigma_0 + \sigma_1 - \frac{1}{2} \rho_0 (1 + \tau) = 1 - d_u^0 - \frac{1}{2} d_s^0 + o(1) > 0$$

under (H4), and a fortiori $\sigma_0 + \sigma_1 - \frac{1}{2} \rho_1 (1 + \tau) > 0$. Thus, A_2 and \tilde{A}_2 are larger when \tilde{I}_{s-1} is larger; the largest case is $\tilde{I}_{s-1} = I_0$, which gives

$$(9.223) \quad \max(A_2, \tilde{A}_2) \leq \hat{A}_2 := 2 \left(C \varepsilon_0^{\frac{1}{2}(1-\tau)} |P_u| \right)^{\rho_1}.$$

We, then, set

$$(9.224) \quad \begin{aligned} \chi_1(z) &= \sum_{h>0} \hat{A}_2^h z^{A_1 h} (1-z)^{-h-2} \\ &= \hat{A}_2 z^{A_1} (1-z)^{-2} (1-z - \hat{A}_2 z^{A_1})^{-1}. \end{aligned}$$

The (partial) sum of $b_+(\ell) b_-(\ell)$ is, thus, not larger than $C\tilde{B}$ times the coefficient of z^{A_0} in the power series for $\chi_1(z)$. Recall that A_0, A_1 were defined in (9.188), (9.189).

We estimate this coefficient by Cauchy integration on the circle $\{|z| = 1 - A_0^{-1} - \hat{A}_2\}$, on which we have

$$(9.225) \quad |1-z|^{-2} \leq (\hat{A}_2 + A_0^{-1})^{-2} \leq A_0^2,$$

$$(9.226) \quad |1-z - \hat{A}_2 z^{A_1}|^{-1} \leq A_0,$$

$$(9.227) \quad |\chi_1(z)| \leq \hat{A}_2 A_0^3,$$

$$(9.228) \quad |z^{-A_0}| \leq C(1 + \hat{A}_2)^{A_0}.$$

The (partial) sum of $b_+(\ell)b_-(\ell)$ is therefore dominated by

$$(9.229) \quad C(1 + \widehat{A}_2)^{A_0} \widehat{A}_2 A_0^3 \widetilde{B}$$

We now have to sum over sequences $\widehat{I}_0 \subset \cdots \subset \widehat{I}_{s-1}$ and depth s ; but (9.229) is independent of these data and the same remarks as in the proof of Proposition 33 apply. So, we finally obtain for the sum of $b_+(\ell)b_-(\ell)$ over subclasses with $s > 0$ and $h > 0$, a bound by

$$(9.230) \quad C(1 + \widehat{A}_2)^{A_0} \widehat{A}_2 A_0^3 2^{S_{\max}} \widetilde{B}$$

with $S_{\max} := 2\tau^{-1} \log\left(\frac{\log |\widetilde{I}|^{-1}}{\log \varepsilon_0^{-1}}\right)$ (cf. (9.198)).

We have here, from (9.200), (9.201)

$$(9.231) \quad A_0 \leq C\tau^{-1} \log |\widetilde{I}|^{-1},$$

$$(9.232) \quad (1 + \widehat{A}_2)A_0 \leq |\widetilde{I}|^{-C\tau^{-1}\widehat{A}_2},$$

$$(9.233) \quad \widehat{A}_2 \leq \varepsilon_0^{\frac{1}{2}\rho_1}.$$

As ε_0 can be made as small as we want with regard to τ , the term in (9.230) is bounded by

$$(9.234) \quad \varepsilon_0^\sigma |\widetilde{I}|^{-\varepsilon_0^\sigma} (\log |\widetilde{I}|^{-1})^{2\tau^{-1}} \widetilde{B}$$

for some fixed $\sigma > 0$.

We summarize the calculations in this subsection in

Proposition 34. *The quantity $\widehat{B} = \sum_L b_+(\ell)b_-(\ell)$ is bounded by $\widehat{B}_1 + \widehat{B}_2 + \widehat{B}_3$, with*

$$\begin{aligned} \widehat{B}_1 &= (\#\mathcal{C}_+(\widetilde{I}))(\#\mathcal{C}_-(\widetilde{I}))(\log |\widetilde{I}|^{-1})^2, \\ \widehat{B}_2 &= \left(\frac{\log |\widetilde{I}|^{-1}}{\log \varepsilon_0^{-1}}\right)^{2\tau^{-1}} (C|\widetilde{I}|^{1+\tau} x^{-1})^{d_s^*} (\#\mathcal{C}_+(\widetilde{I})) \sum_{\mathcal{C}_-(\widetilde{I})} (\max_{\widetilde{I}} |P_\omega|)^{d_s^*}, \\ \widehat{B}_3 &= \varepsilon_0^\sigma |\widetilde{I}|^{-\varepsilon_0^\sigma} (\log |\widetilde{I}|^{-1})^{2\tau^{-1}} \widetilde{B}. \end{aligned}$$

9.12 End of the Induction Step for (SR3)_s

Of the two inequalities that were assumed in Subsection 9.8 to make work the argument, the first has been the subject of Proposition 33. The second is

$$(9.144) \quad \widehat{B} |\widetilde{I}|^{\tau(2-d_s^+ - d_u^+ - 6\tau)} \leq CB$$

From Proposition 34, it is sufficient to prove the same inequality with \widehat{B} replaced by \widehat{B}_i , $i = 1, 2, 3$.

First consider \widehat{B}_3 . For $x \leq \widetilde{x}_{cr}$, we have

$$(9.235) \quad \widetilde{B} = \widetilde{B}_0 = |\widetilde{I}|^{-\tau(2\sigma_0 + \sigma_1)} B_0 \leq |\widetilde{I}|^{-\tau(2\sigma_0 + \sigma_1)} B,$$

and, therefore,

$$(9.236) \quad \widehat{B}_3 |\widetilde{I}|^{\tau(2-d_s^+ - d_u^+ - 6\tau)} \leq \varepsilon_0^\sigma (\log |\widetilde{I}|^{-1})^{2\tau-1} |\widetilde{I}|^\omega B,$$

with $\omega = \tau(2 - d_s^+ - d_u^+ - 6\tau - 2\sigma_0 - \sigma_1) - \varepsilon_0^\sigma$.

We choose the exponents σ_0, σ_1 in order to have

$$(9.237) \quad \omega > \tau^2$$

which means

$$(9.238) \quad 2\sigma_0 + \sigma_1 < 2 - d_s^+ - d_u^+ - 7\tau - \varepsilon_0^\sigma \tau^{-1}$$

which is compatible with the previous conditions on σ_0, σ_1 .

As ε_0 can be chosen arbitrarily small with respect to τ , (9.144) holds for \widehat{B}_3 .

Next consider \widehat{B}_1 . We assume that the sizes of $\mathcal{C}_+(\widetilde{I}), \mathcal{C}_-(\widetilde{I})$ are controlled by (SR1)_s, (SR1)_u. Then, we have

$$(9.239) \quad \widehat{B}_1 |\widetilde{I}|^{\tau(2-d_s^+ - d_u^+ - 6\tau)} \leq C (\log |\widetilde{I}|^{-1})^2 \left(\frac{|\widetilde{I}|}{\varepsilon_0} \right)^{2-2d_s^+ - 2d_u^+ - 2\tau} \varepsilon_0^{-\tau(d_s^0 + d_u^0)} |\widetilde{I}|^{\tau(2-d_s^+ - d_u^+ - 6\tau)}.$$

The right-hand will be smaller than CB_0 as soon as

$$(9.240) \quad \left(\frac{x}{\varepsilon_0 |P_u|} \right)^{\rho_0} \leq \left(\frac{|\widetilde{I}|}{\varepsilon_0} \right)^{2\sigma_0 + \sigma_1 - 2 + 2d_s^+ + 2d_u^+ + A\tau} \varepsilon_0^{A\tau} (\log |\widetilde{I}|^{-1})^{-2},$$

for some fixed constant $A > 0$.

We would like (9.240) to be a consequence of $x \leq \widetilde{x}_{cr}$, but unfortunately this is only true if \widetilde{I} is not too large. Observe that

$$(9.241) \quad \rho_0^{-1} (2\sigma_0 + \sigma_1 - 2 + 2d_s^+ + 2d_u^+ + A\tau) - (\rho_0 - \rho_1)^{-1} \sigma_0 = o(1),$$

where $(\rho_0 - \rho_1)^{-1} \sigma_0$ is the exponent appearing in the definition (9.31) of \widetilde{x}_{cr} . Write

$$(9.242) \quad \sigma := 2\sigma_0 + \sigma_1 - 2 + 2d_s^+ + 2d_u^+ + A\tau = d_s^+ + d_u^+ + o(1).$$

We choose the exponent ρ_1 in order to have

$$(9.243) \quad \rho_1 > \rho_0 \left(1 - \frac{\sigma_0}{\sigma} \right) + \kappa,$$

with $\kappa > 0$ small to guarantee that (9.22) holds. Then, we have

$$(9.244) \quad \rho_0^{-1} \sigma < \frac{\sigma_0}{\rho_0 - \rho_1 + \kappa} < \frac{\sigma_0}{(\rho_0 - \rho_1)} - C^{-1} \kappa$$

and therefore (9.240) holds as soon as

$$(9.245) \quad x \leq \tilde{x}_{cr} \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{-C^{-1}\kappa} \varepsilon_0^{A\tau} (\log |\tilde{I}|^{-1})^{-\frac{2}{\rho_0}}$$

Keeping $\kappa \gg \tau$, the right-hand side is larger than \tilde{x}_{cr} when $|\tilde{I}| < \varepsilon_0^{1+C\kappa^{-1}\tau}$, C fixed large enough. Thus, we are able to conclude that (9.144) holds for \widehat{B}_1 except in the range

$$(9.246) \quad \begin{aligned} \varepsilon_0 &\geq |\tilde{I}| \geq \varepsilon_0^{1+C\kappa^{-1}\tau}, \\ \tilde{x}_{cr} &\geq x \geq \tilde{x}_{cr} \varepsilon_0^{2A\tau}. \end{aligned}$$

We shall deal directly with this case below. Before that, we consider (9.144) for \widehat{B}_2 . In this case, we assume that (SR1)_s and (SR2)_u hold. A small calculation shows that (9.144) holds as soon as

$$(9.247) \quad \left(\frac{x}{\varepsilon_0 |P_u|} \right)^{\rho_0 - d_s^*} \leq \left(\frac{|\tilde{I}|}{\varepsilon_0} \right)^{\widehat{\sigma}} \varepsilon_0^{\widehat{A}\tau} \left(\frac{\log |\tilde{I}|^{-1}}{\log \varepsilon_0^{-1}} \right)^{-2\tau^{-1}}$$

with

$$(9.248) \quad \widehat{\sigma} = -2 + 2d_s^+ + d_u^+ - d_s^* + 2\sigma_0 + \sigma_1 + \widehat{A}\tau$$

and \widehat{A} a fixed positive constant. Here, both $\rho_0 - d_s^*$ and $\widehat{\sigma}$ are $o(1)$.

With κ as above, i.e., $\kappa = o(1)$, $\tau = o(\kappa)$, we choose the exponents in order to have

$$(9.249) \quad \rho_0 > d_s^* + \kappa^2,$$

$$(9.250) \quad \widehat{\sigma} < -2\kappa,$$

which corresponds to

$$(9.251) \quad 2\sigma_0 + \sigma_1 < 2 - 2d_s^+ - d_u^+ + d_s^* - \widehat{A}\tau - 2\kappa.$$

If (\tilde{I}, x) is not in the range

$$(9.252) \quad \begin{aligned} \varepsilon_0 &\geq |\tilde{I}| \geq \varepsilon_0^{1+C\kappa^{-1}\tau}, \\ \tilde{x}_{cr} &\geq x \geq \tilde{x}_{cr} \varepsilon_0^{C\kappa^{-1}\tau}, \end{aligned}$$

then, (9.247) follows from $x \leq \tilde{x}_{cr}$.

The final step in the inductive proof of (SR3)_s is, therefore, the proof of

Proposition 35. *Assume that*

$$|\tilde{I}| \leq \varepsilon_0^{1+C\tau\kappa^{-1}}$$

and that (SR1)_{ŝ}, (SR1)_û, (SR2)_û hold for some candidate $I \subset \tilde{I}$. Then (SR3)_s holds for I in the range

$$\tilde{x}_{cr} \geq x \geq \tilde{x}_{cr} \varepsilon_0^{C\tau\kappa^{-1}}.$$

Proof. Any element (P, Q, n) in $Bi_+(I, I, I; x)$ satisfies $P \subset P_s, Q \subset Q_u$. Using (H1), this implies, for all $t \in I$:

$$(9.253) \quad |P| \leq C\varepsilon_0|P_u|.$$

It, then, follows from Corollary 6 in Subsection 6.7 that for $x \geq \tilde{x}_{cr} \varepsilon_0^{C\tau\kappa^{-1}}$, one must have $(P, Q, n) \in \mathcal{R}(I_0)$.

As P is thin I -critical, there exists $(P_\alpha, Q_\alpha, n_\alpha) \in \widehat{\mathcal{C}}_+(I)$ with $P \subset P_\alpha$. Similarly, there exists $(P_\omega, Q_\omega, n_\omega) \in \widehat{\mathcal{C}}_-(I)$ with $Q \subset Q_\omega$; see Subsection 9.1 for the definitions of $\widehat{\mathcal{C}}_-(I), \widehat{\mathcal{C}}_+(I)$.

Let us estimate the number of possible (P, Q, n) for fixed P_α, Q_ω .

If $n \leq n_\alpha + n_\omega$, there is at most one for each value of n . If $n \geq n_\alpha + n_\omega$, we can write (as $(P, Q, n) \in \mathcal{R}(I_0)$), (P, Q, n) as a simple composition

$$(P, Q, n) = (P_\alpha, Q_\alpha, n_\alpha) * (P', Q', n') * (P_\omega, Q_\omega, n_\omega)$$

and conclude that there are no more than

$$(9.254) \quad C \left(\frac{x}{|I||P_\omega|} \right)^{-d_s^0}$$

possible P' (at this scale, the dependence of dimension on the parameter is not relevant).

The total number is, therefore, at most

$$(9.255) \quad C \log\left(\frac{\varepsilon_0|P_u|}{x}\right) (\#\widehat{\mathcal{C}}_-(I)) (\#\widehat{\mathcal{C}}_+(I)) + C (\#\widehat{\mathcal{C}}_+(I)) \left(\frac{x}{|I|}\right)^{-d_s^*} \sum_{\widehat{\mathcal{C}}_-(I)} |P_\omega|^{d_s^*}.$$

In view of (SR1) $_{\hat{s}}$, (SR1) $_{\hat{u}}$, (SR2) $_{\hat{u}}$, this is not greater than

$$(9.256) \quad C \left(\log \frac{\varepsilon_0|P_u|}{x} \right) \left(\frac{|I|}{\varepsilon_0} \right)^{2-2d_s^+ - 2d_u^+ - 2\tau} + C \left(\frac{|I|}{\varepsilon_0} \right)^{2-2d_s^+ - d_u^+ - 2\tau + d_s^*} \left(\frac{x}{\varepsilon_0|P_u|} \right)^{-d_s^*}$$

and this should be smaller than CB_0 with

$$(9.257) \quad B_0 = \left(\frac{x}{\varepsilon_0|P_u|} \right)^{-\rho_0} \left(\frac{|I|}{\varepsilon_0} \right)^{2\sigma_0 + \sigma_1}$$

and

$$(9.258) \quad \frac{x}{\varepsilon_0|P_u|} \leq \left(\frac{|I|}{\varepsilon_0} \right)^{\frac{\sigma_0}{\rho_0 - \rho_1}}.$$

The second part of (9.256) is as required as $\rho_0 > d_s^*$ (cf. (9.65) and we ask that

$$(9.259) \quad 2\sigma_0 + \sigma_1 < 2 - 2d_s^+ - d_u^+ - 2\tau + d_s^*.$$

In the first part of (9.256), we bound the logarithmic term by a small negative power of $\frac{x}{\varepsilon_0|P_u|}$ and we use (9.258) to conclude.

This ends the proof of (SR3) $_s$ in this case. □

The induction step for $(\text{SR3})_s$ is complete. We have proved

Theorem 2. *Assume that all parameter intervals which contain \tilde{I} are strongly regular. Then, all candidates but a proportion not larger than $|\tilde{I}|^{\tau^2}$ satisfy $(\text{SR3})_s$.*

9.13 The Induction Step for $(\text{SR3})_u$

We have already explained for $(\text{SR3})_u$ the cases where I_α or I_ω contains \tilde{I} , and the case where $I = I_\alpha = I_\omega$ and x is large (cf. Proposition 31).

The case $I = I_\alpha = I_\omega$, x small which has been treated for $(\text{SR3})_s$ in the Subsections 9.8–9.13 is completely similar for $(\text{SR3})_u$. It is only not completely symmetric because we have assumed that $d_s^0 \geq d_u^0$ and thus the formulas for the exponents are not symmetric. So, one has only to be careful with the inequalities involving the exponents. For instance, in (9.222) we had

$$\sigma_0 + \sigma_1 - \frac{1}{2} \rho_0(1 + \tau) > 0.$$

As $\rho'_0 = \frac{d_u^0}{d_s^0} \leq \rho_0$, we still have

$$\sigma_0 + \sigma_1 - \frac{1}{2} \rho'_0(1 + \tau) > 0.$$

Checking everything in this way is rather tedious, and we leave this to the reader.

Remark. *At several points, we have asked that the exponents $\rho_0, \rho_1, \rho'_0, \rho'_1, \sigma_0, \sigma_1$ of $(\text{SR3})_s, (\text{SR3})_u$ should satisfy some inequalities; one could worry whether these inequalities are compatible between themselves (it is easy to check that each is compatible with (9.21) through (9.24) and (9.28) through (9.29)). But we are always bounding $\rho_0, \rho_1, \rho'_0, \rho'_1$ from below and σ_0, σ_1 from above, hence the compatibility is obvious.*

10 The Well-Behaved Part of the Dynamics for Strongly Regular Parameters

10.1 Prime Elements and Prime Decomposition

In the last two sections, we fix a strongly regular parameter, i.e. the intersection of a decreasing sequence $(I_m)_{m \geq 0}$ of strongly regular parameter intervals.

The sequence $\mathcal{R}(I_m)$ is increasing and we set

$$(10.1) \quad \mathcal{R} = \bigcup_{m \geq 0} \mathcal{R}(I_m).$$

Definition. An element $(P, Q, n) \in \mathcal{R}$ is *prime* if $n > 0$ and it cannot be written as a *simple* composition of two shorter elements.

Obviously, for any $(a, a') \in \mathcal{B}$, the element $(P_{aa'}, Q_{aa'}, 1)$ is prime. Such elements are called *trivial* primes. Non trivial primes are those of length bigger than 1.

There are only finitely many trivial primes. On the other hand, there are typically countably many non trivial ones.

Proposition 36. *Any element $(P, Q, n) \in \mathcal{R}$ with $n > 0$ can be uniquely written as a simple composition of a finite sequence of prime elements.*

Proof. The existence of such a decomposition is clear. We have to show it is unique. Assume on the opposite that we can write

$$(10.2) \quad \begin{aligned} (P, Q, n) &= (P_1, Q_1, n_1) * \cdots * (P_r, Q_r, n_r) \\ &= (P'_1, Q'_1, n'_1) * \cdots * (P'_s, Q'_s, n'_s) \end{aligned}$$

It is sufficient to show that $(P_1, Q_1, n_1) = (P'_1, Q'_1, n'_1)$. This is true if $n_1 = n'_1$. Assume for instance that $n_1 < n'_1$. Then we have $P \subset P'_1 \subset P_1$ with $P'_1 \neq P_1$. By Proposition 8 in Subsection 6.5, we can write

$$(10.3) \quad (P'_1, Q'_1, n'_1) = (P_1, Q_1, n_1) * (\tilde{P}, \tilde{Q}, \tilde{n})$$

which contradicts the fact that (P'_1, Q'_1, n'_1) is prime. □

Remark. *In the prime decomposition*

$$(10.4) \quad (P, Q, n) = (P_1, Q_1, n_1) * (P_r, Q_r, n_r),$$

P_1 can be characterized as the thinnest prime rectangle containing P .

We will denote by \mathcal{P} the set of prime elements of \mathcal{R} . We denote by \mathcal{R}^* the set of elements of \mathcal{R} of length > 0 .

Let (P, Q, n) be an element of \mathcal{R}^* and let

$$(10.5) \quad (P, Q, n) = (P_1, Q_1, n_1) * \cdots * (P_r, Q_r, n_r),$$

be its prime decomposition. We define

$$(10.6) \quad \begin{aligned} T^+((P, Q, n)) &= (P_2, Q_2, n_2) * \cdots * (P_r, Q_r, n_r), \\ T^-((P, Q, n)) &= (P_1, Q_1, n_1) * \cdots * (P_{r-1}, Q_{r-1}, n_{r-1}), \end{aligned}$$

if $r > 1$. When (P, Q, n) is prime, with $P \subset R_a$ and $Q \subset R_{a'}$, we set

$$(10.7) \quad \begin{aligned} T^+((P, Q, n)) &= (R_{a'}, R_{a'}, 0) \\ T^-((P, Q, n)) &= (R_a, R_a, 0). \end{aligned}$$

For $S = (P, Q, n) \in \mathcal{R}$, we write $S * \mathcal{R}$, resp. $\mathcal{R} * S$, for the set of elements which can be written as $(P, Q, n) * (P', Q', n')$, resp. $(P', Q', n') * (P, Q, n)$, for some $(P', Q', n') \in \mathcal{R}$. We have partitions

$$(10.8) \quad \mathcal{R}^* = \bigsqcup_{\mathcal{P}} S * \mathcal{R} = \bigsqcup_{\mathcal{P}} \mathcal{R} * S.$$

Moreover, for any $S \in \mathcal{P}$, the restriction of T^+ , resp. T^- , to $S * \mathcal{R}$, resp. $\mathcal{R} * S$, is a bijection onto \mathcal{R} , inverse of $S' \mapsto S * S'$, resp. $S' \mapsto S' * S$.

10.2 Number of Factors in a Prime Decomposition

We write $r(S)$ for the number of factors in the prime decomposition of an element S of \mathcal{R} (setting $r(S) = 0$ if S has length 0). Let $(P, Q, n), (P', Q', n')$ be elements of \mathcal{R} such that P' is a child of P . When P' is a simple child, it is obtained by simple composition of P with an element of length 1 and we have

$$(10.9) \quad r(P', Q', n') = r(P, Q, n) + 1.$$

Proposition 37. *If P' is a non-simple child of P , we have*

$$r(P', Q', n') \leq r(P, Q, n)$$

Proof. Let

$$(10.10) \quad (P, Q, n) = (P_1, Q_1, n_1) * \cdots * (P_r, Q_r, n_r)$$

be the prime decomposition of (P, Q, n) .

Let $(\widehat{P}, \widehat{Q}, \widehat{n})$ be the element of R such that

$$(10.11) \quad (P', Q', n') \in (P, Q, n) \square (\widehat{P}, \widehat{Q}, \widehat{n}),$$

(cf. Proposition 5 in Subsection 6.4). There exists $m \geq 0$ such that Q and \widehat{P} are I_m -transverse. Define

$$(10.12) \quad \begin{aligned} (P^i, Q^i, n^i) &= (P_i, Q_i, n_i) * \cdots * (P_r, Q_r, n_r) \\ &= (T^+)^{i-1} (P, Q, n), \quad 1 \leq i \leq r. \end{aligned}$$

We have an increasing sequence

$$(10.13) \quad Q = Q^1 \subset Q^2 \subset \cdots \subset Q^r = Q_r.$$

Let r' be the largest integer in $\{1, \dots, r\}$ such that $Q^{r'}$ and \widehat{P} are I_m -transverse for some $m \geq 0$ (and then for all large enough m). Define $(\widetilde{P}, \widetilde{Q}, \widetilde{n}) \in \mathcal{R}$ by the condition $Q' \subset \widetilde{Q}$ and

$$(10.14) \quad (\widetilde{P}, \widetilde{Q}, \widetilde{n}) \in (P^{r'}, Q^{r'}, n^{r'}) \square (\widehat{P}, \widehat{Q}, \widehat{n}).$$

We then have

$$(10.15) \quad (P', Q', n') = (P_1, Q_1, n_1) * \cdots * (P_{r'-1}, Q_{r'-1}, n_{r'-1}) * (\widetilde{P}, \widetilde{Q}, \widetilde{n}).$$

The assertion of the proposition is, thus, a consequence of

Lemma 12. $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$ is prime.

Proof. Assume by contradiction that we can write

$$(10.16) \quad (\widetilde{P}, \widetilde{Q}, \widetilde{n}) = (\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1) * (\widetilde{P}_2, \widetilde{Q}_2, \widetilde{n}_2)$$

with $\widetilde{n}_1, \widetilde{n}_2 > 0$. Define

$$(10.17) \quad (P'_1, Q'_1, n'_1) = (P_1, Q_1, n_1) * \cdots * (P_{r'-1}, Q_{r'-1}, n_{r'-1}) * (\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1).$$

We have $P' \subset P'_1$, $P' \neq P'_1$, hence $P \subset P'_1$. We also have

$$(10.18) \quad (P', Q', n') = (P'_1, Q'_1, n'_1) * (\widetilde{P}_2, \widetilde{Q}_2, \widetilde{n}_2).$$

By Proposition 8 in Subsection 6.5, there exists (P^*, Q^*, n^*) such that

$$(10.19) \quad (P, Q, n) = (P'_1, Q'_1, n'_1) * (P^*, Q^*, n^*).$$

We claim that there exists $j \in \{1, \dots, r\}$ such that

$$(10.20) \quad (P'_1, Q'_1, n'_1) = (P_1, Q_1, n_1) * \cdots * (P_j, Q_j, n_j).$$

Indeed, if j is the smallest integer such that

$$(10.21) \quad n'_1 \leq n_1 + \cdots + n_j,$$

it follows from Proposition 8 that we can write

$$(10.22) \quad (P_1, Q_1, n_1) * \cdots * (P_j, Q_j, n_j) = (P'_1, Q'_1, n'_1) * (\bar{P}, \bar{Q}, \bar{n})$$

for some $(\bar{P}, \bar{Q}, \bar{n}) \in \mathcal{R}$. By the same proposition, we can also write

$$(10.23) \quad (P_j, Q_j, n_j) = (\bar{P}', \bar{Q}', \bar{n}') * (\bar{P}, \bar{Q}, \bar{n})$$

for some $(\bar{P}', \bar{Q}', \bar{n}') \in \mathcal{R}$. We have $\bar{n} < n_j$ by the definition of j , hence $\bar{n}' > 0$. As (P_j, Q_j, n_j) is prime, we have $\bar{n} = 0$ which implies our claim.

The integer j satisfies $r' \leq j \leq r$. We have

$$(10.24) \quad (\tilde{P}_2, \tilde{Q}_2, \tilde{n}_2) \in (P^{j+1}, Q^{j+1}, n^{j+1}) \square (\hat{P}, \hat{Q}, \hat{n}),$$

which contradicts the definition of r' .

This concludes the proof of the lemma and also of the proposition. □

10.3 A Weighted Estimate on the Number of Children

We present in this subsection a variation over the estimates in Subsection 8.1, which will be important in the definition of a transfer operator.

We fix a constant $\kappa \in (0, 1)$ close to 1, but independent of ε_0 . We set

$$(10.25) \quad d_s^- = d_s^0 - C\varepsilon_0,$$

with a constant C sufficiently large so d_s^- is smaller than the transverse Hausdorff dimension of the stable foliation $W^s(K)$ for the parameter that we are dealing with.

For $S = (P, Q, n)$, we set

$$(10.26) \quad \|P\| = |P|^{d_s^-} \kappa^{r(S)},$$

(we will also write $r(P)$ instead of $r(S)$).

Proposition 38. *For any $m \geq 1$, any $(P, Q, n) \in \mathcal{R}$, we have*

$$\sum_{P'} \|P'\| \leq C \kappa^{\frac{m}{2}} \|P\|$$

where the sum in the left-hand side is over elements (P', Q', n') such that P' is a descendent of the m^{th} generation of P .

We will first state a Lemma, then prove the proposition from the Lemma, and finally prove the Lemma.

Lemma 13. *Let $\varepsilon_1 > 0$. If ε_0 is small enough, we have*

$$\sum_{P'} \|P'\| \leq \varepsilon_1 \|P\|$$

for all $(P, Q, n) \in \mathcal{R}$, where the sum in the left-hand side is over non-simple children of P .

Proof of the Proposition.

Let $m_0 \geq 1$ be an integer to be determined later. Consider all chains

$$(10.27) \quad P = P^0 \supset P^1 \supset \dots \supset P^{m_0} = P'$$

where P is given and P^{i+1} is a child of P^i . If P^{i+1} is, for each i , a simple child of P^i , one has $r(P') = m_0 + r(P)$ and the corresponding part of the sum in Proposition 38 satisfies

$$(10.28) \quad \sum \|P'\| \leq C \kappa^{m_0} \|P\|$$

(as long as $m_0 = o(\varepsilon_0^{-1})$). We choose m_0 such that in (10.28) we have From the lemma above, it follows that for every \tilde{P} , we have

$$(10.29) \quad 2C \kappa^{m_0} \leq \kappa^{\frac{m_0}{2}}.$$

$$(10.30) \quad \sum \|\tilde{P}'\| \leq C \|\tilde{P}\|$$

where the sum is over all children of \tilde{P} . Using the lemma again, when we sum over chains such that P_{i+1} is a non-simple child of P_i for some i , we obtain

$$(10.31) \quad \sum \|\tilde{P}'\| \leq m_0 C^{m_0-1} \varepsilon_1 \|P\|.$$

Taking ε_1 small enough, we obtain

$$(10.32) \quad \sum \|P'\| \leq \kappa^{\frac{m_0}{2}} \|P\|$$

where the sum is now over all chains. The proposition follows immediately from (10.32) and (10.30).

Proof of Lemma 13. Let $(P, Q, n) \in \mathcal{R}$. Any non-simple child P' of P is obtained as

$$(10.33) \quad (P', Q', n') \in (P, Q, n) \square (P_1, Q_1, n_1)$$

and we denote by \tilde{P}_1 , the parent of P_1 . One has

$$(10.34) \quad |P'| \leq C |P| |P_1| \delta(Q, P_1)^{-\frac{1}{2}}.$$

Therefore, we will have

$$(10.35) \quad \|P\|^{-1} \sum \|P'\| \leq C \sum |P_1|^{d_s^-} \kappa^{r(P')-r(P)} \delta^{-\frac{1}{2} d_s^-}.$$

From the proof of Proposition 37, there is an increasing sequence

$$(10.36) \quad Q = Q^1 \subset Q^2 \subset \dots \subset Q^{r(P)}$$

such that $r(P')$ is the largest integer for which Q^r and P_1 are I_m -transverse for large enough m . If

$$(10.37) \quad |Q^r| \leq \delta(Q, P_1)^2$$

then, by Proposition 10, Q^r and P_1 are I_m -transverse for large m and thus $r(P') \geq r$. On the other hand, there exists $\kappa^* \in (0, 1)$ such that

$$(10.38) \quad |Q^r| \leq \kappa^* |Q^{r+1}|$$

for $r < r(P)$. We infer that

$$(10.39) \quad r(P) - r(P') \leq C \log(\delta(Q, P_1))^{-1}.$$

Therefore, if κ is close enough to 1, we have

$$(10.40) \quad \kappa^{r(P')-r(P)} \leq (\delta(Q, P_1))^{-\frac{1}{6} d_s^-}$$

and the right-hand side of (10.35) is bounded by

$$(10.41) \quad C \sum |P_1|^{d_s^-} (\delta(Q, P_1))^{-\frac{2}{3} d_s^-}$$

Using (R7), this is smaller than

$$(10.42) \quad C \sum |P_1|^{\frac{1}{3} d_s^-}.$$

To estimate this sum, we first fix the parent \tilde{P}_1 and sum over children P_1 ; it follows from Proposition 21 that the corresponding sum is bounded by $C |\tilde{P}_1|^{\frac{1}{3} d_s^-}$ and (10.42) is not greater than

$$(10.43) \quad C \sum |\tilde{P}_1|^{\frac{1}{3} d_s^-}.$$

For each integer m , let us count now how many \tilde{P}_1 may satisfy

$$(10.44) \quad 2^{-m} |P_s| \geq |\tilde{P}_1| \geq 2^{-m-1} |P_s|.$$

As Q is transverse to P_1 but not to \tilde{P}_1 , we must have, by Proposition 10:

$$(10.45) \quad \delta(Q, \tilde{P}_1) \leq C 2^{-m(1-\eta)}$$

which shows that there are no more than $C 2^{m\eta}$ such \tilde{P}_1 's. This implies that the sum (10.43) is at most of order $C \varepsilon_0^{\frac{1}{3} d_s^-}$, which yields the statement of the lemma. \square

Remark.

1. In the lemma, the value $d_s^- = d_s^0 - C\varepsilon_0$ that has been used to define $\|P\|$ is irrelevant. The assertion of the lemma is still true if we replace d_s^- by any positive number bounded away from 0.

2. In the proposition, we can replace d_s^- by a slightly lower value (assuming, as usual, that ε_0 is arbitrarily small): if we take

$$(10.46) \quad \widehat{d}_s^- = d_s^0 - o(\log \kappa^{-1}),$$

the same argument works and the result of the proposition is still valid.

Corollary 14. *Let $\varepsilon_1 > 0$. If ε_0 is small enough, we have*

$$\sum n|P|^{d_s^-} < \varepsilon_1$$

where the sum in the left-hand side is over non-trivial primes (P, Q, n) .

Proof. Let (P, Q, n) be a non trivial prime. We have, by Proposition 13 of Subsection 7.1

$$(10.47) \quad n \leq \left(\log(C|P|^{-1}) \right)^{\frac{\log 2}{\log 3/2}}.$$

Consequently, we have

$$(10.48) \quad n|P|^{d_s^-} \leq |P|^{\widehat{d}_s^-},$$

with \widehat{d}_s^- as in (10.46) and $d_s^0 - \widehat{d}_s^-$ being independent of ε_0 (as P is a non trivial prime, one has $|P| < \varepsilon_0$). Observe also that the thinnest $(\tilde{P}, \tilde{Q}, \tilde{n}) \in \mathcal{R}(I_0)$ with $P \subset \tilde{P}$ satisfies $\tilde{Q} \subset Q_u$ hence $|\tilde{P}| \leq \varepsilon_0^\alpha$ for some fixed positive α .

We apply the proposition, taking the Remark 2 above into account and using \widehat{d}_s^- instead of d_s^- ; we obtain

$$(10.49) \quad \begin{aligned} \sum n|P|^{d_s^-} &\leq \sum |P|^{\widehat{d}_s^-} \\ &= \kappa^{-1} \sum \|P\| \\ &\leq C \kappa^{-1} (1 - \kappa^{\frac{1}{2}})^{-1} \sum \|\tilde{P}\|, \end{aligned}$$

where, in the last sum, $(\tilde{P}, \tilde{Q}, \tilde{n})$ runs through the elements of $\mathcal{R}(I_0)$ with $|\tilde{P}|$ of the order of ε_0^α . We have

$$(10.50) \quad \begin{aligned} \sum \|\tilde{P}\| &\leq C \sum \varepsilon_0^{\alpha \widehat{d}_s^-} \kappa^{C^{-1} \log \varepsilon_0^{-1}} \\ &\leq C \varepsilon_0^{\alpha(d_s^- - d_s^0) + C^{-1} \log \kappa^{-1}}, \end{aligned}$$

and the exponent is positive from (10.46). Putting this into (10.49) yields Corollary 14. \square

Remark. We do not know the range of d for which

$$\sum_{\mathcal{P}} |P|^d$$

is convergent. The Corollary shows at least that there are relatively few primes: the sum

$$\sum_{\mathcal{R}(I_0)} |P|^{d_s^-} \quad \text{is convergent.}$$

10.4 Stable Curves

Let $(P_k, Q_k, n_k)_{k \geq 0}$ be a sequence of elements of \mathcal{R} such that P_{k+1} is strictly contained in P_k for $k \geq 0$. Let R_a be the rectangle of the Markov partition which contains P_0 . The vertical part of the boundary of P_k is the union of two graphs $\{x_a = \varphi_k^\pm(y_a)\}$, and the φ_k^\pm are uniformly bounded in the C^2 topology. Moreover, there exists $\kappa^* \in (0, 1)$ such that $|P_{k+1}| \leq \kappa^* |P_k|$ for all $k \geq 0$. It follows that both sequences φ_k^\pm converge in the C^1 topology to the same limit φ_∞ , which is of class $C^{1+\text{Lip}}$, where Lip stands for Lipschitz. We state this as

Proposition 39. *The intersection $\bigcap_{k \geq 0} P_k$ is the graph $\{x_a = \varphi_\infty(y_a)\}$ of a $C^{1+\text{Lip}}$ function. Moreover, the $C^{1+\text{Lip}}$ norm of φ_∞ is bounded independently of the sequence $(P_k)_{k \geq 0}$.*

Definition.

1. A *stable curve* is the intersection $\omega = \bigcap_{k \geq 0} P_k$ of a decreasing sequence of vertical-like rectangles as above. An *unstable curve* is the intersection $\omega' = \bigcap_{k \geq 0} Q'_k$ of a decreasing sequence of horizontal-like strips.
2. The *set* of stable curves, resp. unstable curves, is denoted by \mathcal{R}_+^∞ , resp. \mathcal{R}_-^∞ . The *union* of stable curves, resp. unstable curves, is denoted by $\tilde{\mathcal{R}}_+^\infty$, resp. $\tilde{\mathcal{R}}_-^\infty$.
3. Any stable curve $\omega \subset R_a$ has a *canonical defining sequence* characterized by the following conditions: $P_0 = R_a$ and, for each k , P_{k+1} is a child of P_k .
4. Two stable curves are equal or disjoint. Hence there is a *canonical projection*

$$\pi : \tilde{\mathcal{R}}_+^\infty \mapsto \mathcal{R}_+^\infty.$$

We will now define dynamics on a part of the sets $\mathcal{R}_+^\infty, \tilde{\mathcal{R}}_+^\infty$.

Let \mathcal{N}_+ be the set of stable curves ω which are contained in infinitely many prime elements and let \mathcal{D}_+ be the complementary subset in \mathcal{R}_+^∞ . For $(P, Q, n) \in \mathcal{P}$, denote by $\mathcal{R}_+^\infty(P)$ the set of stable curves $\omega \in \mathcal{D}_+$ such that P is the thinnest prime containing ω .

We, thus, have partitions

$$(10.51) \quad \mathcal{R}_+^\infty = \mathcal{N}_+ \bigsqcup \mathcal{D}_+,$$

$$(10.52) \quad \mathcal{D}_+ = \bigsqcup_{\mathcal{P}} \mathcal{R}_+^\infty(P).$$

We denote by $\tilde{\mathcal{N}}_+$, $\tilde{\mathcal{D}}_+$, $\tilde{\mathcal{R}}_+^\infty(P)$ the respective pre-images by π .

Let $(P, Q, n) \in \mathcal{P}$, $\omega \in \mathcal{R}_+^\infty(P)$. For any (P_k, Q_k, n_k) with $\omega \subset P_k \subset P$, we can write (cf. Remark after Proposition 36)

$$(10.53) \quad (P_k, Q_k, n_k) = (P, Q, n) * (P'_k, Q'_k, n'_k)$$

for some $(P'_k, Q'_k, n'_k) \in \mathcal{R}$; we have

$$(10.54) \quad T^+(P_k, Q_k, n_k) = (P'_k, Q'_k, n'_k)$$

and we define $\omega' = T^+(\omega)$ to be the stable curve obtained by the intersection of the P'_k when P_k decrease to ω . We have

$$(10.55) \quad g^n(P_k) \subset P'_k, \quad g^n(\omega) \subset \omega'$$

and we also define

$$(10.56) \quad \tilde{T}^+ / \tilde{\mathcal{R}}_+^\infty(P) = g^n / \tilde{\mathcal{R}}_+^\infty(P).$$

We, thus, have a commutative diagram

$$(10.57) \quad \begin{array}{ccc} \tilde{\mathcal{D}}_+ & \xrightarrow{\tilde{T}^+} & \tilde{\mathcal{R}}_+^\infty \\ \downarrow \pi & & \downarrow \pi \\ \mathcal{D}_+ & \xrightarrow{T^+} & \mathcal{R}_+^\infty \end{array}$$

We observe that for $(P, Q, n) \in \mathcal{P}$ with $Q \subset R_a$, the image $T^+(\omega)$ of any $\omega \in \mathcal{R}_+^\infty(P)$ is contained in R_a .

Conversely, let $(P, Q, n) \in \mathcal{P}$ with $Q \subset R_a$ and let $\omega' \in \mathcal{R}_+^\infty$, $\omega' \subset R_a$. For any (P'_k, Q'_k, n'_k) with $\omega \subset P'_k$, we define (P_k, Q_k, n_k) by (10.53); the intersection ω of the P'_k s, when P'_k decrease to ω' , is the unique stable curve in $\mathcal{R}_+^\infty(P)$ such that $T^+(\omega) = \omega'$.

Thus, T^+ induces a bijection from $\mathcal{R}_+^\infty(P)$ on the set $\mathcal{R}_+^\infty(a)$ of stable curves contained in R_a . For $\omega \in \mathcal{R}_+^\infty(P)$, we have

$$(10.58) \quad \tilde{T}^+(\omega) = \omega' \cap Q.$$

10.5 Topology and Geometry of \mathcal{R}_+^∞ and $\tilde{\mathcal{R}}_+^\infty$

Each stable curve is a compact subset of $R = \cup R_a$. Therefore, \mathcal{R}_+^∞ may be viewed as a subset of the set of non empty compact subsets of R endowed with the Hausdorff topology. The topology induced on \mathcal{R}_+^∞ can also be viewed directly: for any $\omega = \cap P_k$ in \mathcal{R}_+^∞ , a basis of neighbourhoods of ω is obtained by considering for each k the set V_k of stable curves contained in P_k .

Equipped with this topology, \mathcal{R}_+^∞ is a Cantor set. Each $\mathcal{R}_+^\infty(P)$, $P \in \mathcal{P}$, is a closed subset, and also a Cantor set. The restriction of T^+ to each $\mathcal{R}_+^\infty(P)$ is a homeomorphism onto $\mathcal{R}_+^\infty(a)$ (with $Q \subset R_a$).

However, the subset \mathcal{N}_+ may be dense and the map T^+ in general is not continuous on the whole of \mathcal{D}_+ . We will see in the sequel that \mathcal{N}_+ is, in some appropriate sense, negligible.

For each $\omega \in \mathcal{R}_+^\infty(a)$, we denote by φ_ω the $C^{1+\text{Lip}}$ map such that $\omega = \{x_a = \varphi_\omega(y_a)\}$; for each $a \in \mathcal{a}$, each $y_a^0 \in I_a^u$, the map

$$(10.59) \quad \begin{aligned} \phi_{y_a^0} : \mathcal{R}_+^\infty(a) &\mapsto I_a^s \\ \omega &\mapsto \varphi_\omega(y_a^0) \end{aligned}$$

is a homeomorphism onto its image. Letting y_a^0 vary, we get an homeomorphism from $\mathcal{R}_+^\infty(a) \times I_a^u$ onto $\tilde{\mathcal{R}}_+^\infty(a)$.

The regularity of the partial foliation $\tilde{\mathcal{R}}_+^\infty(a)$ is given by

Proposition 40. *There exists $C > 0$ such that, for all $a \in \mathcal{a}$, all distinct $\omega, \omega' \in \mathcal{R}_+^\infty(a)$, all $y, y' \in I_a^u$, (cf. Subsection 2.1) we have*

$$(10.60) \quad \left| \log \frac{\varphi_\omega(y) - \varphi_{\omega'}(y)}{\varphi_\omega(y') - \varphi_{\omega'}(y')} \right| \leq C |y - y'|.$$

In particular, the homeomorphisms $\phi_{y'} \circ \phi_y^{-1}$ are bi-Lipschitzian, uniformly in y, y' .

Proof. The calculations that support the proof will be found in Appendix B. Write $\omega = \cap P_k$, $\omega' = \cap P'_k$, where $(P_k)_{k \geq 0}, (P'_k)_{k \geq 0}$ are the canonical sequences associated with ω, ω' . We write φ_k, φ'_k for $\varphi_k^+, \varphi_k'^+$. Let ℓ be the largest integer such that $P_\ell = P'_\ell$. There are two cases.

Case 1. $P_{\ell+1}$ or $P'_{\ell+1}$ is a simple child of P_ℓ . Let (A, B) be the implicit representation for (P_ℓ, Q_ℓ, n_ℓ) . We have in this case, for all $y \in I_a^u$

$$(10.61) \quad C^{-1}|P_\ell| \leq |\varphi_\omega(y) - \varphi_{\omega'}(y)| \leq C|P_\ell|.$$

To prove (10.60), we observe (cf. Appendix B) that we can write, for $k > \ell$,

$$(10.62) \quad \begin{aligned} \varphi_k(y) &= A(y, \psi_k(y)), \\ \varphi'_k(y) &= A(y, \psi'_k(y)), \end{aligned}$$

with ψ_k, ψ'_k uniformly C^1 -bounded and

$$(10.63) \quad |\psi_k(y) - \psi'_k(y)| \geq C > 0.$$

We then have,

$$(10.64) \quad \varphi'_k(y) - \varphi_k(y) = \int_0^1 A_x\left(y, (1-t)\psi_k(y) + t\psi'_k(y)\right) dt [\psi'_k(y) - \psi_k(y)].$$

From (10.64) we deduce (10.60) with $\varphi_\omega, \varphi_{\omega'}$ replaced by φ_k, φ'_k . We then let k goes to $+\infty$ to obtain (10.60) for $\varphi_\omega, \varphi_{\omega'}$.

Case 2. $P_{\ell+1}$ and $P'_{\ell+1}$ are non-simple children of P_ℓ . In this case (10.61) does not hold. The computation in Appendix B shows that

$$(10.65) \quad \left| \log \frac{\varphi_{\ell+1}(y) - \varphi'_{\ell+1}(y)}{\varphi_{\ell+1}(y') - \varphi'_{\ell+1}(y')} \right| \leq C |y - y'|.$$

By the same proof as in Case 1, we have

$$(10.66) \quad \left| \log \frac{\varphi_{\ell+1}(y) - \varphi_\omega(y)}{\varphi_{\ell+1}(y') - \varphi_\omega(y')} \right| \leq C |y - y'|,$$

$$(10.67) \quad \left| \log \frac{\varphi'_{\ell+1}(y) - \varphi'_\omega(y)}{\varphi'_{\ell+1}(y') - \varphi'_\omega(y')} \right| \leq C |y - y'|.$$

As we have

$$(10.68) \quad |\varphi_{\ell+1}(y) - \varphi'_{\ell+1}(y)| \leq C |\varphi_\omega(y) - \varphi_{\omega'}(y)|,$$

$$(10.69) \quad |\varphi_{\ell+1}(y) - \varphi_\omega(y)| \leq C |\varphi_\omega(y) - \varphi_{\omega'}(y)|,$$

$$(10.70) \quad |\varphi'_{\ell+1}(y) - \varphi_{\omega'}(y)| \leq C |\varphi_\omega(y) - \varphi_{\omega'}(y)|,$$

the inequality in the proposition follows. \square

The result of Proposition 40 implies that the transverse Hausdorff dimension $d_s = d_s(g)$ of $\tilde{\mathcal{R}}_+^\infty$ is well-defined, being equal to the Hausdorff dimension of $\phi_y(\mathcal{R}_+^\infty(a))$ for any $a \in \mathcal{a}, y \in I_a^u$. We have just proved that it does not depend on y . That it does not depend on a is seen as follows: for $(a, a') \in \mathcal{B}$, g sends $\tilde{\mathcal{R}}_+^\infty \cap P_{aa'}$ into $\tilde{\mathcal{R}}_+^\infty \cap R_{a'}$; the transverse Hausdorff dimension of $\tilde{\mathcal{R}}_+^\infty \cap R_{a'}$ is therefore not smaller than that of $\tilde{\mathcal{R}}_+^\infty \cap R_a$; as this is true for all $(a, a') \in \mathcal{B}$, the conclusion follows.

We will also identify below in this section the transverse Hausdorff dimension d_s through a transfer operator in the classical manner of Bowen, Ruelle and Sinai.

Another partial control on the geometry of $\tilde{\mathcal{R}}_+^\infty$ is obtained via the usual estimate on the graph transform, in the setting of uniformly hyperbolic dynamics. Let $\omega, \hat{\omega}$ be two stable curves and $j > 0$. Assume that ω and $\hat{\omega}$ belong to the domain of $(T^+)^j$. We say that ω and $\hat{\omega}$ belong to the same component of the domain of $(T^+)^j$ if, for each $0 \leq i < j$, there exists a prime P_i such that $(T^+)^i(\omega)$ and $(T^+)^i(\hat{\omega})$ belong to $\mathcal{R}_+^\infty(P_i)$.

Proposition 41. *There exists $\theta_0 \in (0, 1)$ such that, if $\omega, \widehat{\omega}$ belong to the same component of the domain of $(T^+)^j$, we have*

$$|D\varphi_\omega(y) - D\varphi_{\widehat{\omega}}(y)| \leq C\theta_0^j$$

for all y .

Proof. We leave the proof of this very standard estimate to the reader. □

10.6 Transverse Dilatation

This subsection is a preparation for the definition of a transfer operator in the next subsection. The weight function in this transfer operator is, up to a coboundary term, given by a transverse dilatation.

Let $(P, Q, n) \in \mathcal{P}$, $\omega = \{x = \varphi_\omega(y)\}$ a stable curve in $\mathcal{R}_+^\infty(P)$, $\omega' = T^+(\omega)$ its image. Let (A, B) be the implicit representation of (P, Q, n) .

For $z = (\varphi_\omega(y), y) \in \omega$, let

$$(10.71) \quad v_\omega(z) = \frac{\partial}{\partial y} + D\varphi_\omega(y) \frac{\partial}{\partial x}$$

be the normalized tangent vector to ω at z .

The matrix of $D\widetilde{T}_+$ at z , computed in the bases $(\frac{\partial}{\partial x}, v_\omega(z))$ at z , $(\frac{\partial}{\partial x}, v_{\omega'}(z'))$ at $z' = \widetilde{T}^+(z)$, is lower triangular; the first diagonal coefficient is

$$(10.72) \quad A_x^{-1}(y, x') \left(1 - B_x(y, x') D\varphi_{\omega'}(y') \right)$$

We denote by $\widetilde{b}(z)$ the logarithm of the absolute value of this coefficient. As φ_ω is $C^{1+\text{Lip}}$ uniformly in ω and $g^n : P \mapsto Q$ has bounded distortion, we have, for all $z, z^* \in \omega$

$$(10.73) \quad |\widetilde{b}(z) - \widetilde{b}(z^*)| \leq C|z - z^*|.$$

Let $j > 0$ and let $\omega, \widehat{\omega}$ be stable curves which belong to the same component of the domain of $(T^+)^j$. Let z, \widehat{z} be point of $\omega, \widehat{\omega}$, respectively, with the same y coordinate. It follows from Proposition 41 that one has

$$(10.74) \quad |\widetilde{b}(z) - \widetilde{b}(\widehat{z})| \leq C\theta_0^j.$$

We also have, from the definition of \widetilde{b} :

$$(10.75) \quad \widetilde{b}(z) + \log |P| \leq C.$$

We want to get rid of the dependence of \tilde{b} on the y coordinate along ω by adding a coboundary term. We define

$$(10.76) \quad \mathcal{D}_+^\infty = \bigcap_{j \geq 0} \text{Dom}(T^+)^j = \bigcap_{j \geq 0} (T^+)^{-j}(\mathcal{D}_+).$$

For each a , fix some $y_a^0 \in I_a^u$. Then, for $\omega \in \mathcal{D}_+^\infty$, $\omega \subset R_a$, $z \in \omega$, define

$$(10.77) \quad \Delta b(z) = \sum_{i \geq 0} \left(\tilde{b}((\tilde{T}^+)^i(z)) - \tilde{b}((\tilde{T}^+)^i(z^0)) \right),$$

where z^0 is the point on ω with vertical coordinate equal to y_a^0 .

From the cone condition, we have, for $i \geq 0$:

$$(10.78) \quad \|(\tilde{T}^+)^i(z) - (\tilde{T}^+)^i(z^0)\| \leq C\lambda^{-i}.$$

The series defining Δb is uniformly convergent from (10.73), (10.78), and Δb is bounded on $\tilde{\mathcal{D}}_+^\infty := \pi^{-1}(\mathcal{D}_+^\infty)$.

Write z^1 for the point on $T_+(\omega) \in R_{a'}$ with vertical coordinate $y_{a'}^0$. We have

$$(10.79) \quad \Delta b(z) - \Delta b(\tilde{T}^+(z)) = \tilde{b}(z) - b(\omega),$$

with

$$(10.80) \quad b(\omega) = \tilde{b}(z^0) + \sum_{i \geq 0} \left[\tilde{b}((\tilde{T}^+)^{i+1}(z^0)) - \tilde{b}((\tilde{T}^+)^i(z^1)) \right].$$

We call b the (logarithmic) mean transverse dilatation.

Proposition 42. *The mean transverse dilatation b , which differs from \tilde{b} on \mathcal{D}_+^∞ by the coboundary of the bounded function Δb , satisfies*

$$(10.81) \quad |b(\omega) - b(\hat{\omega})| \leq C\theta_1^j$$

if $\omega, \hat{\omega}$ belong to the same component of the domain of $(T^+)^j$. Here θ_1 is a fixed constant in $(0, 1)$ larger than θ_0 .

Proof. We have only to prove (10.81). Let $z^0 \in \omega$, $z^1 \in T^+(\omega)$ as above and let $\hat{z}^0 \in \hat{\omega}$, $\hat{z}^1 \in T^+(\hat{\omega})$ be similarly defined. We have, for $i \geq 0$:

$$(10.82) \quad |\tilde{b}((\tilde{T}^+)^{i+1}(z^0)) - \tilde{b}((\tilde{T}^+)^i(z^1))| \leq C\lambda^{-i},$$

$$(10.83) \quad |\tilde{b}((\tilde{T}^+)^{i+1}(\hat{z}^0)) - \tilde{b}((\tilde{T}^+)^i(\hat{z}^1))| \leq C\lambda^{-i},$$

From (10.74), we also have

$$(10.84) \quad |\tilde{b}(z^0) - \tilde{b}(\hat{z}^0)| \leq C\theta_0^j.$$

For $0 \leq i < j - 1$, we compare $\tilde{b}((\tilde{T}^+)^{i+1}(z^0))$ and $\tilde{b}((\tilde{T}^+)^{i+1}(\hat{z}^0))$ as follows: let \hat{z}^{i+1} be the point on $T_+^{i+1}(\hat{\omega})$ with the same y -coordinate as $(\tilde{T}^+)^{i+1}(z^0)$; one has

$$(10.85) \quad \left| \tilde{b}(\hat{z}^{i+1}) - \tilde{b}((\tilde{T}^+)^{i+1}(z^0)) \right| \leq C \theta_0^{j-i-1},$$

$$(10.86) \quad \|\hat{z}^{i+1} - (\tilde{T}^+)^{i+1}(\hat{z}^0)\| \leq C \lambda^{i+1-j}.$$

Dealing in the same way with the terms involving z^1, \hat{z}^1 , we obtain (10.81) with

$$(10.87) \quad \theta_1 = \max(\theta_0^{1/2}, \lambda^{-1/2})$$

proving the proposition. □

10.7 Definition of a Transverse Operator

As Δb is bounded, it follows from (10.75) that

$$(10.88) \quad |b(\omega) + \log |P|| \leq C$$

for all $(P, Q, n) \in \mathcal{P}$, $\omega \in \mathcal{R}_+^\infty(P)$.

It is then a consequence of Corollary 14 in Subsection 10.3 that the series

$$(10.89) \quad \sum_{T^+\omega'=\omega} \exp(-db(\omega'))$$

over pre-images ω' of a given stable curve ω is converging, uniformly in ω , for $d \geq d_s^-$. We will, therefore, define a transfer operator L_d for $d \geq d_s^-$ as follows: for a bounded function h defined on \mathcal{D}_+^∞ , for $\omega \in \mathcal{D}_+^\infty$, we set

$$(10.90) \quad L_d h(\omega) = \sum_{T^+\omega'=\omega} \exp(-db(\omega')) h(\omega').$$

We can also view this sum over pre-images as a sum over inverse branches of T^+ , which are in one-to-one correspondence with the primes (P, Q, n) such that Q and ω belong to the same rectangle R_a . Accordingly, we split the series in two parts: a finite sum corresponding to the trivial primes, (cf. Subsection 10.1), which we denote by L_d^0 and which is defined for all values of d , and a perturbative term which we denote by ΔL_d . The formula (10.90) defines a bounded operator from the space of bounded functions on \mathcal{D}_+^∞ into itself, but to have nice spectral properties we need, as usual, to restrict to spaces of slightly more regular functions.

Let θ be a constant with

$$(10.91) \quad \theta_1 < \theta < 1$$

where θ_1 comes from Proposition 42 and satisfies $\theta_1 > \lambda^{-1}$ (cf. 10.87). Denote by E the space of bounded functions h on \mathcal{D}_+^∞ which satisfy, for some constant $C > 0$,

$$(10.92) \quad |h(\omega) - h(\widehat{\omega})| \leq C\theta^j$$

whenever $\omega, \widehat{\omega}$ belong to the same component of the domain of $(T^+)^j$. We denote by $\|h\|_\infty$ the usual norm on bounded functions, by $|h|_E$ the best possible C and set

$$(10.93) \quad \|h\|_E = \max(|h|_E, \|h\|_\infty).$$

It is clear that E is a Banach space.

Proposition 43. *For $d \geq d_s^-$, L_d restricts to a bounded operator on E . Moreover, the norm of the perturbative part ΔL_d is as small as we want if ε_0 is small enough.*

Proof. Let $h \in E$, $\omega, \widehat{\omega} \in \mathcal{D}_+^\infty$, $j > 0$. Assume that $\omega, \widehat{\omega}$ belong to the same component of the domain of $(T^+)^j$. Let (P, Q, n) be a prime such that $Q, \omega, \widehat{\omega}$ belong to the same rectangle R_a , and let $\omega_1, \widehat{\omega}_1$ be the inverse images of $\omega, \widehat{\omega}$ by T^+ corresponding to this inverse branch. By the definition of $| \cdot |_E$, we have

$$(10.94) \quad |h(\omega_1) - h(\widehat{\omega}_1)| \leq |h|_E \theta^{j+1}.$$

From Proposition 42, we have

$$(10.95) \quad |b(\omega_1) - b(\widehat{\omega}_1)| \leq C\theta_1^{j+1}.$$

It follows from (10.88) that

$$(10.96) \quad |\exp(-db(\omega_1)) - \exp(-db(\widehat{\omega}_1))| \leq C(d) |P|^d \theta_1^{j+1}.$$

Putting together (10.94) and (10.96), we have

$$(10.97) \quad |h(\omega_1)\exp(-db(\omega_1)) - h(\widehat{\omega}_1)\exp(-db(\widehat{\omega}_1))| \leq C(d) |P|^d (\theta^{j+1}|h|_E + \theta_1^{j+1}\|h\|_\infty).$$

Summing over (non trivial) primes yields for $d \geq d_s^-$:

$$(10.98) \quad |\Delta L_d h|_E < \varepsilon_1 \|h\|_E,$$

$$(10.99) \quad |L_d h|_E < C \|h\|_E,$$

where ε_1 can be made arbitrarily small if ε_0 is small enough, according to Corollary 14. The same estimates (for $d \geq d_s^-$) for $\|\Delta L_d h\|_\infty$ and $\|L_d h\|_\infty$ are easier and can be seen directly. The proposition follows. \square

10.8 Spectral Properties of the Transfer Operator

Let us denote by $\mathcal{R}_+^\infty(K)$ the set of stable curves ω which are intersections of a sequence of rectangles belonging to $\mathcal{R}(I_0)$; these stable curves are precisely those which meet the initial horseshoe K .

Observe that $\mathcal{R}_+^\infty(K) \subset \mathcal{D}_+^\infty$. Denote by E_K the space of bounded functions h on $\mathcal{R}_+^\infty(K)$ which satisfy

$$(10.100) \quad |h(\omega) - h(\widehat{\omega})| \leq C\theta^j,$$

whenever $\omega, \widehat{\omega}$ belong to the same component of the domain of $(T^+)^j$. Define $\|h\|_{E_K}, \|h\|_{E_K}$ as above, which makes E_K a Banach space.

Let $h \in E$; the restriction of h to $\mathcal{R}_+^\infty(K)$ belongs to E_K and we have

$$(10.101) \quad \|h/\mathcal{R}_+^\infty(K)\|_{E_K} \leq \|h\|_E.$$

The formula for L_d^0 defines a bounded operator, still denoted by L_d^0 , on E_K and we have a commutative diagram

$$(10.102) \quad \begin{array}{ccc} E & \xrightarrow{L_d^0} & E \\ r \downarrow & & \downarrow r \\ E_K & \xrightarrow{L_d^0} & E_K \end{array}$$

where $r : E \mapsto E_K$ is the restriction operator. The bounded operator $L_d^0 : E_K \mapsto E_K$ is the subject of the classical theory by Bowen, Ruelle, Sinai for uniformly hyperbolic systems.

Let us recall some standard results of this theory.

a) There is a direct sum invariant decomposition

$$(10.103) \quad E_K = \mathbb{R} h'_d \oplus H'_d$$

depending analytically on the parameter d , such that h'_d is a positive eigenfunction, with associated eigenvalue $\lambda'_d > 0$, and such that

$$(10.104) \quad sp(L_d^0/H'_d) \subset \{|z| < \lambda'_d\}.$$

b) There exists a (unique) probability measure μ'_d on $\mathcal{R}_+^\infty(K)$ such that

$$(10.105) \quad H'_d = \{h \in E_K, \int h d\mu'_d = 0\}.$$

One normalizes h'_d to have $\int h'_d d\mu'_d = 1$. Then, the probability measure $\nu'_d = h'_d \mu'_d$ is invariant under the restriction of T^+ to $\mathcal{R}_+^\infty(K)$ (observe that \widetilde{T}^+ on $\widetilde{\mathcal{R}}_+^\infty(K)$ is just the restriction of g).

Let E^0 be the kernel of the restriction operator $r : E \mapsto E_K$. It is invariant under L_d^0 .

Lemma 14. *One has, for all $d \in \mathbb{R}$.*

$$sp(L_d^0/E^0) \subset \{|z| \leq \theta \lambda'_d\}.$$

Proof. Let $h \in E^0$, $j \geq 0$. We have

$$(10.106) \quad (L_d^0)^j h(\omega) = \sum_{(T^+)^j(\omega')=\omega}^0 h(\omega') \exp(-db^{(j)}(\omega')),$$

where the symbol \sum^0 indicates that we only consider inverse branches of T^+ associated with trivial primes. The notation $b^{(j)}$ denotes the Birkhoff sum

$$(10.107) \quad b^{(j)}(\omega') = \sum_{0 \leq i < j} b((T^+)^i(\omega')).$$

We observe that in the sum in (10.106), each ω' belongs to the same component of the domain of $(T^+)^j$ as a stable curve in $\mathcal{R}_+^\infty(K)$. As h belongs to E^0 , this implies that for such a ω' we have

$$(10.108) \quad |h(\omega')| \leq |h|_E \theta^j.$$

On the other hand, we have

$$(10.109) \quad \sum^0 \exp(-db^{(j)}(\omega')) \leq C \lambda'_d{}^j,$$

and it follows that

$$(10.110) \quad \|(L_d^0)^j h\|_\infty \leq C \lambda'_d{}^j \theta^j \|h\|_E.$$

Let $\widehat{\omega} \in \mathcal{R}_+^\infty$ belong to the same component of the domain of $(T^+)^j$ as ω . Denote by $\widehat{\omega}'$ the inverse image of $\widehat{\omega}$ associated to the same sequence of trivial primes as ω' . We have

$$(10.111) \quad |h(\omega') - h(\widehat{\omega}')| \leq |h|_E \theta^{j+\ell},$$

and, from Proposition 42

$$(10.112) \quad |b^{(j)}(\omega') - b^{(j)}(\widehat{\omega}')| \leq C \theta_1^\ell.$$

Using also (10.108) and (10.109), we obtain

$$(10.113) \quad |(L_d^0)^j h(\omega) - (L_d^0)^j h(\widehat{\omega})| \leq C \theta^{j+\ell} \lambda'_d{}^j \|h\|_E,$$

which implies the statement of the Lemma. □

We deduce from Lemma 14 that there is a unique function in E , still denoted by h'_d , which restricts to h'_d on $\mathcal{R}_+^\infty(K)$ and satisfies

$$(10.114) \quad L_d^0(h'_d) = \lambda'_d h'_d.$$

Moreover, defining a supplementary hyperplane by

$$(10.115) \quad H_d'' = r^{-1}(H_d') \oplus E^0,$$

we have that H_d'' is invariant under L_d^0 and

$$(10.116) \quad sp(L_d^0/H_d'') \subset \{|z| \leq \lambda_d''\},$$

where $\lambda_d'' < \lambda_d'$ is independent of ε_0 .

Using Proposition 43, we now consider L_d itself, assuming that ε_0 is small enough and $d \geq d_s^-$.

As the norm of the perturbation part ΔL_d is arbitrarily small, we conclude that L_d has a positive eigenfunction h_d , with associated eigenvalue λ_d arbitrarily close to λ_d' , and an invariant supplementary hyperplane H_d satisfying

$$(10.117) \quad sp(L_d/H_d) \subset \{|z| < \lambda_d\}.$$

Moreover, h_d , λ_d and H_d depend analytically on d for $d > d_s^-$ because L_d does. We check that

$$(10.118) \quad h_d \geq c^{-1} > 0.$$

Indeed, the sequence $h^{(n)} = \lambda_d^{-n} L_d^n(1)$ converge to a positive multiple of h_d . We have

$$(10.119) \quad h^{(n)}(\omega) = \sum_{(T^+)^n(\omega')=\omega} \exp(-db^{(n)}(\omega')).$$

Let $\omega, \widehat{\omega}$ be elements of \mathcal{D}_+^∞ in the same rectangle R_a ; let $\omega', \widehat{\omega}'$ be pre-images of $\omega, \widehat{\omega}$ by $(T^+)^n$ associated with the same sequence of primes. We have (cf. 10.112)

$$(10.120) \quad |b^{(n)}(\omega') - b^{(n)}(\widehat{\omega}')| \leq C,$$

and it follows that

$$(10.121) \quad C^{-1} \leq (h^{(n)}(\omega'))^{-1} h^{(n)}(\omega) \leq C.$$

This implies (10.118). One normalizes h_d in order to have

$$(10.122) \quad h_d = \lim_{n \rightarrow +\infty} \lambda_d^{-n} L_d^n(1).$$

Denote then by μ_d the linear form on E with kernel H_d normalized by $\mu_d(h_d) = 1$. We have, for all $h \in E$

$$(10.123) \quad \lim_{n \rightarrow \infty} \lambda_d^{-n} L_d^n h = \mu_d(h) h_d.$$

As L_d is a positive operator, μ_d is positive. Observe also that for all $(P, Q, n) \in \mathcal{R}$, the characteristic function χ_P (equal to 1 if $\omega \subset P$, 0 otherwise) belongs to E and satisfies $L^n \chi_P > 0$ everywhere for some $n > 0$. Therefore, there exists a unique probability measure on \mathcal{R}_+^∞ , still denoted by μ_d , which coincides with μ_d on the intersection of E with $C(\mathcal{R}_+^\infty)$.

10.9 The Gibbs Measure

From the defining property (10.123) of μ_d , we have, for all $h \in E$

$$(10.124) \quad \mu_d(L_d h) = \lambda_d \mu_d(h).$$

We will now check the classical Jacobian property for μ_d .

Let $(P, Q, n) \in \mathcal{R}$. Let $P_1 \supset P$ be the thinnest prime containing P . The application T^+ is injective on the set $\mathcal{R}_+^\infty(P)$ formed by the $\omega \in \mathcal{R}_+^\infty(P_1)$ which are contained in P . The image of this set by T^+ is exactly the set of stable curves contained in P' , with $T^+(P, Q, n) = (P', Q', n')$.

Let h be a function in E which vanishes outside $\mathcal{R}_+^\infty(P)$. Then, $L_d h$ vanishes on any curve not contained in P' , and satisfies

$$(10.125) \quad L_d h(T^+ \omega) = h(\omega) \exp(-db(\omega))$$

for $\omega \in \mathcal{R}_+^\infty(P)$. The relation (10.124) for h is the Jacobian property.

Consider in particular the case where h is the characteristic function of $\mathcal{R}_+^\infty(P)$. We then obtain

$$(10.126) \quad \lambda_d \mu_d(\mathcal{R}_+^\infty(P)) = \int_{P'} \exp(-db((T^+)^{-1} \omega)) d\mu_d(\omega).$$

We now will specify the value of d by asking that

$$(10.127) \quad \lambda_d = 1.$$

Indeed, we have, for $d \geq d_s^-$

$$(10.128) \quad \frac{\partial}{\partial d} \lambda_d < 0.$$

This follows from the formula for L_d and the fact that $b^{(n)}$ increases at least linearly with n (cf. (10.107)).

We also see easily that

$$(10.129) \quad \lim_{d \rightarrow +\infty} \lambda_d = 0.$$

Finally, we have

$$(10.130) \quad \lambda_{d_s^-} > 1.$$

Indeed, d_s^- was chosen in order to be smaller than the transverse Hausdorff dimension of $W^s(K)$. This means that the eigenvalue $\lambda'_{d_s^-}$ for $\mathbf{L}_{d_s^-}^0$ on E_K satisfies $\lambda'_{d_s^-} > 1$. As $\Delta L_{d_s^-}$ is also a nonnegative operator, we have $\lambda_{d_s^-} \geq \lambda'_{d_s^-}$. Therefore, (10.130) holds, and it follows from (10.128)–(10.130) that

(10.127) holds for a unique value of d . We will denote this value by d_s . We shall indeed see that d_s is the transverse Hausdorff dimension of $\widetilde{\mathcal{R}}_+^\infty$ which we were able to define in Subsection 10.5.

We just write μ for the measure μ_{d_s} and h^* for the eigenfunction h_{d_s} .

Proposition 44. *For any $(P, Q, n) \in \mathcal{R}$, we have*

$$C^{-1}|P|^{d_s} \leq \mu(\{\omega \subset P\}) \leq C|P|^{d_s}.$$

Proof. Let

$$(10.131) \quad (P, Q, n) = (P_1, Q_1, n_1) * \cdots * (P_r, Q_r, n_r)$$

be the prime decomposition of (P, Q, n) . If $\omega \in \mathcal{D}_+^\infty$ satisfies $(T^+)^i(\omega) \in \mathcal{R}_+^\infty(P_{i+1})$ for $0 \leq i < r$, we have, from (10.88), (see the definition of $b^{(r)}$ in (10.107)):

$$(10.132) \quad C^{-1}|P|^{d_s} \leq \exp(-d_s b^{(r)}(\omega)) \leq C|P|^{d_s}.$$

It, then, follows from the Jacobian property that

$$(10.133) \quad \begin{aligned} \mu(\{\omega \subset P\}) &\geq C^{-1}|P|^{d_s} \mu(\{\omega \in \mathcal{R}_+^\infty(a)\}), \\ &\geq C^{-1}|P|^{d_s}, \end{aligned}$$

where R_a is the rectangle containing Q .

For the opposite inequality, we have also to take into account the other inverse branches of T_+^r when we estimate $L_{d_s}^r(\chi_P)$, where χ_P is the characteristic function of $\{\omega \subset P\}$. For $0 \leq i \leq r$, let

$$(10.134) \quad (P^i, Q^i, n^i) = (P_{i+1}, Q_{i+1}, n_{i+1}) * \cdots * (P_r, Q_r, n_r)$$

(with $(P^r, Q^r, n^r) = (R_a, R_a, 0)$). We have

$$(10.135) \quad L_{d_s} \chi_P = \chi_P^1 + \Delta \chi_P^1$$

where

$$(10.136) \quad \chi_P^1(\omega^1) = \begin{cases} 0 & \text{if } \omega^1 \not\subset P^1 \\ \exp(-d_s b(\omega^0)) & \text{if } \omega^1 = T^+(\omega^0) \text{ for some } \omega^0 \in \mathcal{R}_+^\infty(P) \end{cases}$$

and

$$(10.137) \quad \Delta \chi_P^1 \leq C \sum |P_1^*|^{d_s},$$

where the sum runs over prime elements (P_1^*, Q_1^*, n_1^*) with P_1^* contained in P and distinct from P . By Proposition 38 in Subsection 10.3, we obtain

$$(10.138) \quad \mu(\Delta \chi_P^1) \leq C|P|^{d_s} \kappa^{\frac{r-1}{2}}.$$

If $r > 1$, we write similarly

$$(10.139) \quad L_d \chi_P^1 = \chi_P^2 + \Delta \chi_P^2,$$

where χ_P^2 is associated with the inverse branch defined by the prime P_2 and vanishes outside P^2 . The perturbative term satisfies

$$(10.140) \quad \Delta \chi_P^2 \leq C |P_1|^{d_s} \sum |P_2^*|^{d_s},$$

where the sum now is over primes P_2^* contained in P^1 and distinct from P^1 . Proposition 38 now gives

$$(10.141) \quad \begin{aligned} \mu(\Delta \chi_P^2) &\leq C |P_1|^{d_s} |P^1|^{d_s} \kappa^{\frac{r-2}{2}} \\ &\leq C |P|^{d_s} \kappa^{\frac{r-2}{2}}. \end{aligned}$$

We iterate this process. At the last step, we have from (10.132)

$$(10.142) \quad \mu(\chi_P^r) \leq C |P|^{d_s}.$$

The contribution of the perturbative terms is bounded by

$$(10.143) \quad \mu\left(\sum_1^r \chi_P^i\right) \leq C |P|^{d_s} \sum_1^r \kappa^{\frac{r-i}{2}} \leq C |P|^{d_s}.$$

□

Corollary 15. *The transverse Hausdorff dimension of $\tilde{\mathcal{R}}_+^\infty$ is $\leq d_s$. More precisely, for any curve γ which is transverse to $\tilde{\mathcal{R}}_+^\infty$, the Hausdorff measure in dimension d_s of the intersection of γ with $\tilde{\mathcal{R}}_+^\infty$ is finite.*

We will see below that the transverse Hausdorff dimension is equal to d_s .

Proof. Let $\delta > 0$, choose a finite collection of disjoint rectangles P_i with $|P_i| \leq \delta$ for each i and $\tilde{\mathcal{R}}_+^\infty \subset \cup P_i$. We have

$$(10.144) \quad \begin{aligned} 1 = \sum \mu(P_i) &\geq C^{-1} \sum |P_i|^{d_s} \\ &\geq C^{-1} \sum [\text{diam}(\gamma \cap P_i)]^{d_s} \end{aligned}$$

and the statement of the Corollary follows. □

The following statement shows that the dynamics T^+ is only undefined on a small set.

Proposition 45. *The transverse Hausdorff dimension of the set $\tilde{\mathcal{R}}_+^\infty - \tilde{\mathcal{D}}_+^\infty$ is $\leq d_s^- < d_s$.*

Proof. We have

$$(10.145) \quad \widetilde{\mathcal{R}}_+^\infty - \widetilde{\mathcal{D}}_+^\infty = \bigcup_{n \geq 0} (T^+)^{-n}(\mathcal{N}_+)$$

As each $(T^+)^n$ has countably many inverse branches which are Lipschitzian, it is sufficient to prove that the transverse Hausdorff dimension of \mathcal{N}_+ is $\leq d_s^-$. But by the definition of \mathcal{N}_+ , for any $\delta > 0$, the union of prime rectangles P with $|P| < \delta$ contains \mathcal{D} . Proposition 45 then follows from Corollary 14. \square

10.10 Transverse Hausdorff Dimension of $\widetilde{\mathcal{R}}_+^\infty$

Proposition 46. *The transverse Hausdorff dimension of $\widetilde{\mathcal{R}}_+^\infty$ is the number d_s characterized by $\lambda_{d_s} = 1$.*

Remark. *We have already seen that the Hausdorff measure in dimension d_s of the intersection of $\widetilde{\mathcal{R}}_+^\infty$ with a transverse curve is always finite. We do not know whether it is positive or always zero.*

Proof. Let γ be a smooth horizontal-like curve in some R_a . We denote by $[\gamma]$ the set of stable curves which meet γ . We will show that

$$(10.146) \quad \mu([\gamma]) \leq C(\text{diam } \gamma)^{d_s} (\log(\text{diam } \gamma)^{-1})^{C_0}.$$

This, being true for all such γ , clearly implies that the transverse Hausdorff dimension of $\widetilde{\mathcal{R}}_+^\infty$ is $\geq d_s$, which is sufficient to prove the proposition in view of Corollary 15.

Clearly, we may assume that $\mu([\gamma]) > 0$. Define $(P_0, Q_0, n_0) \in \mathcal{R}$ to be the element such that P_0 is the thinnest rectangle containing any stable curve in $[\gamma]$. There are at least two children of P_0 which contain a stable curve in $[\gamma]$. If one of these children is simple, we must have

$$(10.147) \quad \text{diam } \gamma \geq C^{-1}|P_0|$$

and also

$$(10.148) \quad \mu([\gamma]) \leq \mu(\{\omega \subset P_0\}) \leq C|P_0|^{d_s}$$

by Proposition 44, which gives the required estimate (and even better). This case is said to have complexity 0. In the remaining case, denote by $P_{0,i}$ the (non-simple) children of P_0 which contain a stable curve in $[\gamma]$. Each $P_{0,i}$ is obtained by parabolic composition:

$$(10.149) \quad (P_{0,i}, Q_{0,i}, n_{0,i}) \in (P_0, Q_0, n_0) \square (P_{0,i}^*, Q_{0,i}^*, n_{0,i}^*)$$

and the widths are related through

$$(10.150) \quad C^{-1} \leq |P_{0,i}| |P_0|^{-1} |P_{0,i}^*|^{-1} \delta(Q_0, P_{0,i}^*)^{\frac{1}{2}} \leq C.$$

Let $\delta_0 = \sup_i \delta(Q_0, P_{0,i}^*)$; let γ_1 be any horizontal-like curve with the following property: a stable curve meets γ_1 if and only if it is contained in some $P_{0,i}^*$. For $\ell \geq 0$, denote by $\gamma_{1,\ell}$ a piece of γ_1 with the property that a stable curve meets $\gamma_{1,\ell}$ iff it is contained in some $P_{0,i}^*$ with

$$(10.151) \quad \delta_0 2^{-\ell-1} \leq \delta(Q_0, P_{0,i}^*) \leq \delta_0 2^{-\ell}$$

(if there is no such $P_{0,i}^*$, take $\gamma_{1,\ell} = \emptyset$). We can now write

$$(10.152) \quad \begin{aligned} \mu([\gamma]) &\leq C \sum_i |P_{0,i}|^{d_s} \quad (\text{from Proposition 44}) \\ &\leq C |P_0|^{d_s} \sum_i |P_{0,i}^*|^{d_s} \delta(Q_0, P_{0,i}^*)^{-\frac{1}{2} d_s} \\ &\leq C |P_0|^{d_s} \delta_0^{-\frac{1}{2} d_s} \sum_{\ell \geq 0} 2^{\frac{\ell d_s}{2}} \sum^{(\ell)} |P_{0,i}^*|^{d_s} \\ &\leq C |P_0|^{d_s} \delta_0^{-\frac{1}{2} d_s} \sum_{\ell \geq 0} 2^{\frac{\ell d_s}{2}} \mu([\gamma_{1,\ell}]) \end{aligned}$$

again by Proposition 44. We have written $\sum^{(\ell)}$ for the partial sum over those $P_{0,i}^*$ satisfying (10.151). Assume for some constant $A > 0$, that we have, for each $\ell \geq 0$:

$$(10.153) \quad \mu([\gamma_{1,\ell}]) \leq A(\text{diam } \gamma_{1,\ell})^{d_s}.$$

Observe that we have

$$(10.154) \quad \text{diam } \gamma_{1,\ell} \leq C 2^{-\ell} \text{diam } \gamma_1,$$

$$(10.155) \quad \text{diam } \gamma \geq C^{-1} \delta_0^{-\frac{1}{2}} |P_0| \text{diam } \gamma_1.$$

Making use of (10.153)–(10.155) in (10.152) yields

$$(10.156) \quad \mu([\gamma]) \leq AC(\text{diam } \gamma)^{d_s},$$

which is of the same form as (10.153), but with a worse constant AC instead of A .

To obtain (10.146), it is thus sufficient to define a complexity index $c(\gamma) \in \mathbb{N}$ which satisfies

$$(10.157) \quad c(\gamma) \leq c \log \log |P_0|^{-1},$$

$$(10.158) \quad c(\gamma) = 1 + \sup_{\ell} c(\gamma_{1,\ell}),$$

the case of complexity 0 having already been defined and dealt with.

We want to use (10.158) to give an inductive definition of $c(\gamma)$. This will work if the $\gamma_{1,\ell}$ are in some sense "simpler" than γ . If all $\gamma_{1,\ell}$ have complexity 0, we just set $c(\gamma) = 1$. Assume therefore that some $\gamma_{1,\ell}$ has complexity > 0 . This means that there exists an element $(P_{1,\ell}, Q_{1,\ell}, n_{1,\ell})$ with the following properties:

- each $P_{0,i}^*$ related to $\gamma_{1,\ell}$ through (10.151) is contained in some non-simple child of $P_{1,\ell}$;
- at least two non-simple children of $P_{1,\ell}$ contain some $P_{0,i}^*$.

For any parameter interval I containing the given parameter value, $P_{1,\ell}$ is I -critical (cf. Proposition 5 in Subsection 6.4). We now distinguish two cases.

Case 1. P_0 is also I -critical (for any I as above).

Let I^* be the largest parameter interval for which we have

$$(10.159) \quad |P_0| > |I^*|^\beta.$$

Then, (P_0, Q_0, n_0) cannot be I^* -bicritical as I^* is β -regular, hence Q_0 cannot be I^* -critical. This implies

$$(10.160) \quad \delta_0 2^{-\ell} \geq C^{-1} |I^*| \geq C^{-1} |P_0|^{\beta^{-1}(1+\tau)}.$$

As $P_{1,\ell}$ is not transverse to Q_0 (because $P_{0,i}$ was a child of P_0), we must have from Proposition 10 in Subsection 6.6 that

$$(10.161) \quad |P_{1,\ell}| \geq C^{-1} (\delta_0 2^{-\ell})^{(1-\eta)^{-1}}.$$

Comparing with (10.160), this guarantees that

$$(10.162) \quad |P_0| \ll |P_{1,\ell}|^{\frac{1}{2}(1+\beta)}.$$

This means indeed that every $P_{1,\ell}$ (such that the complexity of $\gamma_{1,\ell}$ is > 0) is indeed simpler than P_0 and allows us to use (10.158) to define inductively $c(\gamma)$. Observe that the hypothesis of case 1 is always satisfied by the $P_{1,\ell}$. The inequality (10.157) follows from (10.162).

Case 2. P_0 is I -transverse for I small enough.

From Case 1, we have already defined the complexity indices $c(\gamma_{1,\ell})$ using (10.158) and again we define $c(\gamma)$ by (10.158). We have to check (10.157) in this case. This will hold if we have, for each ℓ ,

$$(10.163) \quad \log \log |P_{1,\ell}|^{-1} \leq c \log \log |P_0|^{-1}.$$

But (10.161) still holds. We also have

$$(10.164) \quad |Q_0| \ll \delta_0 2^{-\ell} \quad (\text{from (R7)}),$$

$$(10.165) \quad \log |Q_0|^{-1} \leq C n_0,$$

$$(10.166) \quad \log \log |P_0|^{-1} \geq \frac{\log \frac{3}{2}}{\log 2} \log n_0 - C,$$

from Proposition 13 of Subsection 7.1. Putting this together, we obtain (10.163). The proof of the proposition is now complete. \square

10.11 Invariant Measures

From the Gibbs measure μ , which is not invariant but has the Jacobian property, we define a T^+ -invariant measure ν on \mathcal{D}_+^∞ by

$$(10.167) \quad d\nu = h^* d\mu.$$

The measure ν is actually a probability measure on \mathcal{R}_+^∞ by Proposition 44 and Corollary 14 (see proof of Proposition 45), which implies that

$$(10.168) \quad \mu(\mathcal{R}_+^\infty - \mathcal{D}_+^\infty) = 0.$$

As h^* is bounded and bounded away from 0 (cf. (10.118)), the statement of Proposition 44 is also valid for ν instead of μ . To check that ν is indeed T_+ -invariant, we first observe that, if $h_0, h_1 \in E$, the product $h_0 h_1$ also belongs to E ; indeed we have

$$(10.169) \quad |h_0 h_1|_E \leq \|h_0\|_\infty |h_1|_E + |h_0|_E \|h_1\|_\infty.$$

In particular, for any $h \in E$, $h h^*$ also belongs to E . Let $h \in E$. We write

$$(10.170) \quad \begin{aligned} \int h(T^+\omega) d\nu(\omega) &= \int h(T^+\omega) h^*(\omega) d\mu(\omega) \\ &= \sum_P \int h(T^+\omega) h^*(\omega) \chi_P^*(\omega) d\mu(\omega), \end{aligned}$$

where χ_P^* is the characteristic function of $\mathcal{R}_+^\infty(P)$. The Jacobian property gives

$$(10.171) \quad \int h(T^+\omega) h^*(\omega) \chi_P^*(\omega) d\mu(\omega) = \int h(\omega) h^*(\omega') \exp(-d_s b(\omega')) d\mu(\omega)$$

where ω' is the image of ω under the inverse branch of T_+ associated with P . Summing over P and using that h^* is L_{d_s} -invariant gives

$$(10.172) \quad \int h(T^+\omega) d\nu(\omega) = \int h(\omega) d\nu(\omega).$$

But $E \cap C(\mathcal{R}_+^\infty)$ is dense in the space of continuous functions $C(\mathcal{R}_+^\infty)$; the invariance of ν follows.

Let us now check that the invariant measure ν is ergodic. Let $A \subset \mathcal{R}_+^\infty$ be a T^+ -invariant Borel subset with $\nu(A) > 0$ and A^c its complement. Let $\varepsilon > 0$. We will prove that there exists $a \in A$ such that

$$(10.173) \quad \nu(A \cap \mathcal{R}_+^\infty(a)) \geq (1 - \varepsilon) \nu(\mathcal{R}_+^\infty(a)).$$

As $\varepsilon > 0$ is arbitrary, this easily implies $\nu(A) = 1$.

As $\nu(A) > 0$, we can find (P, Q, n) such that

$$(10.174) \quad \nu(\{\omega \subset P\} \cap A^c) \leq \varepsilon' \nu(\{\omega \subset P\}),$$

where $\varepsilon'\varepsilon^{-1}$ is small. Let r be the number of factors in the prime decomposition of (P, Q, n) . Up to a set of measure 0, we have

$$(10.175) \quad \{\omega \subset P\} = \bigcup_{0 \leq j \leq r} \bigcup_{P_j} (T^+)^{-j}(\mathcal{R}_+^\infty(P_j)) \pmod{0}$$

where P_j runs through prime elements satisfying $P_j \subset (T^+)^j(P)$ and $(T^+)^{-j}$ is the inverse branch of $(T^+)^j$ whose image contains P . From (10.174), there exists $0 \leq j \leq r$ and P_j such that

$$(10.176) \quad \nu(A^c \cap (T^+)^{-j}(\mathcal{R}_+^\infty(P_j))) \leq \varepsilon' \nu((T^+)^{-j}(\mathcal{R}_+^\infty(P_j))).$$

We apply the Jacobian property, taking (10.120) into account to get (10.173) with $\varepsilon = C\varepsilon'$. We have proved that ν is ergodic. We summarize:

Proposition 47. *The measure $d\nu = h^* d\mu$ is T^+ -invariant, ergodic. It satisfies, for all $(P, Q, n) \in \mathcal{R}$:*

$$C^{-1}|P|^{d_s} \leq \nu(\{\omega \subset P\}) \leq C|P|^{d_s}$$

and $\nu(\mathcal{D}_+^\infty) = 1$.

We will now lift ν to obtain a \tilde{T}^+ -invariant probability measure on $\tilde{\mathcal{R}}_+^\infty$.

Proposition 48. *There exists a unique probability measure $\tilde{\nu}$ on $\tilde{\mathcal{R}}_+^\infty$ which is \tilde{T}^+ -invariant and projects onto ν under π . It is ergodic.*

Proof. The arguments are standard.

Existence. Denote by $\mathcal{M}(\nu)$ the set of probability measures on \mathcal{R}_+^∞ which project onto ν . This is a compact set for the weak topology, invariant under \tilde{T}^+ because ν is T^+ -invariant. One obtains a \tilde{T}^+ -invariant measure in $\mathcal{M}(\nu)$ by taking any $\tilde{\nu}_0 \in \mathcal{M}(\nu)$ and choosing a weak limit of a subsequence of

$$(10.177) \quad \frac{1}{n} \sum_0^{n-1} [(\tilde{T}^+)^j]^*(\tilde{\nu}_0).$$

Uniqueness. The set of fixed points for the action of \tilde{T}^+ on $\mathcal{M}(\nu)$ is thus non-empty. It is also compact and convex. If it has more than one point, it has at least two distinct extremal points $\tilde{\nu}_0, \tilde{\nu}_1$. As ν is ergodic, $\tilde{\nu}_0$ and $\tilde{\nu}_1$ are also ergodic. Still by the ergodicity of ν , some stable curve ω must meet the basins of both $\tilde{\nu}_0$ and $\tilde{\nu}_1$. But stable curves are contracted exponentially fast under positive iteration by T^+ ; we should thus have $\tilde{\nu}_0 = \tilde{\nu}_1$, a contradiction.

We have already said that $\tilde{\nu}$ is ergodic. □

Finally, we want to "spread" the \tilde{T}^+ -invariant measure $\tilde{\nu}$ in order to obtain a g -invariant measure σ . Let $\Lambda = \Lambda_g$ as in the Introduction (cf. Subsection 1.2).

We first observe that the support of $\tilde{\nu}$ is contained into $\Lambda \cap \tilde{\mathcal{R}}_+^\infty$: if $N \subset \tilde{\mathcal{R}}_+^\infty$ is compact and disjoint from Λ , then N is disjoint from the image of $(\tilde{T}^+)^j$ if j is large enough, hence $\tilde{\nu}(N) = 0$.

Let now h be a continuous, and thus bounded, function on Λ . For $x \in \Lambda \cap \mathcal{D}_+^\infty$, we write

$$(10.178) \quad \tilde{T}^+(x) = g^{N(x)}(x),$$

where $N(x) = n$ if $x \in \tilde{\mathcal{R}}_+^\infty(P)$ with $(P, Q, n) \in \mathcal{P}$. We define:

$$(10.179) \quad Sh(x) = \sum_{0 \leq j < N(x)} h(g^j(x)).$$

The function Sh is defined $\tilde{\nu}$ -almost everywhere. It satisfies:

$$(10.180) \quad |Sh(x)| \leq \|h\|_\infty N(x).$$

By Proposition 47 and Corollary 14 in Subsection 10.3, the function N is $\tilde{\nu}$ -integrable. We have therefore defined an operator

$$(10.181) \quad S: C(\Lambda) \mapsto L^1(\tilde{\nu}).$$

where $C(\Lambda)$ stands for the space of continuous functions on $\Lambda = \Lambda_g$.

We define a finite measure σ on Λ by

$$(10.182) \quad \int h d\sigma = \int Sh d\tilde{\nu},$$

for $h \in C(\Lambda)$. From the definition of Sh , we have

$$(10.183) \quad S(h \circ g) = Sh + h \circ \tilde{T}^+ - h.$$

Thus, the \tilde{T}^+ -invariance of $\tilde{\nu}$ implies that σ is g -invariant. It is ergodic. The Lyapunov exponents of \tilde{T}^+ for $\tilde{\nu}$ are non-zero because \tilde{T}^+ is uniformly hyperbolic. To get the Lyapunov exponents of g for σ we have only to change time, which is possible since N is $\tilde{\nu}$ -integrable.

In the next and last section, we will see that in some appropriate geometric sense, the measure σ captures "most" of the dynamics on Λ , and therefore can be considered as a naturally defined geometric invariant measure on Λ .

We end this section by observing that everything that has been done for T^+ and positive iteration in Section 10, can also be done for T^- and negative iteration, leading to another naturally defined geometric invariant measure σ^- on Λ .

11 Some Further Geometric Properties of the Invariant Set

In this final section we pursue the geometric study of the invariant set $\Lambda = \Lambda_g$ in two directions. First, we will describe in a rather precise way, both from a dynamical and a geometrical point of view, the intersection of an unstable curve in \mathcal{R}^∞ , as defined in Subsection 10.4, with the invariant set Λ . In the second part of the section, we prove that Λ is a saddle-like invariant set in the measure-theoretical sense: both its stable and unstable sets have Lebesgue measure 0; thus, no attractors are present in Λ .

11.1 One-Dimensional Analysis of the Invariant Set

Let $\omega^* \in \mathcal{R}^\infty$ be an unstable curve as defined in Subsection 10.4. Let $(P_k^*, Q_k^*, n_k^*)_{k \geq 0}$ be the canonical sequence associated to ω^* (cf. definition also in Subsection 10.4). We have

$$(11.1) \quad \omega^* = \bigcap_{k \geq 0} Q_k^*,$$

where Q_0^* is a rectangle R_a and Q_{k+1}^* is a child of Q_k^* for each $k \geq 0$. We want to analyze the intersection $\omega^* \cap \Lambda$. In Section 10, we have analyzed the set $\tilde{\mathcal{R}}_+^\infty$ and we know, in particular, that $\omega^* \cap \Lambda$ contains the subset $\omega^* \cap \tilde{\mathcal{R}}_+^\infty$; this last subset has Hausdorff dimension d_s characterized in terms of the transfer operator studied in Section 10; in particular, this dimension is independent of ω^* .

Let us summarize the results of our analysis in this section.

Theorem 3. *The intersection $\omega^* \cap \Lambda$ is the disjoint union of*

- *a, at most countable, family of Cantor sets $\Lambda_i(\omega^*)$,*
- *a, at most countable, set $Cr(\omega^*)$,*
- *an exceptional set $\mathcal{E}(\omega^*)$,*

with the following properties

(i) *For each i , there exists a piece $\omega^*(i)$ of ω^* containing $\Lambda_i(\omega^*)$, an unstable curve ω_i^* and an integer n_i such that*

$$(11.2) \quad g^{n_i}(\omega^*(i)) = \omega_i^*,$$

$$(11.3) \quad g^{n_i}(\Lambda_i(\omega^*)) = \omega_i^* \cap \tilde{\mathcal{R}}_+^\infty.$$

In particular, there is a special index $i = 0$ for which $n_0 = 0$, $\omega^(0) = \omega_0^* = \omega^*$, $\Lambda_0(\omega^*) = \omega^* \cap \tilde{\mathcal{R}}_+^\infty$.*

(ii) For every point $c \in Cr(\omega^*)$, there exists a stable curve $\omega_+(c) \in \mathcal{R}_+^\infty$, an unstable curve $\omega_-(c) \in \mathcal{R}_-^\infty$, a positive integer $n(c)$ such that $g^{n(c)}(c)$ is a quadratic tangency point between $\omega^+(c)$ and $g^{N_0}(\omega^-(c) \cap L_u)$.

(iii) The Hausdorff dimension of $\mathcal{E}(\omega^*)$ is not greater than

$$(11.4) \quad (d_s^0 + d_u^0 - 1) \frac{2d_s^0}{2d_u^0 + d_s^0} + o(1)$$

where the $o(1)$ term is small provided τ is small enough. Consequently, the Hausdorff dimension of $\omega^* \cap \Lambda$ is equal to d_s .

(iv) Every point $x \in \mathcal{E}(\omega^*)$ is the intersection of a decreasing sequence of pieces $(\omega^*(i_n(x)))_{n \geq 0}$.

Remark.

1. The structure will be made more precise in the next subsections. We have tried here to extract the most significant features of our analysis.

2. Even with $d_s^0 + d_u^0 > 1$, it may happen that Λ is a uniformly hyperbolic horseshoe; then, the family $(\Lambda_i(\omega^*))_i$ is finite, $Cr(\omega^*)$ and $\mathcal{E}(\omega^*)$ are empty. When Λ is not uniformly hyperbolic, the family $(\Lambda_i(\omega^*))_i$ is countable and $\mathcal{E}(\omega^*)$ is a Cantor set; it is not clear in this case if $Cr(\omega^*)$ can be empty.

11.2 Parabolic Cores

Let $(P, Q, n) \in \mathcal{R}$, \mathcal{R} as in Subsection 10.1.

Definition. The *parabolic core* of P , denoted by $c(P)$, is the set of points of $W^s(\Lambda, \widehat{R})$ which belong to P but not to any child of P . The parabolic core of Q , denoted by $c(Q)$, is the set of points of $W^u(\Lambda, \widehat{R})$ which belong to Q but not to any child of Q .

We have partitions

$$(11.5) \quad R \cap W^s(\Lambda, \widehat{R}) = \bigsqcup_{\mathcal{R}} c(P) \sqcup \widetilde{\mathcal{R}}_+^\infty,$$

$$(11.6) \quad R \cap W^u(\Lambda, \widehat{R}) = \bigsqcup_{\mathcal{R}} c(Q) \sqcup \widetilde{\mathcal{R}}_-^\infty.$$

If R_a is the rectangle which contains ω^* , we also have

$$(11.7) \quad \omega^* \cap \Lambda = \bigsqcup_{P \subset R_a} (\omega^* \cap c(P)) \sqcup (\omega^* \cap \widetilde{\mathcal{R}}_+^\infty).$$

The parabolic core is empty if and only if P is I -decomposable for a small enough parameter interval containing the given strongly regular parameter value. In particular, $c(P)$ is empty if Q is I -transverse. Thus, the union in (11.5), (11.7) can be restricted to those $(P, Q, n) \in \mathcal{R}$ such that Q is I -critical for all I .

We will denote by $C(\omega^*)$ the set of elements $(P, Q, n) \in \mathcal{R}$ such that $c(P) \cap \omega^*$ is not empty. For any $(P, Q, n) \in C(\omega^*)$, Q is I -critical for all I .

11.3 Decomposition of $c(P) \cap \omega^*$

Let $(P, Q, n) \in C(\omega^*)$. For $k \geq 0$, set

$$(11.8) \quad (P_k, Q_k, n_k) = (P_k^*, Q_k^*, n_k^*) * (P, Q, n),$$

$$(11.9) \quad \omega_P^* = \bigcap_{k \geq 0} Q_k.$$

The unstable curve ω_P^* is contained in Q and we have

$$(11.10) \quad g^n(\omega^* \cap c(P)) \subset \omega_P^* \cap L_u.$$

We define a tree $\mathcal{A}(\omega^*, P)$ as follows. The vertices are the rectangles $P' \subset P_s$ with the following property: for any parameter interval I (containing the given parameter value, say t), for any $Q_k \supset \omega_P^*$, Q_k and P' are not I -separated, and Q_k and the parent of P' are I -critically related.

We connect two vertices by an (oriented) edge if one is the parent of the other. We say that a vertex P' is *critical* if, for all I and $Q_k \supset \omega_P^*$, Q_k and P' are I -critically related. Otherwise, we say that P' is *transverse*. The parent of a vertex is always a critical vertex, except if this vertex is P_s , the *root* of the tree. When P' is a transverse vertex, the smallest integer k such that Q_k, P are I -transverse for I small enough is called the *level* of P' .

Let P' be a critical vertex; then, for every parameter interval $I \ni t$, P' is I -critical and, therefore, decomposable.

Let P' be a transverse vertex of level 0. We have $Q_0 = Q$. Therefore, the parabolic composition $(P, Q, n) \square (P', Q', n')$ is well defined and produces two children of P .

Let P' be a transverse vertex of level $k > 0$. For all $m \geq k$, the parabolic composition $(P_m, Q_m, n_m) \square (P', Q', n')$ is well-defined and produces two elements $(P_m^\pm, Q_m^\pm, n_m^\pm)$. The formulas

$$(11.11) \quad \begin{aligned} \omega_{P, P', +}^* &:= \cap Q_m^+, \\ \omega_{P, P', -}^* &:= \cap Q_m^-, \end{aligned}$$

define unstable curves $\omega_{P,P',\pm}^*$ contained in Q' . We also define pieces $\omega^*(P, P', \pm)$ through

$$(11.12) \quad g^{n_{P,P'}}(\omega^*(P, P', \pm)) = \omega_{P,P',\pm}^*,$$

$$(11.13) \quad n_{P,P'} := n + n' + N_0.$$

Lemma 15. *Let x be a point in $\omega^* \cap c(P)$, $y = g^{n+N_0}(x)$. Either y belong to a transverse vertex of level > 0 or it belongs to an infinite decreasing sequence of critical vertices.*

Proof. We have $g^n(x) \in L_u$ (cf. (11.10)), $y \in L_s \subset P_s$, and P_s is the root and a critical vertex of the tree $\mathcal{A}(\omega^*, P)$. We assume that the first possibility in the statement of the lemma does not hold and construct, starting with P_s , a sequence of critical vertices containing y .

Assume that y belongs to a critical vertex P' . As P' is indecomposable and $y \in W^s(\Lambda)$, y belongs to some child P'_1 of P' . This rectangle is a vertex of the tree: otherwise, Q_k and P'_1 would be I -separated if I and Q_k are thin enough, and then $g^{N_0}(g^n(\omega^*) \cap L_u) \cap P'_1$ (which contains y) would be empty. The vertex P'_1 cannot be transverse of level 0 because, as remarked above, the parabolic composition of (P, Q, n) and (P'_1, Q'_1, n'_1) would produce a child of P containing x , contradicting the hypothesis that $x \in c(P)$. Finally P'_1 cannot be transverse of level > 0 by hypothesis. It must be a critical vertex, and the induction step is complete. \square

Proposition 49. *There is at most one point $x \in \omega^* \cap c(P)$ such that $y = g^{n+N_0}(x)$ belongs to a decreasing sequence of critical vertices. When such a point exists, the intersection of this decreasing sequence of vertices is a stable curve which intersects $g^{N_0}(L_u \cap \omega_P^*)$ at y as a quadratic tangency point.*

Proof. Let x be a point in $\omega^* \cap c(P)$ such that $y = g^{n+N_0}(x)$ belongs to a decreasing sequence $(P'_\ell)_{\ell \geq 0}$ of critical vertices. Denote by ω_+ the stable curve which is the intersection of these critical vertices. For all parameter intervals I , all $k \geq 0$, $\ell \geq 0$, Q_k and P'_ℓ are I -critically related. This implies that

$$(11.14) \quad \lim_{\substack{k \rightarrow +\infty \\ \ell \rightarrow +\infty}} \delta(Q_k, P'_\ell) = 0.$$

For large k and ℓ , let γ_k (resp. (γ'_ℓ)) be the image in Q_k (resp. the inverse image in P'_ℓ) of the intersection of P_k with an horizontal curve (resp. the intersection of Q'_ℓ with a vertical curve). By (11.14), the distance between the vertical-like curve γ'_ℓ and the tip of the parabolic-like curve $g^{N_0}(\gamma_k)$ goes to zero as k, ℓ go to $+\infty$. Passing to the limit, we see that ω_+ has a tangency with $g^{N_0}(\omega_P^* \cap L_u)$. This tangency is quadratic in the following sense (cf. also the remark after the end of the proof): First, $g^{N_0}(\omega_P^* \cap L_u)$ is contained, with the exception of the tangency point, in one of the components of $P_s - \omega_+$; moreover, the angle between the tangent lines to $\omega_+(x)$, $g^{N_0}(L_u \cap \omega_P^*)$ at points on these curves at the same distance and on the same side of the tangency point is of the

same order as this distance to the tangency point. This is a consequence of the uniform estimates (3.21), (3.22) in Subsection 3.5.

As ω_+ and $g^{N_0}(L_u \cap \omega_P^*)$ meet at only one point, this point must be y . If x' is a point with the same property as x , and we construct ω'_+ in the same way as ω_+ , we must have $\omega_+ = \omega'_+$ because otherwise $g^{N_0}(L_u \cap \omega_P^*) \cap \omega_+$ or $g^{N_0}(L_u \cap \omega_P^*) \cap \omega'_+$ is empty. But, then, we have $y' := g^{n+N_0}(x') = y$ and $x' = x$. \square

Remark. *Calculations involving partial derivatives of higher order for the maps (A, B) , which implicitly represent elements of \mathcal{R} , show that stable curves and unstable curves are actually of class C^∞ , with uniform estimates in the C^k topology for all k . Then, quadratic tangency can be taken in the usual sense. However, the calculations involved, especially when considering parabolic composition, are quite long and not very interesting; we decided to stick to the C^{1+Lip} regularity class, where the notion of "quadratic" tangency, as explained in the proof of Proposition 49, still makes sense.*

It is easy to see exactly when a point $x \in \omega^* \cap c(P)$ with the property specified in Proposition 49 does exist: a necessary and sufficient condition is that the tree $\mathcal{A}(\omega^*, P)$ is infinite. In this case, the point x will be a point of the set $Cr(\omega^*)$ in the statement of Theorem 3 and the point $y = g^{n+N_0}(x)$ is said to be *critical*.

Summarizing what we have established so far, two cases may happen:

1) The tree $\mathcal{A}(\omega^*, P)$ is finite. Then, the intersection $\omega^* \cap c(P)$ is the finite disjoint union of the sets

$$(11.15) \quad \omega^*(P, P', \pm) \cap \Lambda$$

where P' runs through the vertices of the tree which are transverse of level > 0 . The image under $g^{n_{P, P'}}$ of the set (11.15) is the intersection $\omega_{P, P', \pm}^* \cap \Lambda$.

2) The tree $\mathcal{A}(\omega^*, P)$ is infinite. Then, the intersection $\omega^* \cap c(P)$ is the countable disjoint union of the sets $\omega^*(P, P', \pm) \cap \Lambda$ as above and a single point $x \in Cr(\omega^*)$. The point $x = x_P$ is the limit of the pieces $\omega^*(P, P', \pm)$ (whose diameters goes to 0 as $|P'|$ goes to 0).

11.4 The Structure of $\omega^* \cap \Lambda$

We are now ready to prove all the statements in Theorem 3, stated above in Subsection 11.1, with the exception of (iii) (the estimate on the Hausdorff dimension of $\mathcal{E}(\omega^*)$).

The structure of $\omega^* \cap \Lambda$ that we are looking for, which is roughly described in Theorem 3, is obtained by iterating the partition (11.7) and the decomposition of $\omega^* \cap c(P)$ described in Subsection 11.3.

At the first step, we have partitioned $\omega^* \cap \Lambda$ into the following subsets:

- the intersection $\omega^* \cap \widetilde{\mathcal{R}}_+^\infty$; points in this set are said of type I;
- for each $(P, Q, n) \in C(\omega^*)$ such that $\mathcal{A}(P, \omega^*)$ is infinite, a point x_P such that $y_P = g^{n+N_0}(x_P)$ is critical; such points x_P are said of type II;
- for each $(P, Q, n) \in C(\omega^*)$, each vertex (P', Q', n') of $\mathcal{A}(\omega^*, P)$ which is transverse of level bigger than 0, each $\varepsilon \in \{+, -\}$, the intersection $\omega^*(P, P', \varepsilon) \cap \Lambda$; the image of this set under $g^{n_{P, P'}}$ is the intersection $\omega_{P, P', \varepsilon}^* \cap \Lambda$ of another unstable curve with Λ .

The intersection $\omega_{P, P', \varepsilon}^* \cap \Lambda$ will be analyzed in the same way that $\omega^* \cap \Lambda$.

Consider a point $z_0 \in \omega^* \cap \Lambda$. If it is of type I, it belongs to the set $\Lambda_0(\omega^*) := \omega^* \cap \widetilde{\mathcal{R}}_+^\infty$ of the statement of Theorem 3. If it is of type II, it belongs to $Cr(\omega^*)$. Assume now that it is of type III. Then, it belongs to some $\omega^*(P, P', \varepsilon) \cap \Lambda$ as above. Define

$$(11.16) \quad z_1 = g^{n_{P, P'}}(z_0),$$

which belongs to $\omega_{P, P', \varepsilon}^* \cap \Lambda =: \omega_1^*$. This point may in turn be of type I, II, III with respect to ω_1^* . The process stops if z_1 is of type I or II; if z_1 is of type III, it belongs to some piece $\omega_1^*(P_1, P'_1, \varepsilon_1)$; we define

$$(11.17) \quad z_2 = g^{n_{P_1, P'_1}}(z_1),$$

which belongs to $\omega_2^* \cap \Lambda$, with

$$(11.18) \quad \omega_2^* := g^{n_{P_1, P'_1}}(\omega_1^*(P_1, P'_1, \varepsilon_1)).$$

Iterating this process lead to one of three possible outcomes:

- 1) the z_k 's are defined and of type III for all $k \geq 0$; the corresponding initial points z_0 form the set $\mathcal{E}(\omega^*)$.
- 2) the z_k 's are defined for $0 \leq k \leq \ell$ and z_ℓ is of type I, i.e. it belongs to $\widetilde{\mathcal{R}}_+^\infty$; let $(P_k, P'_k, \varepsilon_k)$ for $0 \leq k < \ell$ be the data involved in the definitions of the z_k 's. We collect together the initial points z_0 's with the same set of data; such a set form one of the Cantor sets $\Lambda_i(\omega^*)$ in Theorem 3.
- 3) the z_k 's are defined for $0 \leq k \leq \ell$ and z_ℓ is of type II. Then z_0 belongs to the set $Cr(\omega^*)$.

We have now completely defined the partition of $\omega^* \cap \Lambda$ described in Theorem 3. The properties (i), (ii), (iv) follow immediately from the definitions.

11.5 Hausdorff Dimension of the Exceptional Set $\mathcal{E}(\omega^*)$

The self-similar structure apparent in the definition of $\mathcal{E}(\omega^*)$ is the key to obtain an estimate of the dimension of this set. More specifically, we have

$$(11.19) \quad \mathcal{E}(\omega^*) = \bigsqcup_{(P, P', \varepsilon)} g^{-n_{P, P'}} (\mathcal{E}(\omega^*_{P, P', \varepsilon})),$$

where $\varepsilon \in \{+, -\}$, P runs through $C(\omega^*)$ and P' through vertices of $\mathcal{A}(\omega^*, P)$ which are transverse of level > 0 .

Lemma 16. *The maps*

$$g^{n_{P, P'}} : \omega^*(P, P', \varepsilon) \rightarrow \omega^*_{P, P', \varepsilon}$$

have uniformly bounded distortion.

Proof. Let k be an integer larger than the level of the transverse vertex P' . Then, the parabolic composition of (P_k, Q_k, n_k) (cf. (11.8)) and (P', Q', n') is defined and produces an element (P'_k, Q'_k, n'_k) such that Q'_k contains $\omega^*_{P, P', \varepsilon}$. Let γ_k^* be an horizontal segment in P_k^* , γ_k its image under $g^{n_k^*}$, γ'_k the image of $\gamma_k^* \cap P'_k$ under $g^{n'_k}$.

The affine-like maps

$$(11.20) \quad g^{n_k^*} : P_k^* \rightarrow Q_k^*, \quad g^{n'_k} : P'_k \rightarrow Q'_k,$$

have bounded distortion, hence the one-dimensional map

$$(11.21) \quad g^{n_k^*} \circ (g^{n'_k})^{-1} : \gamma'_k \rightarrow \gamma_k$$

have also uniformly bounded distortion. Letting k go to $+\infty$, γ'_k converge to $\omega^*_{P, P', \varepsilon}$ and γ_k to ω^* in the $C^{2-\varepsilon}$ -topology for all $\varepsilon > 0$. The statement of the lemma follows. \square

Lemma 17. *Let*

$$\delta(\omega^*_P, P') = \lim_{k \rightarrow +\infty} \delta(Q_k, P')$$

We have

$$C^{-1} \leq \frac{\text{diam} \omega^*(P, P', \varepsilon)}{|P| |P'| (\delta(\omega^*_P, P'))^{-\frac{1}{2}}} \leq C$$

Proof. As in the proof of Lemma 16, we write

$$(11.22) \quad g^{n_{P, P'}} = g^{n'_k} \circ (g^{n_k^*})^{-1}.$$

From the estimate (3.27) for parabolic composition in Subsection 3.5, we have

$$(11.23) \quad C^{-1} \leq \frac{\text{diam} \gamma_k^*}{|P_k|, |P'| (\delta(Q_k, P'))^{-\frac{1}{2}}} \leq C$$

We also have, from the estimates on simple composition

$$(11.24) \quad C^{-1} \leq \frac{\text{diam } \omega^*(P, P', \varepsilon) |P_k^*|}{\text{diam } \gamma_k^*} \leq C,$$

$$(11.25) \quad C^{-1} \leq \frac{|P_k|}{|P| |P_k^*|} \leq C.$$

Multiplying these three inequalities yields the Lemma. \square

Let us introduce

$$(11.26) \quad \chi(d) = \sum_{(P, P', \varepsilon)} [\text{diam } \omega^*(P, P', \varepsilon)]^d.$$

If we are able, for some value of d , to show that the series defining χ is convergent and $\chi(d)$ is small, then by (11.19) and Lemma 16, we will deduce that the Hausdorff dimension of $\mathcal{E}(\omega^*)$ is $\leq d$.

In order to study χ , we will first fix P in $C(\omega^*)$ and sum over (P', ε) . As ε takes only two values, and in view of Lemma 17, we define, for $P \in C(\omega^*)$:

$$(11.27) \quad \chi_P(d) = \sum_{P'} |P'|^d \delta(\omega_P^*, P')^{-\frac{1}{2}d}.$$

We will then have

$$(11.28) \quad \chi(d) \leq C \sum_P |P|^d \chi_P(d).$$

In the sum (11.27), P' is a transverse vertex of level > 0 , and we therefore must have

$$(11.29) \quad \delta(\omega_P^*, P') \leq \delta_{\max} := \min(\varepsilon_0, C|Q|^{1-\eta}).$$

In the series (11.27), we first sum over those P' such that

$$(11.30) \quad 2^{-\ell} \delta_{\max} \geq \delta(\omega_P^*, P') \geq 2^{-\ell-1} \delta_{\max}$$

for some fixed $\ell \geq 0$. This allow us to write

$$(11.31) \quad \chi_P(d) \leq C \delta_{\max}^{-\frac{1}{2}d} \sum_{\ell \geq 0} 2^{\frac{\ell d}{2}} \left(\sum^{(\ell)} |P'|^d \right),$$

where $\sum^{(\ell)}$ means that P' is constrained by (11.30). We divide $\sum^{(\ell)}$ into two parts.

In the first part, denoted by $\sum_1^{(\ell)}$, we consider only those P' such that its parent \tilde{P}' satisfies

$$(11.32) \quad |\tilde{P}'| \leq 2^{-\ell} \delta_{\max}.$$

To estimate $\sum_1^{(\ell)} |P'|^d$, first observe that, with d bounded away from 0, it follows from Proposition 21 in Subsection 8.1 that the sum of $|P'|^d$ over children of a fixed parent \tilde{P}' is bounded by $C|\tilde{P}'|^d$. We must therefore bound $\sum_1^{(\ell)} |\tilde{P}'|^d$.

Also, as \tilde{P}' is a critical vertex, \tilde{P}' cannot be very thin: from Proposition 10 in Subsection 6.6, we have

$$(11.33) \quad |\tilde{P}'| \geq C^{-1}[\delta(\omega_P^*, P')]^{(1-\eta)^{-1}} \geq C^{-1}(\delta_{\max} 2^{-\ell})^{(1-\eta)^{-1}}.$$

Finally, the number of \tilde{P}' with $|\tilde{P}'|$ of order $2^{-m-\ell} \delta_{\max}$ is at most $C2^m$ and the integer m here is restricted by (11.33) to the range

$$(11.34) \quad 1 \leq 2^m \leq C(\delta_{\max} 2^{-\ell})^{-\eta(1-\eta)^{-1}}.$$

We, therefore, obtain for d bounded away from 0 and 1,

$$(11.35) \quad \begin{aligned} \sum_1^{(\ell)} |P'|^d &\leq C \sum_1^{(\ell)} |\tilde{P}'|^d \\ &\leq C \delta_{\max}^d 2^{-\ell d} \sum_m 2^{m(1-d)} \\ &\leq C (\delta_{\max} 2^{-\ell})^{d-\eta}. \end{aligned}$$

In the second part of $\sum^{(\ell)}$, denoted by $\sum_2^{(\ell)}$, we have on the opposite

$$(11.36) \quad |\tilde{P}'| > 2^{-\ell} \delta_{\max}.$$

As $|\tilde{P}'| > \delta(\omega_P^*, P')$, the number of possibilities for \tilde{P}' is now bounded. As each P' is a transverse vertex, we must have (by (R7))

$$(11.37) \quad |P'| \leq C(2^{-\ell} \delta_{\max})^{(1-\eta)^{-1}}.$$

In particular, from (11.36), (11.37), P' is a non-simple child of \tilde{P}' . From Proposition 21 in Subsection 8.1, the number of P' with $|P'|$ of order $2^{-m} \varepsilon_0$ is at most $2^{Cm\eta}$.

We have

$$(11.38) \quad \sum_2^{(\ell)} |P'|^d \leq \varepsilon_0^d \sum_m 2^{-m(d-C\eta)} \leq C \varepsilon_0^{-C\eta} (2^{-\ell} \delta_{\max})^{d-C\eta}.$$

Putting (11.35) and (11.38) together yields

$$(11.39) \quad \sum_{C(\omega^*)}^{\ell} |P'|^d \leq C(\delta_{\max} 2^{-\ell})^{d-C'\eta}$$

and introducing this in (11.31) allow us to estimate χ_P :

$$(11.40) \quad \chi_P(d) \leq C \delta_{\max}^{\frac{1}{2}d-C'\eta}.$$

Finally, we obtain

$$(11.41) \quad \chi(d) \leq C \sum_{C(\omega^*)} |P|^d [\min(\varepsilon_0, |Q|)]^{\frac{1}{2}d-C\eta}.$$

We do not know exactly the set $C(\omega^*)$, but we know that if $(P, Q, n) \in C(\omega^*)$, the parabolic core $c(P)$ is non-empty and Q must be I -critical for all parameter intervals I containing the given parameter value.

We use Hölder's inequality to separate the P and Q in (11.41): for any $p, q > 1$ such that

$$(11.42) \quad \frac{1}{p} + \frac{1}{q} = 1,$$

we have

$$(11.43) \quad \chi(d) \leq C \chi_+(d)^{\frac{1}{p}} \chi_-(d)^{\frac{1}{q}}$$

where

$$(11.44) \quad \chi_+(d) = \sum_{Q \text{ critical}} |P|^{dp},$$

$$(11.45) \quad \chi_-(d) = \sum_{Q \text{ critical}} \min(\varepsilon_0, |Q|)^{\left(\frac{1}{2}d - C\eta\right)q}.$$

We will choose d, p, q (satisfying (11.42) in order to have $\chi_+(d)$ bounded and $\chi_-(d)$ small (when ε_0 is small). For such a choice, we can conclude that the Hausdorff dimension of $\mathcal{E}(\omega^*)$ is $\leq d$.

We now use that the parameter value is strongly regular, more precisely that the eight estimates (SR1), (SR2) of Subsection 9.2 on the size of the critical locus are satisfied.

It is not difficult to deduce from (SR1) $_{\hat{u}}$ that, if

$$(11.46) \quad \frac{1}{2}dq > d_s^0 + d_u^0 - 1$$

then, $\chi_-(d)$ will be small.

From (SR2) $_{\hat{u}}$, one can also deduce that if

$$(11.47) \quad dp \frac{d_u^0}{d_s^0} > d_s^0 + d_u^0 - 1$$

then χ_+ will be bounded. The relations (11.42), (11.46), (11.47) are compatible exactly when

$$(11.48) \quad d > (d_s^0 + d_u^0 - 1) \frac{2d_s^0}{2d_u^0 + d_s^0}.$$

We observe that the right hand size is always $< d_s^0$. This ends the proof of Theorem 3.

Remark. *The inequalities (11.46), (11.47) should be understood in the following sense: the difference between the left and right-hand sides is much larger than τ (which is itself much larger than η).*

11.6 The Stable and Unstable Sets of Λ

Our goal at the end of this final section is to prove that the invariant set Λ is a saddle-like object in the following measure-theoretical sense:

Theorem 4. *For a strongly regular parameter, both the stable set $W^s(\Lambda)$ and the unstable set $W^u(\Lambda)$ have Lebesgue measure 0.*

The situation is symmetrical and we will deal with the stable set.

We have:

$$(11.49) \quad W^s(\Lambda) = \bigsqcup_{n \geq 0} g^{-n}(W^s(\Lambda, \widehat{R}) \cap R).$$

Therefore, it is sufficient to show that $W^s(\Lambda, \widehat{R}) \cap R$ has Lebesgue measure 0. We write

$$(11.50) \quad R \cap W^s(\Lambda, \widehat{R}) = \bigsqcup_{n \geq 0} \left(W^s(\Lambda, \widehat{R}) \cap R \cap g^{-n}(\widetilde{\mathcal{R}}_+^\infty) \right) \sqcup \mathcal{E}^+,$$

with

$$(11.51) \quad \mathcal{E}^+ = \{z \in W^s(\Lambda, \widehat{R}) \cap R, g^n(z) \notin \widetilde{\mathcal{R}}_+^\infty \text{ for all } n \geq 0\}.$$

We have seen in Section 10 that $\widetilde{\mathcal{R}}_+^\infty$ is Lipschitzian with transverse Hausdorff dimension d_s . Therefore, the Hausdorff dimension of $\widetilde{\mathcal{R}}_+^\infty$ is $1 + d_s$ and its Lebesgue measure is 0. The same is true of $g^{-n}(\widetilde{\mathcal{R}}_+^\infty)$. We have to prove that the Lebesgue measure of \mathcal{E}^+ is equal to 0.

11.7 Decomposition of \mathcal{E}^+

By the definition of \mathcal{E}^+ and of the parabolic cores, we can write

$$(11.52) \quad \mathcal{E}^+ = \bigsqcup_{P_0} \mathcal{E}^+(P_0),$$

where

$$(11.53) \quad \mathcal{E}^+(P_0) = \mathcal{E}^+ \cap c(P_0)$$

and (P_0, Q_0, n_0) runs through the set \mathcal{C}_- of elements of R with $c(P_0) \neq \emptyset$. In particular, Q_0 is I -critical for all I containing the given parameter value.

For any such P_0 , we have

$$(11.54) \quad g^{n_0}(\mathcal{E}^+(P_0)) \subset Q_0 \cap L_u \cap \mathcal{E}^+$$

$$(11.55) \quad g^{n_0+N_0}(\mathcal{E}^+(P_0)) \subset L_s \cap \mathcal{E}^+.$$

For $P_1 \in \mathcal{C}_-$, define

$$(11.56) \quad \mathcal{E}^+(P_0, P_1) = \{z \in \mathcal{E}^+(P_0), g^{n_0+N_0}(z) \in c(P_1)\}.$$

We have a partition

$$(11.57) \quad \mathcal{E}^+(P_0) = \bigsqcup_{P_1} \mathcal{E}^+(P_0, P_1).$$

At step k , we have a partition

$$(11.58) \quad \mathcal{E}^+ = \bigsqcup_{P_0, \dots, P_k} \mathcal{E}^+(P_0, \dots, P_k)$$

where the (P_i, Q_i, n_i) runs through \mathcal{C}_- . We write

$$(11.59) \quad \begin{aligned} m_0 &= n_0, \\ m_1 &= n_0 + N_0 + n_1 \\ &\vdots \\ m_j &= n_0 + N_0 + n_1 + N_0 + \dots + n_{j-1} + N_0 + n_j \\ &= m_{j-1} + N_0 + n_j. \end{aligned}$$

For $0 \leq j \leq k$, we have

$$(11.60) \quad g^{m_j}(\mathcal{E}^+(P_0, \dots, P_k)) \subset Q_j \cap L_u \cap \mathcal{E}^+,$$

$$(11.61) \quad g^{m_j+N_0}(\mathcal{E}^+(P_0, \dots, P_k)) \subset L_s \cap \mathcal{E}^+.$$

We define, for $P_{k+1} \in \mathcal{C}_-$

$$(11.62) \quad \mathcal{E}^+(P_0, \dots, P_k, P_{k+1}) = \{z \in \mathcal{E}^+(P_0, \dots, P_k), g^{m_k+N_0}(z) \in c(P_{k+1})\}$$

and we have

$$(11.63) \quad \mathcal{E}^+(P_0, \dots, P_k) = \bigsqcup_{P_{k+1}} \mathcal{E}^+(P_0, \dots, P_k, P_{k+1}).$$

However, in order to have $\mathcal{E}^+(P_0, \dots, P_k) \neq \emptyset$ strong restrictions on the P_i must take place. We have already mentioned that $(P_i, Q_i, n_i) \in \mathcal{C}_-$. This is the only restriction on (P_0, Q_0, n_0) . But, from (11.55), P_1 must meet P_s and we also know that Q_1 is critical. As the parameter is regular, we must have

$$(11.64) \quad \max(|P_1|, |Q_1|) \leq \varepsilon_0^\beta.$$

Assume that $\mathcal{E}^+(P_0, \dots, P_{k+1})$ is not-empty. We already know that Q_{k+1} is critical. It is also true that, for any parameter interval I containing the given parameter value, Q_k and P_{k+1} cannot be I -transverse: if they were, parabolic composition would produce children of P_k whose union contains $\mathcal{E}_+(P_0, \dots, P_{k+1})$, in contradiction with the definition of the parabolic core of P_k .

Let I be the smallest parameter interval satisfying

$$(11.65) \quad |I| \geq 2(\max |Q_k|, |P_{k+1}|)^{1-\eta}.$$

As Q_k, P_{k+1} are not I -transverse, it follows from Proposition 9 in Subsection 6.6 that P_{k+1} is I -critical. Therefore, we must have

$$(11.66) \quad \max(|P_{k+1}|, |Q_{k+1}|) < |I|^\beta$$

which implies

$$(11.67) \quad \max(|P_{k+1}|, |Q_{k+1}|) \leq C|Q_k|^{\tilde{\beta}},$$

with $\tilde{\beta} = \beta(1-\eta)(1+\tau)^{-1}$. Taking $\hat{\beta} < \tilde{\beta}$ but close to β and ε_0 sufficiently small, (11.67) and (11.64) give

$$(11.68) \quad \max(|P_j|, |Q_j|) \leq \varepsilon_0^{\hat{\beta}^j}.$$

11.8 Size and Area of Parabolic Cores

Proposition 50. *Let $(P, Q, n) \in \mathcal{C}_-$. With Leb standing for Lebesgue measure, we have*

$$(11.69) \quad \text{diam}(g^n(c(P))) \leq C|Q|^{\frac{1}{2}(1-\eta)}$$

$$(11.70) \quad \text{Leb}(g^n(c(P))) \leq C|Q|^{\frac{3}{2}-\frac{1}{2}\eta}$$

$$(11.71) \quad \text{Leb}(c(P)) \leq C|P||Q|^{\frac{1}{2}(1-\eta)}.$$

Remark. *A posteriori, $c(P)$, which is contained in $W^s(\Lambda, \hat{R})$, will have zero Lebesgue measure. However, we estimate here the diameter and Lebesgue measure of a larger set, as will be apparent in the proof.*

Proof. We start with a general observation on an affine-map with implicit representation (A, B) . The Jacobian of the map is the product $A_x^{-1}B_y$. The distortion of Lebesgue measure under the map, which is produced by the oscillation of the logarithm of the Jacobian, is, therefore, controlled by the distortion of the affine-like map in the sense of Subsection 3.2. In particular, the distortion of Lebesgue measure by the restriction of iterates corresponding to the elements of \mathcal{R} is uniformly bounded.

Thus, the third inequality (11.71) in the proposition is a consequence of the second. On the other hand, as $g^n(c(P)) \subset Q$, the second inequality (11.70) is an obvious consequence of the first. We have, therefore, only to prove (11.69). Set

$$(11.72) \quad Z = g^{n+N_0}(c(P)).$$

This set is contained in $g^{N_0}(Q \cap L_u) \cap W^s(\Lambda, \widehat{R})$, and, a fortiori, in P_s . Let $R(Z)$ be the set of $(P^*, Q^*, n^*) \in \mathcal{R}$ with $P^* \cap Z \neq \emptyset$, $P^* \subset P_s$.

By the definition of the parabolic core, Q and P^* cannot be I -transverse (for any parameter interval I). Assume that $|P^*| > 2|Q|$; it follows from Proposition 10 in Subsection 6.6 that P^* is I -critical (for any I); this implies that P^* is I -decomposable if I is small enough.

We conclude that

$$(11.73) \quad Z \subset \bigcup_{\substack{(P^*, Q^*, n^*) \in R(Z) \\ |P^*| \leq 2|Q|}} [P^* \cap g^{N_0}(Q \cap L_u)].$$

We replace $P^* \cap g^{N_0}(Q \cap L_u)$ by the larger Jordan domain $V(P^*, Q)$ defined as follows (see figure 9): the boundary of $V(P^*, Q)$ is made of one arc in the boundary of P^* and one arc in the boundary of $g^{N_0}(Q \cap L_u)$ and the interior of $V(P^*, Q)$ meets both P^* and $g^{N_0}(Q \cap L_u)$.

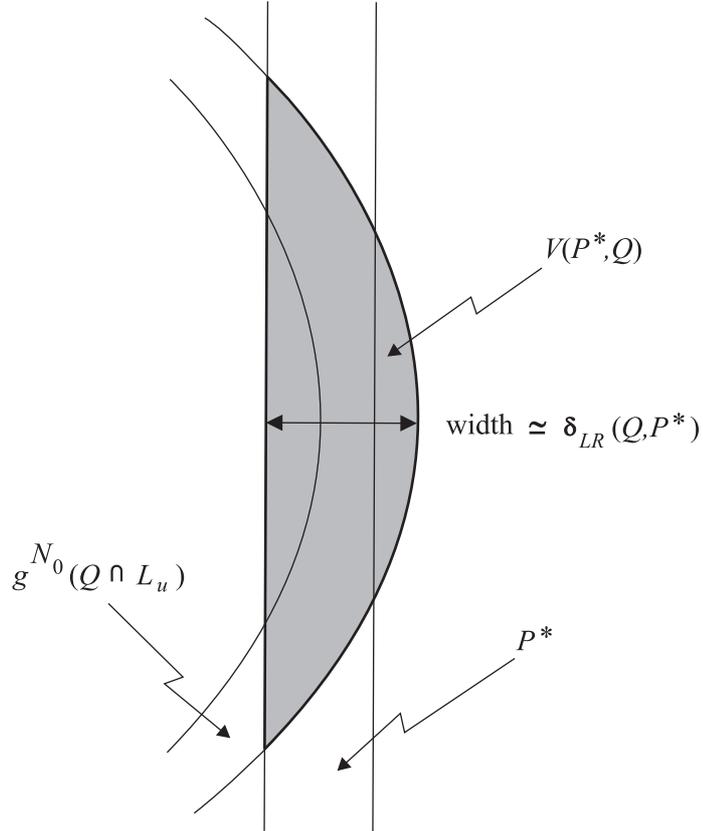


Figure 9

Observe that, if (P_1^*, Q_1^*, n_1^*) , (P_2^*, Q_2^*, n_2^*) belong to $R(Z)$, we have either $V(P_1^*, Q) \subset V(P_2^*, Q)$ or the opposite inclusion. Moreover, the parabolic geometry of the picture (cf. (3.22)) gives the

estimate

$$(11.74) \quad \text{diam } V(P^*, Q) \leq C(\delta_{LR}(Q, P^*))^{\frac{1}{2}}.$$

But, for $|P^*| \leq 2|Q|$, the fact that Q and P^* are not I -transverse (for any I) implies that

$$(11.75) \quad \delta_{LR}(Q, P^*) \leq C|Q|^{1-\eta}.$$

We conclude that

$$(11.76) \quad \begin{aligned} \text{diam } Z &\leq \sup_{\substack{(P^*, Q^*, n^*) \in R(Z) \\ |P^*| \leq 2|Q|}} \text{diam } V(P^*, Q) \\ &\leq C|Q|^{\frac{1}{2}(1-\eta)} \end{aligned}$$

Taking the image by the fixed map g^{-N_0} yields (11.69). □

11.9 Proof of Theorem 4

We will estimate first the Lebesgue measure of each domain $\mathcal{E}^+(P_0, \dots, P_k)$. We have

$$(11.77) \quad g^{m_{k-1}+N_0}(\mathcal{E}^+(P_0, \dots, P_k)) \subset c(P_k).$$

We now use that both the fixed map g^{N_0} and the affine-like iterates $g^{n_j} : P_j \rightarrow Q_j$ (for $0 \leq j < k$) have uniformly bounded distortion with respect to Lebesgue measure. We are, therefore, able to deduce from (11.70) in Proposition 50 that

$$(11.78) \quad \text{Leb}(\mathcal{E}^+(P_0, \dots, P_k)) \leq C^{k+1}|Q_k|^{\frac{3}{2}-\frac{1}{2}\eta} \prod_0^k \frac{|P_j|}{|Q_j|}.$$

By (11.67), we have $|P_{j+1}| \ll |Q_j|$ for $j > 0$ and it is easy to check that this still holds for $j = 0$ (using (11.64) if $|Q_0| \geq \varepsilon_0$, (11.67) otherwise). It then follows from (11.78) that we have (for $k > 0$)

$$(11.79) \quad \text{Leb}(\mathcal{E}^+(P_0, \dots, P_k)) \ll |P_0||Q_k|^{\frac{1}{2}(1-\eta)}.$$

To obtain the estimate for \mathcal{E}^+ , we have to sum over sequences (P_0, \dots, P_k) . We first estimate, when (P_0, Q_0, n_0) and (P_k, Q_k, n_k) are fixed, how many admissible sequences have these two extremities.

The element $(P_{k-1}, Q_{k-1}, n_{k-1})$ must satisfy

$$(11.80) \quad |Q_{k-1}| \geq C^{-1} \max(|P_k| |Q_k|)^{1/\tilde{\beta}}$$

and also that Q_{k-1}, P_k are not I -transverse (for any I). This implies that there is at most

$$(11.81) \quad C|Q_k|^{-\eta/\tilde{\beta}}$$

possibilities for $(P_{k-1}, Q_{k-1}, n_{k-1})$. Iterating leads to a total number of sequences (with P_0, Q_k fixed) which is at most

$$(11.82) \quad C|Q_k|^{-C\eta}.$$

Therefore, we obtain

$$(11.83) \quad \text{Leb}(\mathcal{E}^+) \leq \sum_{P_0, Q_k} |P_0| |Q_k|^{\frac{1}{2}-C\eta}.$$

The sum $\sum |P_0|$ is obviously bounded. The sum $\sum_{Q_k} |Q_k|^{\frac{1}{2}-C\eta}$ is arbitrarily small (when k is large). Indeed, we have from (11.68)

$$(11.84) \quad |Q_k| \leq \varepsilon_0^{\widehat{\beta}^k}.$$

On the other hand, recall that Q_k is I -critical (for any I). As the parameter is strongly regular, we know, for instance, from $(\text{SR1})_{\widehat{u}}$, that the series

$$(11.85) \quad \sum_{Q \text{ critical}} |Q|^d$$

is convergent if $d > d_s^0 + d_u^0 - 1 + o(1)$. As the maximal value taken by $d_s^0 + d_u^0 - 1$ under hypothesis (H4) is $\frac{1}{5}$, we indeed have

$$(11.86) \quad \lim_{k \rightarrow +\infty} \sum_{Q_k} |Q_k|^{\frac{1}{2}-C\eta} = 0,$$

and this concludes the proof of Theorem 4 and, thus, of the Main Theorem in the paper.

We can sum up the results in Sections 10 and 11 by rephrasing our main result as follows:

Theorem 5. *Assume (H1)–(H4). Then, for most $g \in \mathcal{U}_+$, $\Lambda_g \subset W^s(\Lambda_g)$ and $\Lambda_g \subset W^u(\Lambda_g)$ carry geometric invariant measures, à la Sinai-Ruelle-Bowen [Si, Ru, BR], with non-zero Lyapunov exponents. Both $W^s(\Lambda_g)$ and $W^u(\Lambda_g)$ have Lebesgue measure zero and thus Λ_g carries no attractors nor repellers.*

Appendix A

Composition Formulas for Affine-Like Maps

We mostly recall in this appendix the formulas for the simple and parabolic compositions of implicitly defined affine-like maps.

We follow closely [PY2]. The main difference with [PY2] is that we consider maps depending on a parameter t , and we are interested also in some partial derivatives with respect to the parameter.

A.1 Simple Composition. Here we consider a map $F_t : (x_0, y_0) \mapsto (x_1, y_1)$ implicitly defined by

$$(A.1) \quad \begin{cases} x_0 = A(y_0, x_1, t) \\ y_1 = B(y_0, x_1, t). \end{cases}$$

and a map $F'_t : (x_1, y_1) \mapsto (x_2, y_2)$ implicitly defined by

$$(A.2) \quad \begin{cases} x_1 = A'(y_1, x_2, t) \\ y_2 = B'(y_1, x_2, t). \end{cases}$$

The composition $F''_t = F'_t \circ F_t$ is implicitly defined by

$$(A.3) \quad \begin{cases} x_0 = A''(y_0, x_2, t) \\ y_2 = B''(y_0, x_2, t) \end{cases}$$

and we want to relate the partial derivatives of A'' , B'' to those of A , B , A' , B' . Set

$$(A.4) \quad \Delta := 1 - A'_y(y_1, x_2, t)B_x(y_0, x_1, t).$$

When we solve the system (A.1), (A.2) for x_1 , y_1 , we obtain

$$(A.5) \quad \begin{cases} x_1 = X(y_0, x_2, t) \\ y_1 = Y(y_0, x_2, t) \end{cases}$$

where the partial derivatives of X , Y are given by

$$(A.6) \quad \begin{cases} X_x = A'_x \Delta^{-1} \\ X_y = A'_y B_y \Delta^{-1} \\ X_t = (A'_t + A'_y B_t) \Delta^{-1} \\ Y_x = A'_x B_x \Delta^{-1} \\ Y_y = B_y \Delta^{-1} \\ Y_t = (B_t + A'_t B_x) \Delta^{-1}. \end{cases}$$

We have

$$(A.7) \quad \begin{cases} A''(y_0, x_2, t) = A(y_0, X, t) \\ B''(y_0, x_2, t) = B'(Y, x_2, t), \end{cases}$$

which gives

$$(A.8) \quad \begin{cases} A''_x = A_x A'_x \Delta^{-1} \\ B''_y = B'_y B_y \Delta^{-1}, \end{cases}$$

$$(A.9) \quad \begin{cases} A''_y = A_y + A_x X_y \\ B''_x = B'_x + B'_y Y_x, \end{cases}$$

$$(A.10) \quad \begin{cases} A''_t = A_t + A_x X_t \\ B''_t = B'_t + B'_y Y_t. \end{cases}$$

Next, from (A.4), we have

$$(A.11) \quad \begin{cases} -\Delta_x = B_{xx} X_x A'_y + B_x A'_{xy} + B_x A'_{yy} Y_x \\ -\Delta_y = A'_{yy} Y_y B_x + A'_y B_{xy} + A'_y B_{xx} X_y \\ -\Delta_t = B_{xt} A'_y + B_{xx} X_t A'_y + B_x A'_{yt} + B_x A'_{yy} Y_t. \end{cases}$$

Taking logarithmic derivatives in (A.8) gives

$$(A.12) \quad \partial_x \log |A''_x| = \partial_x \log |A'_x| + Y_x \partial_y \log |A'_x| + X_x \partial_x \log |A_x| - \Delta_x \Delta^{-1},$$

$$(A.13) \quad \partial_y \log |A''_x| = \partial_y \log |A_x| + X_y \partial_x \log |A_x| + Y_y \partial_y \log |A'_x| - \Delta_y \Delta^{-1},$$

$$(A.14) \quad \partial_t \log |A''_x| = \partial_t \log |A_x| + \partial_t \log |A'_x| + X_t \partial_x \log |A_x| + Y_t \partial_y \log |A'_x| - \Delta_t \Delta^{-1},$$

$$(A.15) \quad \partial_y \log |B''_y| = \partial_y \log |B_y| + X_y \partial_x \log |B_y| + Y_y \partial_y \log |B'_y| - \Delta_y \Delta^{-1},$$

$$(A.16) \quad \partial_x \log |B''_y| = \partial_x \log |B'_y| + Y_x \partial_y \log |B'_y| + X_x \partial_x \log |B_y| - \Delta_x \Delta^{-1},$$

$$(A.17) \quad \partial_t \log |B''_y| = \partial_t \log |B'_y| + \partial_t \log |B_y| + Y_t \partial_y \log |B'_y| + X_t \partial_x \log |B_y| - \Delta_t \Delta^{-1}.$$

Taking derivatives in (A.9) gives

$$(A.18) \quad \begin{cases} A''_{yy} = A_{yy} + 2A_{xy} X_y + A_{xx} X_y^2 + A_x X_{yy} \\ B''_{xx} = B'_{xx} + 2B'_{xy} Y_x + B'_{yy} Y_x^2 + B'_y Y_{xx}, \end{cases}$$

$$(A.19) \quad \begin{cases} A''_{yt} = A_{yt} + X_t A_{xy} + X_y A_{xt} + X_t X_y A_{xx} + A_x X_{yt} \\ B''_{xt} = B'_{xt} + Y_t B'_{xy} + Y_x B'_{yt} + Y_t Y_x B'_{yy} + B'_y Y_{xt}, \end{cases}$$

where the partial derivatives of X, Y are obtained from (A.6):

$$(A.20) \quad X_{yy} = B_y \Delta^{-1} (A'_{yy} Y_y + A'_y \partial_y \log |B_y| + A'_y X_y \partial_x \log |B_y| - A'_y \Delta_y \Delta^{-1}),$$

$$(A.21) \quad Y_{xx} = A'_x \Delta^{-1} (B_{xx} X_x + B_x \partial_x \log |A'_x| + B_x Y_x \partial_y \log |A'_x| - B_x \Delta_x \Delta^{-1}),$$

$$(A.22) \quad X_{yt} = B_y \Delta^{-1} (A'_{yy} Y_t + A'_{yt} + A'_y \partial_t \log |B_y| + A'_y X_t \partial_x \log |B_y| - A'_y \Delta_t \Delta^{-1}),$$

$$(A.23) \quad Y_{xt} = A'_x \Delta^{-1} (B_{xx} X_t + B_{xt} + B_x \partial_t \log |A'_x| + B_x Y_t \partial_y \log |A'_x| - B_x \Delta_t \Delta^{-1}).$$

A.2 Parabolic Composition. We have now a fold map $G_t = G_+ \circ G_0 \circ G_-$:

$$(x_u, y_u) \xrightarrow{G_-} (w, y_u) \xrightarrow{G_0} (x_s, w) \xrightarrow{G_+} (x_s, y_s),$$

with

$$(A.24) \quad \begin{cases} y_s = Y_s(w, x_s, t) \\ x_u = X_u(w, y_u, t), \end{cases}$$

$$(A.25) \quad w^2 = \theta(y_u, x_s, t).$$

We also have an affine like map $F_0 : (x_0, y_0) \mapsto (x_u, y_u)$ implicitly defined by

$$(A.26) \quad \begin{cases} x_0 = A_0(y_0, x_u, t) \\ y_u = B_0(y_0, x_u, t), \end{cases}$$

and another affine-like map $F_1 : (x_1, y_1) \mapsto (x_s, y_s)$ implicitly defined by

$$(A.27) \quad \begin{cases} x_s = A_1(y_s, x_1, t) \\ y_1 = B_1(y_s, x_1, t). \end{cases}$$

We assume that (PC1), (PC2) in Subsection 3.5 are satisfied. As we have seen in [PY2] and Subsection 3.5, the first step is to write

$$(A.28) \quad \begin{cases} x_u = X(w, y_0, t) \\ y_s = Y(w, x_1, t), \end{cases}$$

where the partial derivatives of X, Y are given by

$$(A.29) \quad \begin{cases} X_w = X_{u,w} \Delta_0^{-1} \\ X_y = X_{u,y} B_{0,y} \Delta_0^{-1} \\ X_t = (X_{u,t} + X_{u,y} B_{0,t}) \Delta_0^{-1} \\ Y_w = Y_{s,w} \Delta_1^{-1} \\ Y_x = Y_{s,x} A_{1,x} \Delta_1^{-1} \\ Y_t = (Y_{s,t} + Y_{s,x} A_{1,t}) \Delta_1^{-1}, \end{cases}$$

$$(A.30) \quad \begin{cases} \Delta_0 = 1 - X_{u,y} B_{0,x} \\ \Delta_1 = 1 - Y_{s,x} A_{1,y}. \end{cases}$$

We set

$$(A.31) \quad \begin{cases} \bar{Y}(w, y_0, t) := B_0(y_0, X, t) \\ \bar{X}(w, x_1, t) := A_1(Y, x_1, t), \end{cases}$$

$$(A.32) \quad C(w, y_0, x_1, t) := w^2 - \theta(\bar{X}, \bar{Y}, t).$$

The partial derivatives are given by

$$(A.33) \quad \begin{cases} \bar{Y}_w = B_{0,x} X_w \\ \bar{Y}_y = B_{0,y} + B_{0,x} X_y = B_{0,y} \Delta_0^{-1} \\ \bar{Y}_t = B_{0,t} + B_{0,x} X_t = (B_{0,t} + B_{0,x} X_{u,t}) \Delta_0^{-1}, \end{cases}$$

$$(A.34) \quad \begin{cases} \bar{X}_w = A_{1,y} Y_w \\ \bar{X}_x = A_{1,x} + A_{1,y} Y_x = A_{1,x} \Delta_1^{-1} \\ \bar{X}_t = A_{1,t} + A_{1,y} Y_t = (A_{1,t} + A_{1,y} Y_{s,t}) \Delta_1^{-1}, \end{cases}$$

$$(A.35) \quad \begin{cases} -C_w = -2w + \theta_x \bar{X}_w + \theta_y \bar{Y}_w \\ -C_x = \theta_x \bar{X}_x \\ -C_y = \theta_y \bar{Y}_y \\ -C_t = \theta_x \bar{X}_t + \theta_y \bar{Y}_t + \theta_t. \end{cases}$$

We solve

$$(A.36) \quad C(w, y_0, x_1, t) = 0$$

to define

$$(A.37) \quad w = W(y_0, x_1, t)$$

(there are two solutions W^+ and W^-).

The corresponding branch of the parabolic composition is implicitly defined by

$$(A.38) \quad \begin{cases} x_0 = A_0(y_0, X(W, y_0, t), t) =: A(y_0, x_1, t) \\ y_1 = B_1(Y(W, x_1, t), x_1, t) =: B(y_0, x_1, t). \end{cases}$$

The partial derivatives of A , B , W are given by

$$(A.39) \quad \begin{cases} A_x = A_{0,x} X_w W_x \\ A_y = A_{0,y} + A_{0,x} (X_y + X_w W_y) \\ A_t = A_{0,t} + A_{0,x} (X_t + X_w W_t) \end{cases}$$

$$(A.40) \quad \begin{cases} B_y = B_{1,y} Y_w W_y \\ B_x = B_{1,x} + B_{1,y} (Y_x + Y_w W_x) \\ B_t = B_{1,t} + B_{1,y} (Y_t + Y_w W_t) \end{cases}$$

$$(A.41) \quad \begin{cases} W_x = -C_x C_w^{-1} \\ W_y = -C_y C_w^{-1} \\ W_t = -C_t C_w^{-1}. \end{cases}$$

Substituting (A.29), (A.41), (A.35), (A.34) in the formulas (A.39)–(A.40) leads to

$$(A.42) \quad \begin{cases} A_x = A_{0,x} A_{1,x} C_w^{-1} \theta_x X_{u,w} \Delta_0^{-1} \Delta_1^{-1} \\ B_y = B_{1,y} B_{0,y} C_w^{-1} \theta_y Y_{s,w} \Delta_0^{-1} \Delta_1^{-1} \end{cases}$$

$$(A.43) \quad \begin{cases} A_y = A_{0,y} + A_{0,x} B_{0,y} \Delta_0^{-1} (X_{u,y} + X_{u,w} \theta_y \Delta_0^{-1} C_w^{-1}) \\ B_x = B_{1,x} + B_{1,y} A_{1,x} \Delta_1^{-1} (Y_{s,x} + Y_{s,w} \theta_x \Delta_1^{-1} C_w^{-1}) \end{cases}$$

$$(A.44) \quad \begin{cases} A_t = A_{0,t} + A_{0,x} \Delta_0^{-1} [X_{u,t} + X_{u,y} B_{0,t} + X_{u,w} C_w^{-1} (\theta_t + \theta_x \bar{X}_t + \theta_y \bar{Y}_t)] \\ B_t = B_{1,t} + B_{1,y} \Delta_1^{-1} [Y_{s,t} + Y_{s,x} A_{1,t} + Y_{s,w} C_w^{-1} (\theta_t + \theta_x \bar{X}_t + \theta_y \bar{Y}_t)] \end{cases}$$

Taking the logarithmic derivatives in the first formula of (A.39), we obtain

$$(A.45) \quad \partial_x \log |A_x| = W_x X_w \partial_x \log |A_{0,x}| + W_x \partial_w \log |X_w| + \partial_x \log |W_x|,$$

$$(A.46) \quad \begin{aligned} \partial_y \log |A_x| &= \partial_y \log |A_{0,x}| + \partial_x \log |A_{0,x}|(X_y + X_w W_y) \\ &\quad + \partial_y \log |X_w| + W_y \partial_w \log |X_w| + \partial_y \log |W_x|, \end{aligned}$$

$$(A.47) \quad \begin{aligned} \partial_t \log |A_x| &= \partial_t \log |A_{0,x}| + \partial_x \log |A_{0,x}|(X_t + X_w W_t) \\ &\quad + \partial_t \log |X_w| + W_t \partial_w \log |X_w| + \partial_t \log |W_x|. \end{aligned}$$

From the second formula in (A.39), one gets

$$(A.48) \quad \begin{aligned} A_{yy} &= A_{0,yy} + 2A_{0,xy}(X_y + X_w W_y) + A_{0,xx}(X_y + X_w W_y)^2 \\ &\quad + A_{0,x}(X_{yy} + 2X_{wy}W_y + X_{ww}W_y^2 + X_w W_{yy}), \end{aligned}$$

$$(A.49) \quad \begin{aligned} A_{yt} &= A_{0,yt} + A_{0,xy}(X_y + X_w W_t) + A_{0,xt}(X_y + X_w W_y) \\ &\quad + A_{0,xx}(X_t + X_w W_t)(X_y + X_w W_y) \\ &\quad + A_{0,x}(X_{yt} + X_{wy}W_t + X_{wt}W_y + X_{ww}W_yW_t + X_w W_{yt}). \end{aligned}$$

The symmetric formulas for B are

$$(A.50) \quad \partial_y \log |B_y| = W_y Y_w \partial_y \log |B_{1,y}| + \partial_y \log |W_y| + W_y \partial_w \log |Y_w|,$$

$$(A.51) \quad \begin{aligned} \partial_x \log |B_y| &= \partial_x \log |B_{1,y}| + \partial_x \log |Y_w| \\ &\quad + \partial_x \log |W_y| + \partial_y \log |B_{1,y}|(Y_x + Y_w W_x) + W_x \partial_w \log |Y_w|, \end{aligned}$$

$$(A.52) \quad \begin{aligned} \partial_t \log |B_y| &= \partial_t \log |B_{1,y}| + \partial_y \log |B_{1,y}|(Y_t + Y_w W_t) \\ &\quad + \partial_t \log |Y_w| + W_t \partial_w \log |Y_w| + \partial_t \log |W_y|, \end{aligned}$$

$$(A.53) \quad \begin{aligned} B_{xx} &= B_{1,xx} + 2B_{1,xy}(Y_x + Y_w W_x) + B_{1,yy}(Y_x + Y_w W_x)^2 \\ &\quad + B_{1,y}(Y_{xx} + 2Y_{wx}W_x + Y_{ww}W_x^2 + Y_w W_{xx}), \end{aligned}$$

$$(A.54) \quad \begin{aligned} B_{xt} &= B_{1,xt} + B_{1,xy}(Y_t + Y_w W_t) + B_{1,yt}(Y_x + Y_w W_x) \\ &\quad + B_{1,yy}(Y_x + Y_w W_x)(Y_t + Y_w W_t) \\ &\quad + B_{1,y}(Y_{xt} + Y_{wx}W_t + Y_{wt}W_x + Y_{ww}W_xW_t + Y_w W_{xt}). \end{aligned}$$

In formulas (A.45)–(A.54), the partial derivatives of order 2 of W are obtained from (A.41):

$$(A.55) \quad \begin{cases} W_{xx} = -C_w^{-1}(C_{ww}W_x^2 + 2C_{wx}W_x + C_{xx}) \\ W_{xy} = -C_w^{-1}(C_{ww}W_xW_y + C_{wx}W_y + C_{wy}W_x + C_{xy}) \\ W_{yy} = -C_w^{-1}(C_{ww}W_y^2 + 2C_{wy}W_y + C_{yy}) \\ W_{xt} = -C_w^{-1}(C_{ww}W_xW_t + C_{wx}W_t + C_{wt}W_x + C_{xt}) \\ W_{yt} = -C_w^{-1}(C_{ww}W_yW_t + C_{wy}W_t + C_{wt}W_y + C_{yt}) \end{cases}$$

The partial derivatives of order 2 of C are obtained from (A.35):

$$(A.56) \quad -C_{ww} = -2 + \theta_x \bar{X}_{ww} + \theta_y \bar{Y}_{ww} + \theta_{xx} \bar{X}_w^2 + 2\theta_{xy} \bar{X}_w \bar{Y}_w + \theta_{yy} \bar{Y}_w^2,$$

$$(A.57) \quad \begin{cases} -C_{wx} = \theta_{xx} \bar{X}_w \bar{X}_x + \theta_{xy} \bar{Y}_w \bar{X}_x + \theta_x \bar{X}_{wx} \\ -C_{wy} = \theta_{yy} \bar{Y}_w \bar{Y}_y + \theta_{xy} \bar{X}_w \bar{Y}_y + \theta_y \bar{Y}_{wy} \end{cases}$$

$$(A.58) \quad -C_{wt} = \theta_{xt} \bar{X}_w + \theta_{xx} \bar{X}_w \bar{X}_t + \theta_{xy} (\bar{X}_w \bar{Y}_t + \bar{Y}_w \bar{X}_t) + \theta_{yy} \bar{Y}_w \bar{Y}_t + \theta_{yt} \bar{Y}_w + \theta_x \bar{X}_{wt} + \theta_y \bar{Y}_{wt},$$

$$(A.59) \quad \begin{cases} -C_{xx} = \theta_{xx} \bar{X}_x^2 + \theta_x \bar{X}_{xx} \\ -C_{xy} = \theta_{xy} \bar{X}_x \bar{Y}_y \\ -C_{yy} = \theta_{yy} \bar{Y}_y^2 + \theta_y \bar{Y}_{yy}, \end{cases}$$

$$(A.60) \quad \begin{cases} -C_{xt} = \theta_{xt} \bar{X}_x + \theta_{xx} \bar{X}_x \bar{X}_t + \theta_{xy} \bar{X}_x \bar{Y}_t + \theta_x \bar{X}_{xt} \\ -C_{yt} = \theta_{yt} \bar{Y}_y + \theta_{yy} \bar{Y}_y \bar{Y}_t + \theta_{xy} \bar{Y}_y \bar{X}_t + \theta_y \bar{Y}_{yt}. \end{cases}$$

The partial derivatives of order 2 of \bar{X} , \bar{Y} are obtained from (A.34):

$$(A.61) \quad \begin{cases} \bar{X}_{ww} = A_{1,yy} Y_w^2 + A_{1,y} Y_{ww} \\ \bar{X}_{wx} = A_{1,xy} Y_w + A_{1,yy} Y_w Y_x + A_{1,y} Y_{wx} \\ \bar{X}_{wt} = A_{1,yt} Y_w + A_{1,yy} Y_w Y_t + A_{1,y} Y_{wt} \\ \bar{X}_{xx} = A_{1,xx} + 2A_{1,xy} Y_x + A_{1,yy} Y_x^2 + A_{1,y} Y_{xx} \\ \bar{X}_{xt} = A_{1,xt} + A_{1,xy} Y_t + A_{1,yt} Y_x + A_{1,yy} Y_x Y_t + A_{1,y} Y_{xt}. \end{cases}$$

$$(A.62) \quad \begin{cases} \bar{Y}_{ww} = B_{0,xx} X_w^2 + B_{0,x} X_{ww} \\ \bar{Y}_{wy} = B_{0,xy} X_w + B_{0,xx} X_w X_y + B_{0,x} X_{wy} \\ \bar{Y}_{wt} = B_{0,xt} X_w + B_{0,xx} X_w X_t + B_{0,x} X_{wt} \\ \bar{Y}_{yy} = B_{0,yy} + 2B_{0,xy} X_y + B_{0,xx} X_y^2 + B_{0,x} X_{yy} \\ \bar{Y}_{yt} = B_{0,yt} + B_{0,xy} X_t + B_{0,xt} X_y + B_{0,xx} X_y X_t + B_{0,x} X_{yt}. \end{cases}$$

Finally, from (A.29), we obtain

$$(A.63) \quad \begin{cases} X_{ww} = \Delta_0^{-1} (X_{u,ww} + 2X_{u,wy} \bar{Y}_w + X_{u,yy} \bar{Y}_w^2 + X_{u,y} B_{0,xx} X_w^2), \\ X_{wy} = \Delta_0^{-1} (X_{u,wy} \bar{Y}_y + X_{u,yy} \bar{Y}_w \bar{Y}_y + X_{u,y} X_w (B_{0,xy} + B_{0,xx} X_y)), \\ X_{yy} = \Delta_0^{-1} (X_{u,yy} \bar{Y}_y^2 + X_{u,y} (B_{0,yy} + 2B_{0,xy} X_y + B_{0,xx} X_y^2)) \\ X_{wt} = \Delta_0^{-1} [X_{u,wt} + X_{u,wy} \bar{Y}_t + X_{u,yy} \bar{Y}_w \bar{Y}_t + X_{u,yt} \bar{Y}_w \\ \quad + X_{u,y} X_w (B_{0,xt} + B_{0,xx} X_t)], \\ X_{yt} = \Delta_0^{-1} [X_{u,yt} \bar{Y}_y + X_{u,yy} \bar{Y}_y \bar{Y}_t \\ \quad + X_{u,y} (B_{0,yt} + B_{0,xy} X_t + B_{0,xt} X_y + B_{0,xx} X_y X_t)], \end{cases}$$

$$(A.64) \quad \left\{ \begin{array}{l} Y_{ww} = \Delta_1^{-1}(Y_{s,ww} + 2Y_{s,wx}\bar{X}_w + Y_{s,xx}\bar{X}_w^2 + Y_{s,x}A_{1,yy}Y_w^2), \\ Y_{wx} = \Delta_1^{-1}(Y_{s,wx}\bar{X}_x + Y_{s,xx}\bar{X}_w\bar{X}_x + Y_{s,x}Y_w(A_{1,xy} + A_{1,yy}Y_x)), \\ Y_{xx} = \Delta_1^{-1}(Y_{s,xx}\bar{X}_x^2 + Y_{s,x}(A_{1,xx} + 2A_{1,xy}Y_x + A_{1,yy}Y_x^2)) \\ Y_{wt} = \Delta_1^{-1}[Y_{s,wt} + Y_{s,wx}\bar{X}_t + Y_{s,xx}\bar{X}_w\bar{X}_t + Y_{s,xt}\bar{X}_w \\ \quad + Y_{s,x}Y_w(A_{1,yt} + A_{1,yy}Y_t)], \\ Y_{xt} = \Delta_1^{-1}[Y_{s,xt}\bar{X}_x + Y_{s,xx}\bar{X}_x\bar{X}_t \\ \quad + Y_{s,x}(A_{1,xt} + A_{1,xy}Y_t + A_{1,yt}Y_x + A_{1,yy}Y_xY_t)]. \end{array} \right.$$

Appendix B

On the Lipschitz Regularity of \tilde{R}_+^∞

B.1 In this appendix, we will perform some calculations and obtain some estimates that uphold the proof of Proposition 40 in Subsection 10.5, showing that the holonomy maps of the partial foliation $\tilde{\mathcal{R}}_+^\infty$ are uniformly Lipschitz.

As always, we have to consider two cases, dealing separately with affine-like maps and parabolic composition, the later case being more involved than the first.

B.2 The Case of Affine-Like Maps. We consider here, with the notations of Subsection 3.1, an affine-like map F with implicit representation (A, B) : the domain of F is a vertical strip P in a rectangle $I_0^s \times I_0^u$, its image is a horizontal strip in a rectangle $I_1^s \times I_1^u$, the respective coordinates being x_0, y_0, x_1, y_1 .

In $I_1^s \times I_1^s$, we are given a one-parameter family of vertical-like curves which are graphs

$$(B.1) \quad \omega(s) = \{x_1 = \varphi(y_1, s)\}, \text{ with } \left| \frac{\partial \varphi}{\partial y_1} \right| \text{ not too large.}$$

We assume that, for all y_1, s

$$(B.2) \quad \frac{\partial \varphi}{\partial s}(y_1, s) > 0$$

and that we have a uniform bound

$$(B.3) \quad \left| \frac{\partial}{\partial y_1} \log \frac{\partial \varphi}{\partial s}(y_1, s) \right| \leq T.$$

We define the vertical-like curves in P by

$$(B.4) \quad \Omega(s) = F^{-1}(\omega(s) \cap Q).$$

We assume that F satisfies a cone condition and we have a good control on the distortion of F . Under these hypotheses, $\Omega(s)$ will be a graph

$$(B.5) \quad \Omega(s) = \{x_0 = \Phi(y_0, s)\}$$

and we want to obtain the analogue of (B.3):

$$(B.6) \quad \left| \frac{\partial}{\partial y_0} \log \left| \frac{\partial \Phi}{\partial s}(y_0, s) \right| \right| \leq T'.$$

The relation between φ and Φ is as follows in the equation

$$(B.7) \quad y_1 = B(y_0, \varphi(y_1, s)),$$

we solve for y_1 to define

$$(B.8) \quad y_1 = \psi(y_0, s).$$

We then have

$$(B.9) \quad \Phi(y_0, s) = A\left(y_0, \varphi(\psi(y_0, s), s)\right).$$

From (B.7), (B.8), we get

$$(B.10) \quad \frac{\partial \psi}{\partial s} = B_x \frac{\partial \varphi}{\partial s} \left(1 - B_x \frac{\partial \varphi}{\partial y_1}\right)^{-1}$$

(where, as always, the functions must be taken at the appropriate arguments); then, from (B.9) we get

$$(B.11) \quad \frac{\partial \Phi}{\partial s} = A_x \frac{\partial \varphi}{\partial s} \left(1 - B_x \frac{\partial \varphi}{\partial y_1}\right)^{-1}.$$

We must now take the logarithmic derivative of the right-hand term with respect to y_0 . It is the sum of three terms:

$$(B.12) \quad Z_1 = \partial_y \log |A_x| + \partial_x \log |A_x| \frac{\partial \varphi}{\partial y_1} \frac{\partial \psi}{\partial y_0},$$

$$(B.13) \quad Z_2 = \frac{\partial}{\partial y_1} \left(\log \frac{\partial \varphi}{\partial s} \right) \frac{\partial \psi}{\partial y_0},$$

$$(B.14) \quad Z_3 = \left(1 - B_x \frac{\partial \varphi}{\partial y_1}\right)^{-1} \left\{ \frac{\partial^2 \varphi}{\partial y_1^2} \frac{\partial \psi}{\partial y_0} B_x + \frac{\partial \varphi}{\partial y_1} \left(B_y \partial_x \log |B_y| + B_{xx} \frac{\partial \varphi}{\partial y_1} \frac{\partial \psi}{\partial y_0} \right) \right\}.$$

In these formulas, we have from (B.7), (B.8)

$$(B.15) \quad \frac{\partial \psi}{\partial y_0} = B_y \left(1 - B_x \frac{\partial \varphi}{\partial y_1}\right)^{-1}.$$

We see that we obtain (B.6) with

$$(B.16) \quad T' \leq aT + C.$$

Here, the constant C is bounded in terms of $\|(1 - B_x \frac{\partial \varphi}{\partial y_1})^{-1}\|_\infty$, the distortion of F , the cone condition, $\|\frac{\partial \varphi}{\partial y_1}\|_\infty$ and $\|\frac{\partial^2 \varphi}{\partial y_1^2}\|_\infty$. The constant a is given by:

$$(B.17) \quad a = \left\| \frac{\partial \psi}{\partial y_0} \right\|_\infty \leq C \|B_y\|_\infty$$

and, therefore, a will be $< \frac{1}{2}$ if $\|Q\| = \|B_y\|_\infty$ is sufficiently small.

Remark. The control of $\frac{\partial \Phi}{\partial y_0}$ is guaranteed by the cone condition and the control of $\frac{\partial^2 \Phi}{\partial y_0^2}$ by the distortion of F .

Thus, the mixed second derivative does not “explode” by taking pre-images by affine-like maps. \square

B.3 The Parabolic Case. We now use the setting and notations of Subsection 3.5. We have intervals $I_0^s, I_0^u, I_u^s, I_u^u, I_s^s, I_s^u$ with respective coordinates $x_0, y_0, x_u, y_u, x_s, y_s$.

We have an affine-like map F with domain $P \subset I_0^s \times I_0^u$, image $Q \subset I_u^s \times I_u^u$, and implicit representation (A, B) . We also have a folding map G with domain $L_u \subset I_u^s \times I_u^u$, image $L_s \subset I_s^s \times I_s^u$. The map G is implicitly defined by the system

$$(B.18) \quad \begin{cases} y_s = Y_s(w, x_s) \\ x_u = X_u(w, y_u) \\ w^2 = \theta(y_u, x_s) \end{cases}$$

with Y_s, X_u, θ as in Subsections 2.3 and 3.5. The affine-like map F should satisfy (cf. (PC1) in Subsection 3.5, (R4) in Subsection 5.3):

$$(B.19) \quad |B_x| < b, \quad |B_{xx}| < b.$$

with $b \ll 1$. As above we are given a one-parameter family of vertical-like curves which are graphs

$$(B.20) \quad \omega(s) = \{x_s = \varphi(y_s, s)\},$$

but these curves are now very close to being exactly vertical

$$(B.21) \quad \left| \frac{\partial \varphi}{\partial y_s} \right| < b, \quad \left| \frac{\partial^2 \varphi}{\partial y_s^2} \right| < b.$$

We also assume as above that

$$(B.22) \quad \frac{\partial \varphi}{\partial s}(y_s, s) > 0$$

and that we have some uniform bound

$$(B.23) \quad \left| \frac{\partial}{\partial s} \log \frac{\partial \varphi}{\partial s} \right| \leq T.$$

We then define the curves $\Omega^\pm(s)$ as the connected components of $F^{-1}(Q \cap G^{-1}(\omega(s) \cap L_s))$.

They should be graphs

$$(B.24) \quad \Omega^\pm(s) = \{x_0 = \Phi^\pm(y_0, s)\}$$

and we want some uniform control

$$(B.25) \quad \left| \frac{\partial}{\partial y_0} \log \left| \frac{\partial \Phi}{\partial s} \right| \right| \leq T'.$$

As in Subsection 3.5 (formulas (3.14), (3.15)) we eliminate y_u to write

$$(B.26) \quad x_u = X(w, y_0).$$

Similarly, in the equation

$$(B.27) \quad y_s = Y_s(w, \varphi(y_s, s)),$$

we solve for y_s to write

$$(B.28) \quad y_s = Y(w, s).$$

We then set

$$(B.29) \quad \bar{X}(w, s) := \varphi(Y(w, s), s),$$

$$(B.30) \quad \bar{Y}(w, y_0) := B(y_0, X(w, y_0)),$$

$$(B.31) \quad C(w, y_0, s) := w^2 - \theta(\bar{Y}(w, y_0), \bar{X}(w, s)).$$

Then, one should solve $C = 0$ to get

$$(B.32) \quad w = W^\pm(y_0, s),$$

$$(B.33) \quad \Phi(y_0, s) = A(y_0, X(W(y_0, s), y_0)),$$

with $W = W^\pm$, $\Phi = \Phi^\pm$.

The relations (B.26) through (B.33) are the equations which allow us to calculate $\frac{\partial}{\partial y_0} \log \left| \frac{\partial \Phi}{\partial s} \right|$.

From (B.33), we have

$$(B.34) \quad \frac{\partial \Phi}{\partial s} = A_x X_w \frac{\partial W}{\partial s}.$$

From the implicit definition of X

$$(B.35) \quad X(w, y_0) = X_u \left(w, B_0(y_0, X(w, y_0)) \right),$$

we get

$$(B.36) \quad X_w = (1 - X_{u,y} B_x)^{-1} X_{u,w}.$$

One has also

$$(B.37) \quad \frac{\partial W}{\partial s} = - \frac{\partial C}{\partial s} C_w^{-1},$$

$$(B.38) \quad \frac{\partial C}{\partial s} = - \theta_x \left(\frac{\partial \varphi}{\partial s} + \frac{\partial \varphi}{\partial y} \frac{\partial Y}{\partial s} \right).$$

From (B.27), (B.28), one obtains

$$(B.39) \quad \frac{\partial Y}{\partial s} = Y_{s,x} \frac{\partial \varphi}{\partial s} \left(1 - Y_{s,x} \frac{\partial \varphi}{\partial y} \right)^{-1},$$

and putting this in (B.38) leads to

$$(B.40) \quad \frac{\partial C}{\partial s} = - \theta_x \frac{\partial \varphi}{\partial s} \left(1 - Y_{s,x} \frac{\partial \varphi}{\partial y} \right)^{-1}.$$

Therefore, introducing (B.36), (B.37), (B.40) into (B.34), we have

$$(B.41) \quad \frac{\partial \Phi}{\partial s} = A_x X_{u,w} (1 - X_{u,y} B_x)^{-1} C_w^{-1} \theta_x \frac{\partial \varphi}{\partial s} \left(1 - Y_{s,x} \frac{\partial \varphi}{\partial y}\right)^{-1}.$$

We now take the logarithmic derivative of this product of seven terms with relation to y_0 . We obtain a sum of seven terms:

$$(B.42) \quad Z_1 = \partial_y \log |A_x| + \partial_x \log |A_x| (X_y + X_w W_y),$$

$$(B.43) \quad Z_2 = (\partial_w \log |X_{u,w}|) W_y + \partial_y \log |X_{u,w}| (\bar{Y}_y + \bar{Y}_w W_y),$$

$$(B.44) \quad Z_3 = (1 - X_{u,y} B_x)^{-1} \left\{ X_{uy} (B_{xy} + B_{xx} (X_y + X_w W_y)) \right. \\ \left. + B_x (X_{u,yw} W_y + X_{u,yy} (\bar{Y}_y + \bar{Y}_w W_y)) \right\},$$

$$(B.45) \quad Z_4 = C_w^{-1} \left[\theta_x \bar{X}_w W_y + \theta_y (\bar{Y}_y + \bar{Y}_w W_y) \right],$$

$$(B.46) \quad Z_5 = \partial_y \log |\theta_x| (\bar{Y}_y + \bar{Y}_w W_y) + \partial_x \log |\theta_x| \bar{X}_w W_y,$$

$$(B.47) \quad Z_6 = \left(\frac{\partial}{\partial y_s} \log \frac{\partial \varphi}{\partial s} \right) Y_w W_y,$$

$$(B.48) \quad Z_7 = \left(1 - Y_{s,x} \frac{\partial \varphi}{\partial y_s}\right)^{-1} \left\{ Y_{s,x} \frac{\partial^2 \varphi}{\partial y_s^2} Y_w W_y + \frac{\partial \varphi}{\partial y_s} (Y_{s,xw} W_y + Y_{s,xx} \bar{X}_w W_y) \right\}.$$

The terms in these expressions that have not yet been introduced are:

$$(B.49) \quad X_y = X_{u,y} B_y (1 - X_{u,y} B_x)^{-1},$$

$$(B.50) \quad W_y = -C_y C_w^{-1} = \theta_y \bar{Y}_y C_w^{-1},$$

$$(B.51) \quad \bar{Y}_y = B_y + B_x X_y,$$

$$(B.52) \quad \bar{Y}_w = B_x X_w,$$

$$(B.53) \quad \bar{X}_w = \frac{\partial \varphi}{\partial y_s} Y_w,$$

$$(B.54) \quad Y_w = Y_{s,w} \left(1 - Y_{s,x} \frac{\partial \varphi}{\partial y_s}\right)^{-1}.$$

We obtain the estimates

$$(B.55) \quad |X_w| \leq C, \quad |Y_w| \leq C,$$

$$(B.56) \quad |\bar{X}_w| \leq Cb, \quad |\bar{Y}_w| \leq Cb,$$

$$(B.57) \quad |X_y| \leq C|Q|, \quad |\bar{Y}_y| \leq C|Q|,$$

$$(B.58) \quad |W_y| \leq C|Q| |C_w^{-1}|,$$

$$(B.59) \quad |Z_i| \leq C + C|Q| |C_w^{-1}| \quad \text{for } i = 1, 2, 3, 5, 7;$$

$$(B.60) \quad |Z_4| \leq C|Q| |C_w^{-1}| + Cb|Q| |C_w^{-1}|^2;$$

$$(B.61) \quad |Z_6| \leq C|Q| |C_w^{-1}| T.$$

Therefore, we obtain (B.25) with

$$(B.62) \quad T' = C|Q| \|C_w^{-1}\|_\infty T + C + C|Q| \|C_w^{-1}\|_\infty + Cb|Q| \|C_w^{-1}\|_\infty^2.$$

Regarding C_w^{-1} , the same considerations as in Subsection 3.5 apply. One checks that C_{ww} is close to two, hence the value of $|C_w^{-1}|$ at the argument where C vanishes is of the order of $|\overline{C}|^{-\frac{1}{2}}$ where

$$(B.63) \quad \overline{C}(y_0, s) = \min_w C(w, y_0, s).$$

In the context of Section 10, we have

$$(B.64) \quad |Q| \ll |\overline{C}| \ll 1,$$

and, therefore,

$$(B.65) \quad T' \leq C|Q|^{\frac{1}{2}}T + C.$$

This shows that T' does not "blow up". □

B.4 In the context of Proposition 40 in Subsection 10.5, Case 2 in the proof, we had an element $(P_\ell, Q_\ell, n_\ell) \in \mathcal{R}$ and two non-simple children $P_{\ell+1}, P'_{\ell+1}$ of P_ℓ . Let $\omega_{\ell+1}, \omega'_{\ell+1}$ be components of the vertical part of the boundary of $P_{\ell+1}, P'_{\ell+1}$ respectively. We wanted to have (cf. (10.65))

$$(B.66) \quad \left| \log \frac{\varphi_{\ell+1}(y) - \varphi'_{\ell+1}(y)}{\varphi_{\ell+1}(y') - \varphi'_{\ell+1}(y')} \right| \leq C|y - y'|$$

where $\omega_{\ell+1} = \{x = \varphi_{\ell+1}(y)\}, \omega'_{\ell+1} = \{x = \varphi'_{\ell+1}(y)\}$.

To prove (10.66), we will imbed both $\omega_{\ell+1}$ and $\omega'_{\ell+1}$ in a one-parameter family of curves

$$(B.67) \quad \Omega(s) = \{x = \Phi(y, s)\}$$

with

$$(B.68) \quad \Omega(s_0) = \omega_{\ell+1}, \quad \Omega(s_1) = \omega'_{\ell+1}$$

$$(B.69) \quad \frac{\partial}{\partial s} \Phi > 0$$

$$(B.70) \quad \left| \frac{\partial}{\partial y} \log \frac{\partial \Phi}{\partial s} \right| \leq C.$$

Indeed, in this case, we can write

$$(B.71) \quad \frac{\varphi_{\ell+1}(y) - \varphi'_{\ell+1}(y)}{\varphi_{\ell+1}(y') - \varphi'_{\ell+1}(y')} = \frac{\int_{s_0}^{s_1} \frac{\partial \Phi}{\partial s}(y, s) ds}{\int_{s_0}^{s_1} \frac{\partial \Phi}{\partial s}(y', s) ds},$$

with, from (B.70), for any s

$$(B.72) \quad e^{-C|y-y'|} \leq \left(\frac{\partial \Phi}{\partial s}(y, s) \right)^{-1} \frac{\partial \Phi}{\partial s}(y', s) \leq e^{C|y-y'|},$$

which yields (B.66). □

Appendix C

A Toy Model for the Transversality Relation

C.1 Our goal in this appendix is to explain why the complicated definition of the transversality relation in Subsection 5.4, is in some way "natural", if we require some useful properties for the proof of our Main Theorem. The toy model that we are considering is an abstract one. It is much simpler than the real situation of Section 5 because the sets in which the relation takes place are well defined to begin with: in Section 5, we need to know the transversality relation in order to construct the classes $\mathcal{R}(I)$.

C.2 A partially ordered set X is a *forest* if, for any $x_0 \in X$, the set $\{x \geq x_0\}$ is finite and totally ordered. A *tree* is a forest with a single maximal element. Let X_1, \dots, X_n be forests and let A be a subset of $X = X_1 \times \dots \times X_n$ (one should think of A as the graph of an n -ary relation). We say that A is *hereditary* if whenever $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$ are such that $y_i \leq x_i$ for all $1 \leq i \leq n$ (abbreviated as $y \leq x$), then $y \in A$ if $x \in A$.

Two points $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$ of X are *coordinate-wise comparable* (c -comparable for short) if for each $i \in \{1, \dots, n\}$, we have $x_i \geq y_i$ or $x_i \leq y_i$. In this case, we set

$$(C.1) \quad x \vee y = (\max(x_i, y_i))_{1 \leq i \leq n}.$$

The set A is *concave* if, whenever $x, y \in A$ are c -comparable, then the point $x \vee y$ also belongs to A .

The intersection of hereditary, resp. concave, subsets of X is hereditary, resp. concave. It follows that any subset $A \subset X$ is contained in a smallest concave hereditary subset, called the $c.h$ -envelope and denoted by \widehat{A} .

Example. When the number of factors $n = 1$, any subset is concave: the $c.h$ -envelope of $A \subset X_1$ is the set of $x \in X_1$ such that $x \leq y$ for some $y \in A$.

C.3 We construct the $c.h$ -envelope when $n = 2$.

Proposition. Let X_1, X_2 be forests and A be a subset of $X = X_1 \times X_2$. Let A_1 be the set of $x \in X$ such that $x = y \vee z$ for some c -comparable $y, z \in A$. The $c.h$ -envelope of A is equal to the set A_2 of $t = (t_1, t_2)$ such that $t_1 \leq x_1, t_2 \leq x_2$ for some $x = (x_1, x_2) \in A_1$.

Proof. It is clear that the set A_2 defined in the proposition is hereditary and it is contained in the $c.h$ -envelope \widehat{A} of A . We have to prove that A_2 is concave. We first prove the

Lemma 18. If $y, z \in A_1$ are c -comparable, $y \vee z$ also belongs to A_1 .

Proof. By the definition of A_1 , we can write $y = y' \vee y''$, $z = z' \vee z''$ with y', y'', z', z'' in A , y' and y'' c -comparable, z' and z'' c -comparable. We may assume that $y_1 \leq z_1$ and $y_2 \geq z_2$, and also $z_1 = z'_1$, $y_2 = y'_2$; then, we have $z'_1 = z_1 \geq y_1 \geq y'_1$ and $y'_2 = y_2 \geq z_2 \geq z'_2$, hence y', z' are c -comparable with $y' \vee z' = y \vee z$. \square

End of the Proof of the Proposition. Let $t', t'' \in A_2$ be c -comparable and let $x', x'' \in A_1$ be such that $x'_i \geq t'_i$, $x''_i \geq t''_i$ for $i = 1, 2$. As X_1 and X_2 are forests, x' and x'' are c -comparable. From the lemma, $x' \vee x''$ belongs to A_1 ; then $t' \vee t''$ belongs to A_2 . \square

C.4 For $n \geq 3$, the situation is more complicated, as the two examples below indicate.

Example 1. Let X_1, X_2, X_3 be forests and let x, y, z be three points of $X = X_1 \times X_2 \times X_3$ such that

$$(C.2) \quad x_1 \geq y_1 \geq z_1,$$

$$(C.3) \quad y_2 \geq x_2, \quad y_2 \geq z_2,$$

$$(C.4) \quad z_3 \geq x_3, \quad z_3 \geq y_3.$$

Let $A = \{x, y, z\} \subset X$. If we define, as in the proposition above,

$$(C.5) \quad A_1 = \{u \vee v, \quad u, v \in A, \quad u, v \text{ } c\text{-comparable}\}$$

and if we assume that x_2, z_2 are *not* comparable and x_3, y_3 are *not* comparable, then we have

$$(C.6) \quad A_1 = \{x, y, z, \quad y \vee z = (y_1, y_2, z_3)\}.$$

On the other hand, the point $w = (x_1, y_2, z_3) = x \vee (y \vee z)$ certainly belongs to the $c.h$ -envelope of A , but does not satisfy $w_i \leq u_i$ ($i = 1, 2, 3$) for any $u \in A_1$. This example shows that the analogue of the proposition above is false for $n = 3$.

Example 2. Let X_1, X_2, X_3 forests and let $x, y, z \in X = X_1 \times X_2 \times X_3$ such that

$$(C.7) \quad x_1 \geq y_1, \quad x_1 \geq z_1$$

$$(C.8) \quad y_2 \geq x_2, \quad y_2 \geq z_2$$

$$(C.9) \quad z_3 \geq x_3, \quad z_3 \geq y_3$$

but none of the pairs (y_1, z_1) , (x_2, z_2) , (x_3, y_3) is made of comparable elements. Let $A = \{x, y, z\}$. The sets $\{u \leq x\}$, $\{v \leq y\}$, $\{w \leq z\}$ are disjoint and their union is the $c.h$ -envelope of A : any $u \leq x$, $v \leq y$ cannot be c -comparable; otherwise, as X_3 is a forest and x, y are larger than $u \wedge v = (\min(u_i, v_i))$, x_3 and y_3 would be comparable. On the other hand, if $u' \leq x$, $u'' \leq x$ and u', u'' c -comparable, then $u' \vee u'' \leq x$.

C.5 We have the following partial result:

Proposition. *Let X_1, \dots, X_n be forests and let A be a subset of $X = X_1 \times \dots \times X_n$. Let A_1 be the set of elements $x \in X$ for which there exists x^1, x^2, \dots, x^n in A with $x_j = x_j^j \geq x_j^i$ for all $1 \leq i, j \leq n$. Let A_2 be the set of elements $y \in X$ such that $y \leq x$ for some x in A_1 . Then A_2 contains the c.h-envelope of A .*

Remark. *Example 2 above shows that A_2 can be strictly larger than the c.h-envelope. Example 1 shows that the straightforward generalization of the case $n = 2$ does not work.*

Proof of the Proposition. It is very similar to the proof of the proposition in C.3 above and left to the reader. □

C.6 We will now see how the definition of the transversality relation in Subsection 5.4 is a natural consequence of the proposition above. As observed earlier, an essential difference with the toy model is that the transversality relation is used to construct the classes $\mathcal{R}(I)$. So, let us just try to define the relation for the starting class $\mathcal{R}(I_0)$ associated to the initial horseshoe K . We would have:

$$(C.10) \quad X_1 = \{(P, Q, n) \in \mathcal{R}(I_0), Q \subset Q_u\},$$

$$(C.11) \quad X_2 = \{(P', Q', n') \in \mathcal{R}(I_0), P \subset P_s\},$$

and X_3 is the set of parameter intervals. All sets are partially ordered by inclusion (of the Q 's for X_1 , of the P 's for X_2), and are obviously *trees*, with respective roots $(P_u, Q_u, n_u), (P_s, Q_s, n_s), I_0$.

We start from an intuitive definition of transversality: for $(P, Q, n) \in X_1, (P', Q', n') \in X_2, I \in X_3$, we write

$$Q \hat{m}_I P'$$

if for all $t \in I$ we have

$$(C.12) \quad \delta(Q, P') \geq 2 \max(I, |Q|^{1-\eta}, |P'|^{1-\eta}).$$

(The number η in the exponent is necessary in order to keep the distortions under control.)

The corresponding subset of $X_1 \times X_2 \times X_3$ is

$$(C.13) \quad A = \{(Q, P', I), Q \hat{m}_I P'\}.$$

This set is hereditary but it is not, a priori, concave. The concavity property (Propositions 4 and 7 in Section 6) is very useful in many places. So, we wish to replace A by a larger set which is hereditary and concave. If we apply the recipe of the proposition in C.5, we are led first to define

a set A_1 and then a set A_2 . According to the proposition, A_1 should be the set of (Q, P', I) for which there exist $Q_2, Q_3 \subset Q, P'_1, P'_3 \subset P', I_1, I_2 \subset I$ satisfying

$$(C.14) \quad Q \widehat{\cap}_{I_1} P'_1, \quad Q_2 \widehat{\cap}_{I_2} P', \quad Q_3 \widehat{\cap}_I P'_3.$$

As I_1, I_2 can be chosen arbitrarily small, we can replace them with single values $t_1, t_2 \in I$ of the parameter; the three conditions in (C.14) become:

– there exists $P'_1 \subset P', t_1 \in I$ such that

$$(C.15) \quad \delta(Q, P'_1) \geq 2 \max(|Q|^{1-\eta}, |P'_1|^{1-\eta})$$

– there exists $Q_2 \subset Q, t_2 \in I$ such that

$$(C.16) \quad \delta(Q_2, P') \geq 2 \max(|Q_2|^{1-\eta}, |P'|^{1-\eta})$$

– there exists $Q_3 \subset Q, P'_3 \subset P'$ such that

$$(C.17) \quad \delta(Q_3, P'_3) \geq 2 \max(|I|, |Q_3|^{1-\eta}, |P'_3|^{1-\eta})$$

for all $t \in I$.

As P'_1, Q_2, Q_3, P'_3 may be chosen arbitrarily thin, it is natural to replace (C.15), (C.16), (C.17) by

$$(C.15)' \quad \delta(Q, P'_1) \geq 2|Q|^{1-\eta} \text{ for } t = t_1;$$

$$(C.16)' \quad \delta(Q_2, P') \geq 2|P'|^{1-\eta} \text{ for } t = t_2;$$

$$(C.17)' \quad \delta(Q_3, P'_3) \geq 2|I| \text{ for all } t \in I.$$

Finally, the largest t value of $\delta(Q, P'_1)$ that one can hope for (by choosing $P'_1 \subset P'$ appropriately) is $\delta_R(Q, P')$; similarly, the largest value of $\delta(Q_2, P')$ that one can hope for is $\delta_L(Q, P')$ and the largest value of $\delta(Q_3, P'_3)$ that one can hope for is $\delta_{LR}(Q, P')$. Notice that we need anyway to eliminate P'_1, Q_2, P'_3, Q_3 from the definition because in $\mathcal{R}(I)$ (instead of $\mathcal{R}(I_0)$), elements are constructed inductively and thinner rectangles are constructed at the end. Replacing $\delta(Q, P'_1)$ by $\delta_R(Q, P')$, $\delta(Q_2, P')$ by $\delta_L(Q, P')$ and $\delta(Q_3, P'_3)$ by $\delta_{LR}(Q, P')$, we obtain the three conditions, (T1), (T2), (T3) in Subsection 5.4. This defines $\overline{\mathfrak{h}}$.

The last step is to go from A_1 to A_2 , taking the hereditary envelope of A_1 , which corresponds exactly to the transition from $\overline{\mathfrak{h}}$ to \mathfrak{h} in Subsection 5.4.

References

- [BC] M. Benedicks and L. Carleson. The dynamics of the Hénon map. *Annals of Math.*, 133:73–169, 1991.
- [BDV] C. Bonatti, L. Diaz, and M. Viana. *Dynamics beyond uniform hyperbolicity*. Encyclopaedia of Math. Sciences, 102. Springer, 2004.
- [BR] R. Bowen and D. Ruelle. The ergodic theory of Axiom A flows. *Invent. Math.*, 29:181–202, 1975.
- [C] E. Colli. Infinitely many coexisting strange attractors. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 15:539–579, 1998.
- [MV] L. Mora and M. Viana. Abundance of strange attractors. *Acta Math.*, 171:1–71, 1993.
- [MPV] C. G. Moreira, J. Palis, and M. Viana. Homoclinic tangencies and fractal invariants in arbitrary dimension. *C. R. Acad. Sci. Paris Sér. I Math.*, 333:475–480, 2001.
- [MY] C. G. Moreira and J.-C. Yoccoz. Stable intersections of regular Cantor sets with large Hausdorff dimensions. *Ann. of Math.*, 154:45–96, 2001. Corrigendum in *Annals of Math.* 154:527, 2001.
- [N] S. Newhouse. The abundance of wild hyperbolic sets and nonsmooth stable sets for diffeomorphisms. *Publ. Math. I.H.E.S.*, 50:101–151, 1979.
- [NP] S. Newhouse and J. Palis. Cycles and bifurcation theory. *Astérisque*, 31:44–140, 1976.
- [P1] J. Palis. A global view of Dynamics and a conjecture on the denseness of finitude of attractors. *Astérisque*, 261:335–347, 2000. *Géométrie complexe et systèmes dynamiques (Orsay, 1995)*.
- [P2] J. Palis. A global perspective for non-conservative dynamics. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 22:485–507, 2005.
- [PT] J. Palis and F. Takens. Hyperbolic and the creation of homoclinic orbits. *Annals of Math.*, 125:337–374, 1987.
- [PY1] J. Palis and J.-C. Yoccoz. Homoclinic tangencies for hyperbolic sets of large Hausdorff dimension. *Acta Math.*, 172:91–136, 1994.
- [PY3] J. Palis and J.-C. Yoccoz. Fers à cheval non uniformément hyperboliques engendrés par une bifurcation homocline et densité nulle des attracteurs. *C.R. Acad. Sci. Paris*, 333:867–871, 2001.

- [PY2] J. Palis and J.-C. Yoccoz. Implicit formalism for affine-like map and parabolic composition. *Global Analysis of Dynamical Systems*, pages 67–87, 2001. Institut of Phys., IOP-London.
- [Ru] D. Ruelle. A measure associated with Axiom A attractors. *Amer. J. Math.*, 98:619–654, 1976.
- [Si] Ya. Sinai. Gibbs measures in ergodic theory. *Russian Math. Surveys*, 27:21–69, 1972.

Jacob Palis
Instituto Nacional de Matemática Pura e Aplicada
Est. Dona Castorina, 110 - Jardim Botânico
22460-320, Rio de Janeiro - RJ, Brazil
jpalis@impa.br

Jean-Christophe Yoccoz
Collège de France - 3 Rue d'Ulm
75005 Paris, France
jean-c.yoccoz@college-de-france.fr